

Editor's LETTER

ACS
chemical
biology

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California-Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Timothy P. Clackson
ARIAD Pharmaceuticals, Inc.

Brian T. Chait
Rockefeller University

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas C. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Fostering Major Breakthroughs

The emphasis on interdisciplinary scientific research is inescapable. In 1999, Metzger and Zare (1) referred to interdisciplinary research as the “mantra of scientific policy”. Evidence in support of this descriptor is abundant. The United States National Academy of Sciences, for example, has organized a task force focused on promoting interdisciplinary research. Private and public funding agencies, universities, research institutes, and professional organizations also have invested heavily in interdisciplinary programs. Many of these organizations are seeking to lower the barriers between scientific disciplines. Centers for interdisciplinary research abound; institutions are being restructured; research space is becoming more integrated. As researchers attracted to chemical biology, our interests naturally fall into the category of interdisciplinary science. Still, one wonders if the cross-talk between different disciplines will yield the desired benefits—major breakthroughs.

One line of evidence that interdisciplinary research will move us forward comes from an analysis of the factors resulting in the major discoveries of the past. J. Rogers Hollingsworth, a professor in the Sociology and History Departments at the University of Wisconsin–Madison, has examined the environment in institutions that have consistently produced major breakthroughs. Additionally, he has searched for characteristics common to scientists who have made such breakthroughs. One of his conclusions is that those successful researchers have “high cognitive complexity”; they easily intuit connections between different fields and readily relate ideas from one field to another (2). A surprising number of these successful scientists had integrated multiple cultural identities from diverse ethnicities, nationalities, or religions. It is perhaps this ability to see the world from different perspectives that allowed the researchers to make contributions with wide-ranging implications (3). Like those between different nations, the cultural barriers between different scientific disciplines (e.g., chemistry and biology) can be substantial. Yet, if our goal is to make major contributions to science, the historical data suggest that it is worthwhile to breach these barriers.

The case for increasing communication between scientists trained in different disciplines is likewise supported by Hollingsworth's analysis. Institutions in which many scientists have made major scientific discoveries share some common features (4). One key attribute is an ingrained culture that facilitates the interaction of researchers from different areas and with different perspectives. Mechanisms such as journal clubs, interdisciplinary seminars, and forums for casual and leisurely interaction are examples of how institutions can foster innovation. Such mechanisms can focus talented researchers on the most critical problems and elicit novel solutions. Although the specific means by which institutions facilitate interactions between scientists in different disciplines does not seem to be important, the benefits of removing boundaries between the disciplines are apparent.

As a forum for the publication of interdisciplinary research, *ACS Chemical Biology* seeks to facilitate conversations between chemists and biologists both in print and on the Web. It is in the latter arena that we are generating a new community, an on-line journal club of sorts, that enables the free and immediate exchange of ideas. One manifestation of our plan to stimulate discussion is the on-line feature “Ask the Expert”. This section was designed with the knowledge that one hallmark of interdisciplinary research is that its

practitioners often are using techniques that are new to them. To lower the barrier to entry, scientists can ask experts questions about fundamental issues, research directions, or techniques and methods. We shall feature different topics and encourage the community to join the discussion. Another manifestation of our plan is the “Chemical Biology WIKI”. This venue also allows for discussion of forefront scientific topics of interest to chemical biologists, as well as policy issues that affect the way we obtain funding, devise and conduct research, and educate and mentor those working at the interface.

In designing our journal, we are inspired by another insight from Hollingsworth’s analysis: Institutions that foster major breakthroughs lack inertia (4). Our vision of *ACS Chemical Biology* is that it will be nimble—it will respond to new scientific directions. We are interested in innovative ideas about how we can best serve our readership and the broader scientific community. We welcome your input.

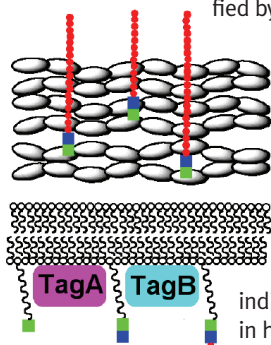


Laura L. Kiessling
Editor-in-Chief

1. Metzger, N.; and Zare, R. N. (1999). Interdisciplinary research: From belief to reality. *Science* 283, 642–643.
2. Hollingsworth, J. R. (in press) *High cognitive complexity and the making of major scientific discoveries. Knowledge, communication and creativity* (Sales, Arnaud, and Fournier, Marcel, Eds.), SAGE Publications, Thousand Oaks, CA.
3. Hollingsworth, J. R.; Hollingsworth, E. J.; and Hage, J. (in press) *The search for excellence: Organizations, institutions, and major discoveries in biomedical science*, Cambridge University Press, Cambridge.
4. Hollingsworth, J. R.; and Hollingsworth, E. J.; (2000) Major discoveries and biomedical research organizations: Perspectives on interdisciplinarity, nurturing leadership, and integrated structure and cultures, in *Practising interdisciplinarity* (Weingart, P., and Stehr, N., Eds.), pp 214–244, University of Toronto Press, Toronto, ON.

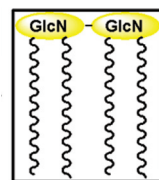
Breaking Down the Wall

In Gram-positive bacteria, three enzymes, TagA, TagB, and TagF, are involved in the synthesis of wall teichoic acids (WTAs), anionic polymer components of the cell wall that are essential for bacterial survival. These three enzymes have traditionally been challenging to study as a result of difficulty in isolating and synthesizing a 55-carbon-long carrier lipid attached to their cell wall precursor substrates. Now, Ginsberg *et al.* (p 25 and Point of View p 14) have developed a chemical and enzymatic approach to generate WTA substrates with a shorter carrier lipid. The authors show that TagA can accept these substrates and generate products that act as substrates for TagB. TagB, in the presence of an appropriate phosphate donor, produces the second expected product, which the authors anticipate will subsequently be a substrate for TagF. These data show that substrate-attached long lipid chains are not necessary for the commitment steps in WTA synthesis, nor is a membrane surface crucial for the reactions. These findings open the door for future structural and molecular studies of these essential enzymes.



Redefining Lipopolysaccharides

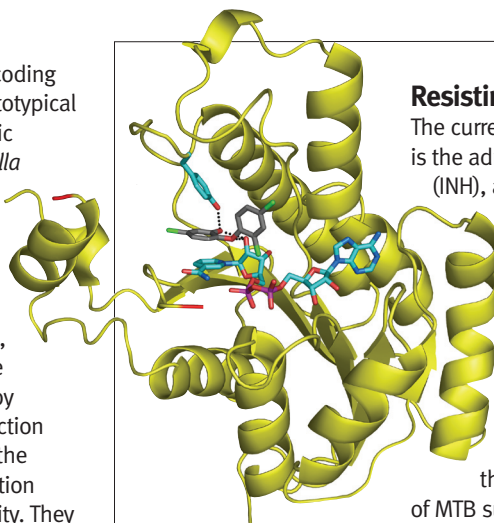
Gram-negative bacteria, unlike their close cousins the Gram-positive bacteria, have an outer membrane (OM) that protects them from the environment. The inner leaflet of the OM contains glycerophospholipids, and the outer leaflet contains lipopolysaccharide (LPS). LPS is also known as endotoxin, a potent proinflammatory molecule in humans associated with septic shock. In bacteria, LPS acts as a permeability barrier for large hydrophobic molecules and polycationic peptides. LPS is divided into three regions: OM-embedded lipid A, an oligosaccharide core, and in some cases an O-specific hydrophilic polysaccharide chain. In *Escherichia coli*, it has long been thought that the minimal LPS structure required for viability contains a lipid A modified



by two 2-keto-3-deoxy-D-manno-octulosonate (Kdo) residues. By using a suppressor strain that lacks Kdo (KPM22), Meredith *et al.* (p 33) redefine the minimal structure of the LPS required for bacterial growth. They show that bacteria lacking Kdo are more susceptible to antibiotics than the wild-type strain and have an OM containing Lipid IV_A, a precursor of LPS lacking glycosylation. These data show that Kdo modification of lipid A is not necessary for the transport of the lipid A backbone to and assembly of the OM. However, the LPS structure is important as a protective barrier. Bacteria containing Lipid IV_A in their outer leaflet do not elicit an immune response, indicating that the LPS structure is also an important component of toxicity in humans. The availability of KPM22 suppressor strain opens the door to further exploration of the role of LPS in bacterial viability and pathogenesis.

Iron Supplements

When iron is limiting, bacteria activate genes coding for iron chelators called siderophores. The prototypical siderophore produced by Gram-negative enteric bacteria such as *Escherichia coli* and *Salmonella typhimurium* is enterobactin (Ent). This small nonribosomal peptide scaffold contains bidentate functional groups to chelate iron scavenged from vertebrate proteins. Some mammalian proteins, such as siderocalin, bind to Ent and inhibit its activity. This binding, in turn, inhibits bacterial growth. Bacteria have found ways around this antimicrobial system by enzymatically tailoring Ent to prevent its interaction with siderocalin. Luo *et al.* (p 29) investigate the effect of two types of modification, C-glycosylation and hydrolysis of the lactone ring, on Ent activity. They find that both types of modification decrease the membrane affinity of Ent. Furthermore, the authors find that the iron acquisition rate is increased in glycosylated Ent. These findings show that bacteria tailor their siderophores to ensure survival in a changing environment.



Resisting Drugs

The current method to treat tuberculosis is the administration of isoniazid (INH), a drug that compromises the cell wall of *Mycobacterium tuberculosis* (MTB), the causative agent of the disease. INH inhibits the enoyl reductase InhA in the type II fatty acid biosynthesis pathway and hinders the biosynthesis of mycolic acids. INH is effective, but the rise of drug-resistant strains of MTB suggest that its effectiveness is short-lived. Using structure-based

drug design Sullivan *et al.* (p 43) synthesize a series of diphenyl ethers that target InhA. These compounds have nanomolar inhibition constants and are effective against both drug-sensitive and drug-resistant strains of MTB. Transcriptional profiling suggests that the improved antibacterial activity of the newly designed inhibitors may result from their ability to evade the detoxification mechanisms triggered by other tuberculosis drugs.

Spotlight

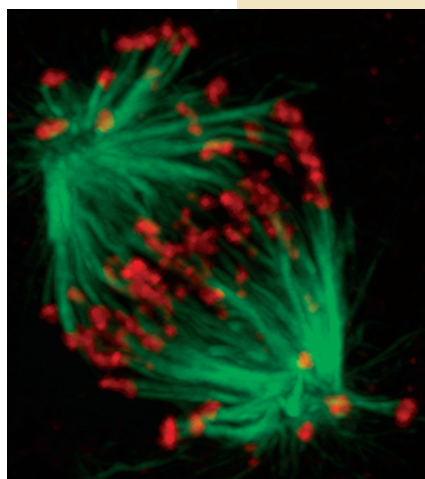


Cannabinoids... Legally

Cannabinoids, such as those found in the *Cannabis* plant, are known to affect cellular function. Endocannabinoids and their receptors are found in neurons, the messenger cells of the central nervous system (CNS), and in microglial cells, the immune cells of the CNS. In microglia, endocannabinoids are believed to protect neurons from inflammatory damage, but the mechanism is not well understood. Eljaschewitsch *et al.* (*Neuron* 2006, 49, 67–79) now elucidate a defined pathway through which the endocannabinoid anandamide (AEA) limits neuronal damage after CNS injury.

As part of their role in immune surveillance, microglia use the mitogen-activated protein kinase (MAPK) pathway to produce inflammatory signals. In healthy brain tissue, AEA release results in induction of the MAPK pathway. In contrast, after primary injury, accumulation of AEA in brain tissue results in protection of neurons from inflammatory damage. To investigate the mechanism of protection, the effects of AEA release on the MAPK pathway in activated microglia were examined. It was discovered that AEA release turns off the MAPK pathway, as evidenced by a decrease in nitric oxide production and a reduction in phosphorylation of extracellular signal-regulated kinase-1/2 (ERK-1/2) and ERK kinase. Furthermore, the authors see an induction of the dephosphorylating enzymes mitogen-activated protein kinase-phosphatases-1 and -2 (MKP-1 and MKP-2) as a direct result of phosphorylation of Histone H3 on the *mkp-1* gene, providing further insight into the mechanism of the protective effect of AEA.

The results suggest that AEA acts as a gatekeeper for signal transduction in microglial cells, imposing a negative feedback loop to control the inflammatory response and neurodegenerative immune reactions after primary brain damage. As the clinical use of cannabinoids or related compounds (including those found in the *Cannabis* plant) for treatment of CNS inflammation and multiple sclerosis are under consideration, these findings represent compelling evidence that the cannabinoid system is a valid target for therapeutic intervention in neuroinflammatory and neurodegenerative disorders. **EG**



Chromosome Congress

During each round of cell division, chromosomes are replicated and remain paired as sister chromatids until later in cell division. The mitotic spindle, a bipolar apparatus in the dividing cell, is responsible for capturing the chromosomes, separating the chromatids, and distributing them into the daughter cells. In preparation for segregation, the chromatids connect to the spindle microtubule bundles (green signal), called kinetochore fibers (K-fibers) that emanate from each spindle at either end of the dividing cell. The captured chromosomes (red signal) move (congress) to the metaphase plate prior to the separation of sister chromatids. If one chromosome is attached through one chromatid to a single K-fiber at one end of the cell, how does the sister chromatid find its K-fiber emanating from the opposite pole? Now Kapoor *et al.* (*Science* 2006, 311, 388–391) use a combination of microscopy, chemical biology, and RNA interference (RNAi) to examine chromosome congression.

They find that chromosomes attached to a K-fiber at one spindle pole slide along neighboring fibers toward the center of the cell. This process brings the unattached sister chromatid in range with a K-fiber from the opposite pole. Using chemical inhibitors that slow mitotic progression, the authors show that this sliding process occurs in ~85% of chromosomes. Combining chemical biology with RNAi methods, the authors determined that CENP-E, a kinesin-7 family microtubule motor, is likely responsible for movement of the unattached chromatid along K-fibers. These data explain why chromosome congression is a cooperative process. As more chromosomes become attached at both spindle

poles, additional K-fiber tracks are generated for the movement of other chromosomes. These results also show that a combination of techniques is required to mechanically dissect complex processes such as mitosis. **EJ**

Neutralizing Neutrophils

The β_2 integrins are cell surface proteins that mediate leukocyte recruitment during the inflammatory response. β_2 integrins are potential drug targets for inflammatory and immune disorders, but their structural and functional complexity has hindered the discovery of effective small molecule inhibitors. Björklund *et al.* (*Biochemistry* 2006, published online 09 February 2006; 10.1021/bi052238b) have developed a novel high throughput screen for the β_2 integrins and have discovered a new class of small molecule inhibitors.

Exploiting the sensitivity and efficiency of phage display technology, the researchers developed a competition assay to identify small molecules that displace phage displaying a known 18-mer peptide, abbreviated DDGW, from the binding domain, or I domain, of the $\alpha_m\beta_2$ integrin. The assay design overcomes the limitations of comparable assays that may lack sensitivity, require labeling reagents that can affect peptide activity, or have undesired avidity effects. The screen yielded several compounds containing a 2-thioxothiazolidin-4-one substructure as specific inhibitors of the interaction between the



DDGW-phage and the I domain. IMB-10 was identified as the most active compound, and additional experiments revealed that this compound stabilized the binding of the I domain to its endogenous ligands, proMMP-9

and fibrinogen. Moreover, molecular modeling and other studies using mutant I domains indicated that IMB-10 shifts the equilibrium of the I domain structure toward the active conformation. Remarkably, leukemia cell migration *in vitro* and leukocyte recruitment *in vivo* were both potently and selectively inhibited by IMB-10, suggesting that stabilization of integrin ligand binding is a viable approach for development of anti-inflammatory agents.

The novel mechanism through which IMB-10 asserts its effects could have advantages over integrin inhibitors that prevent ligand binding. Targeting activated integrins could confer selectivity toward cells already triggered by an inflammatory signal, and upregulation of other integrins to compensate for loss of function is not a concern. These inhibitors could lead to exciting new drugs for immune and inflammatory diseases. **EG**

New Diagnostics for Tiny Motors

An abundance of ATP-dependent enzymes catalyze DNA and RNA rearrangements in the cell. These motors and remodelers play critical roles at all stages of gene replication and expression. The mechanisms coupling ATP hydrolysis with nucleic acid acrobatics have been of interest for decades, yet appropriate assays to view these enzymes in action have remained a barrier. Now, two such enzymes, a bacterial DNA gyrase and a viral RNA helicase, have been tracked at the single molecule level in unique high

resolution assays. In both cases, the experimental design allows direct observation of a nucleic acid substrate during the enzyme's catalytic cycle.

DNA gyrase introduces negative supercoils into DNA. These supercoils are essential because they compact the bacterial chromosome and promote any reaction that involves an untwisting of DNA. There are two basic movements DNA is capable of, bending and twisting. Bending has been studied for a long time, but twisting is more challeng-

ing to examine. To observe this activity, Gore *et al.* (*Nature* 2006, 439, 100–104) engineered molecular tweezers to pull a DNA substrate containing a magnetic bead taut by a magnet. The center of the DNA helix contained a nick to facilitate rotation and a fluorescent bead to visualize rotation. When DNA wraps around the enzyme to form a complex poised for catalysis, approximately one rotation of the DNA is observed. With ATP present, a catalytic cycle causes two rotations of the DNA. This setup

allowed the measurement of a variable rarely accessible to enzymologists, tension. As the DNA tension was increased in tiny increments, both the gyrase initiation rate and the ability to catalyze multiple rotations without dissociation dropped markedly. Interestingly, as DNA tension was increased, the velocity of gyrase catalysis was unchanged. Using this assay, the authors detect two pauses corresponding to two kinetic steps in the gyrase cycle and show that the

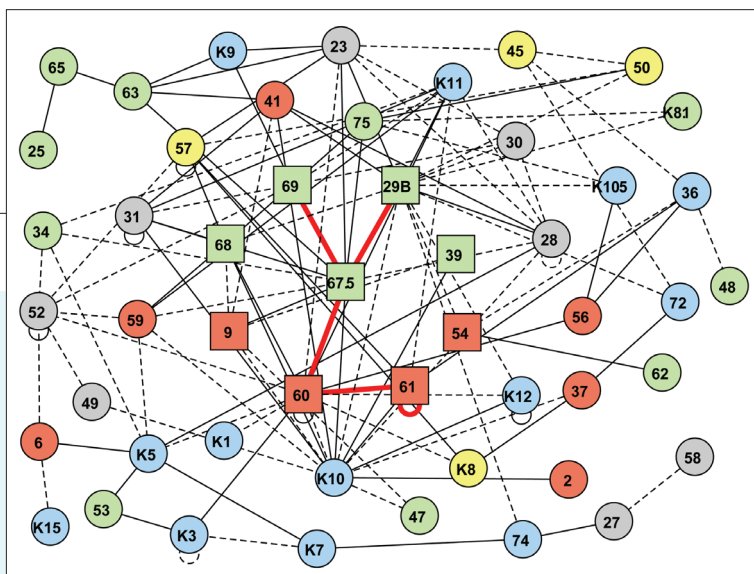
(continued on page 6)

Untangling Webs

Efforts to investigate and treat viral infections have increased our understanding of interactions between viral and host proteins. However, interactions among intraviral proteins, including those in the herpesvirus family, have not yet been well-characterized.

Uetz *et al.* (*Science* 2006, 311, 239–242) have now generated genome-wide intraviral protein interaction maps for Kaposi's sarcoma-associated herpesvirus (KSHV), linked to Kaposi's sarcoma and B-cell lymphomas, and varicella-zoster virus (VSV), a herpesvirus associated with chickenpox and shingles. These data provide insights into the properties of herpesviruses and their interaction with the human protein interactome.

Using yeast two-hybrid technology, the researchers identified 123 and 173 intraviral protein pairs in KSHV and VSV, respectively. They observed that, in contrast to cellular networks which typically exhibit the properties of scale-free networks, where most proteins have few interaction partners but a few have many, viral protein networks emerge as single, highly coupled modules. Further examination of the KSHV interactome coupled with data available from orthologous proteins in other herpesviruses allowed the researchers to predict 114



orthologous intraviral protein interactions in four different herpesviruses.

The researchers connected the KSHV interactome with a prototypical human protein interaction network using 20 predicted interactions between 8 KSHV and 20 human proteins. Notably, when the viral and human interactomes were docked together, the topology of the KSHV network changed from a highly coupled module to a scale-free network of interacting submodules, indicating that the combined virus–host network takes on host network properties. The researchers hypothesize that although the viral and human interactomes have distinct network topologies in their isolated states infection may result in the emergence of new system properties that embody specific features of viral pathogenesis. Further insight into these interactions will enhance understanding of viral mechanisms and may lead to new strategies for treatment of viral infections. **EG**

New Diagnostics for Tiny Motors, *continued*

rate-limiting step of the super-coiling reaction is at the end of the cycle.

Optical tweezers were used to watch a helicase unwind a duplex RNA at high resolution. Dumont *et al.* (*Nature* 2006, 439, 105–108) mounted a hairpin of double-stranded RNA between beads such that unwinding of the RNA increases the bead-to-bead distance. In this case, the test subject was NS3, an ATP-dependent 3' to 5' helicase critical for hepatitis C viral replication.

Upon addition of NS3 and saturating ATP, the RNA duplex was unwound in bursts and pauses like an inchworm rather than a steady train along RNA tracks. An average of 11 base pairs were unwound by NS3 before a pause. By varying ATP, the authors show that exit from a pause is a two-step kinetic mechanism and, as with gyrase, only one step requires ATP binding. At low ATP concentrations, “substeps” separated by subpauses appeared, and these indicated that unwinding

of 11 base pairs occurs in three distinct events. The authors propose a dual-function model for NS3 helicase. In the model, a “translocator” activity moves in 11 base pair steps and monitors double-stranded RNA directly ahead of the unwinding activity. The second activity or “helix opener” moves in 3–5 base pair steps to disrupt the helix. The data indicate that ATP is critical for coordinating both of these inchworming functions.

With both gyrase and NS3, changing the force on the

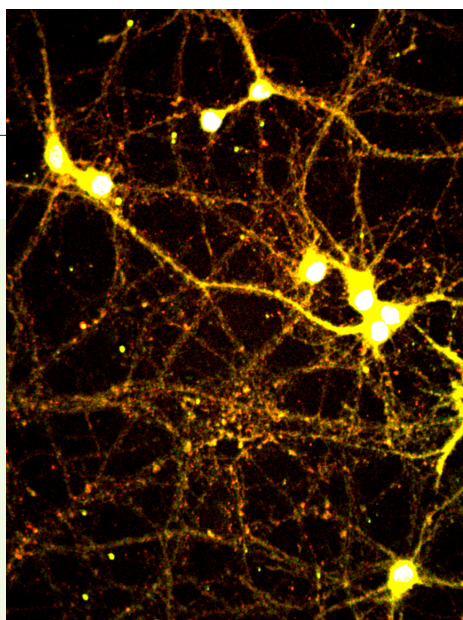
nucleic acid substrate alters some aspects of enzyme activity. In a cell, the force on a DNA or RNA is influenced by proteins which bind, compact or copy the strands. These regulatory contributions will be an interesting direction for future research now that the study of such ATP-dependent molecular machines has such a high resolution. These assays also enable careful mutant studies and dissection of the mysterious role of ATP in such cellular processes. **JU**

Looking at RNA Interactions

RNA binding proteins (RBPs) are essential components of many cellular machines, including those that control pre-messenger RNA (mRNA) splicing, mRNA editing, mRNA transport, translation regulation, and RNA

degradation. Understanding how and when these proteins interact with their target RNA is crucial to studying gene regulation. Few methodologies exist for identifying all of the RBPs that bind to one mRNA of interest. Zielinski *et al.* (*PNAS* 2006, 103, 1557–1562) have now developed a peptide–nucleic acid (PNA)-assisted technique to identify proteins that interact with a specific RNA *in vivo*.

The authors coupled a PNA, a nucleic acid analog in which the sugar-phosphate backbone is replaced with a polyamide backbone, to a cell-penetrating peptide called transportin 10 (TP10), and a photoactivatable amino acid adduct *p*-benzoylphenylalanine (Bpa). The authors chose to apply their new PNA-assisted identification



tool to find RBPs that bind to *ankylosis* RNA, a dendrically localized mRNA coding for a pyrophosphate transporter. After transport into the neuronal cells, the covalent bond between TP10 and the PNA was reduced and the PNA released to hybridize with the target mRNA (red

signal). UV irradiation of the cells activated Bpa and generated a free phenylalanine radical able to crosslink the nearest protein. These PNA-mRNA-protein complexes were isolated and the RBPs identified using mass spectrometry. The authors also examined the effect of external stimuli on these complexes and found that RNA-protein complexes (yellow signal) are remodeled in response to a physiological change, a result consistent with the current thought that RNA-protein interactions are dynamic. This new *in vivo* methodology allows scientists to quantify these changes in RNA-protein interactions which in turn will expand our understanding of RBP-mediated gene regulation. **EJ**

UPCOMING CONFERENCES

HUPO-USA

Human Proteome Organization
March 12–15, 2006
Boston, MA

ACS Spring National Meeting

American Chemical Society
March 26–30, 2006
Atlanta, GA

ASBMB Annual Meeting and Centennial Celebration

Amer. Soc. Biochem.
Mol. Biology
April 1–5, 2006
San Francisco, CA

97th Annual Meeting

American Association for
Cancer Research
April 1–5, 2006
Washington, DC

Spotlights written by Eva Gordon, Evelyn Jabri, and Jason Underwood.

The Japanese Society for Chemical Biology

Hiroyuki Osada*

Discovery Research Institute, RIKEN, Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan

Traditionally, the fusion of chemistry and biology research has prospered in Japan. For example, many years ago, there was a debate on the cause of beriberi; one side supported a pathogenic fungi theory and the other, a nutrition deficiency theory. When Umetaro Suzuki discovered oryzanin in rice bran, however, the altercation was terminated. Oryzanin is now known as vitamin B1, and its discovery opened the frontier of vitaminology. Another scientific achievement that occurred in Japan was the isolation of gibberellin from a fungal strain which caused rice bakanae disease (abnormal elongation of rice seedlings).

This research was the origin of later phytohormone study.

Natural product chemistry, including the identification of the chemicals which govern important biological phenomena, was traditionally a Japanese specialty. Although many Japanese scientists were involved in natural product research, they had few opportunities where they could all come together and discuss their work. This was because Japan's researchers were divided into different societies such as The Chemical Society of Japan, The Japanese Biochemical Society, The Pharmaceutical Society of Japan, Japan Society for Bioscience, Biotechnology, and Agrochemistry. The annual meetings of these societies were invariably held on almost the same days, so researchers had few opportunities to gather and exchange information.

One exception, the Symposium on the Chemistry of Natural Products, has a nearly 50-year history and continues to play an important role in giving a common platform to Japanese organic chemists. This symposium, however, focuses on chemistry more than biology.

With the development of molecular biology, it has become necessary to create a new platform dealing equally with chemistry and biology to reflect current research progress. The launching of new journals in chemical biology has inspired us to organize a new society for discussing and exchanging information on Chemical Biology. In response to an appeal by M. Hagiwara (Tokyo Medical and Dental University), researchers belonging to the University of Tokyo, the Tokyo Institute of Technology, RIKEN, and other such institutions gathered together to form the Japanese Society for Chemical Biology (<http://www.tmd.ac.jp/jcb>) in May 2005. Nagano (University of Tokyo) has graciously agreed to serve as the chairman of the Society.

The first annual meeting will be held May 8–9, 2006, in Tokyo, Japan. More than 1000 researchers with different expertise will participate in the meeting, and editors of *ACS Chemical Biology* and *Nature Chemical Biology* will be invited to the international symposium. We expect that the new society will bring like-minded scientists together to foster communication between the chemists and biologists and expand collaborations between academia and industry. We look forward to an exciting meeting and welcome your participation in this new Japanese endeavor.



Photo courtesy of Andrew Feltz

*To whom correspondence should be addressed.

E-mail: hisyo@riken.jp.

Published online February 17, 2006

10.1021/cb0600039

© 2006 by American Chemical Society

Chemical Biology and the NIH

John M. Schwab^{†,‡,*}, and Jeremy M. Berg[‡]

[†]Division of Pharmacology, Physiology, and Biological Chemistry, [‡]National Institute of General Medical Sciences, National Institutes of Health, 45 Center Drive, Bethesda, Maryland 20892

In recent years it has become increasingly clear to NIH staff that many of the most exciting opportunities for biomedical science are to be found along the borders that define the traditional scientific disciplines. Chemical biology, a recently forged specialty that applies logic and methods from chemistry to the study of complex and challenging questions in biology, captures the essence of this integrated approach to scientific discovery. As biomedical scientists as well as long-time members of the American Chemical Society, we are pleased that the ACS has recognized this important trend by launching this promising new journal and by endowing it with first-class editorial leadership.

At the invitation of the editors, in future issues of *ACS Chemical Biology*, we will address certain topics of great mutual interest to chemical biologists and NIH. These will include grant funding trends, opportunities for collaborative science, and mentoring. We welcome your suggestions for additional topics and look forward to continued, productive interactions between NIH and the chemical biology community.



Courtesy of National Institutes of Health.

There are many non-research related topics that are of interest to chemical biologists. In this section of *ACS Chemical Biology* we will focus on these topics and provide unique perspectives from policy makers, funding agencies, and you the research scientists. We welcome your suggestions for topics we should discuss. Please email your ideas to the executive editor at chembiol@acs.org.

*To whom correspondence should be addressed.
E-mail: schwabj@nigms.nih.gov

Published online February 17, 2006

10.1021/cb0600052

© 2006 by American Chemical Society

Gateways to Collaboration

Profiles provide insights into the lives, backgrounds, career paths, and futures of scientists who serve as Experts on *ACS Chemical Biology's* online Ask the Expert feature. Readers are encouraged to submit questions to the Experts at www.acschemicalbiology.org. The editors will post the most interesting exchanges on the website.

As the name of this new journal implies, science is becoming more and more interdisciplinary. Researchers from what were once separate fields are increasingly combining their knowledge and talents to tackle problems that those from any one discipline would be hard-pressed to achieve on their own. Three investigators from the University of California at Berkeley epitomize this ever more popular style of collaboration: biologists Ehud Isacoff and Richard Kramer, and chemist Dirk Trauner.

For the past 5 years, these researchers have mingled their unique ideas and diverse skill sets, culminating most recently in an article on a redesigned glutamate receptor that responds to light (1).

Ehud Isacoff

Born in 1959 in Darmstadt, Germany, Isacoff initially thought that the idea of entering a scientific field had been solely his own. In hindsight, however, he realized that he was likely influenced by his father, who had been a high school electronics teacher in Israel. "As in previous generations, it's a case of keeping to the family business. My father studied electronics of wires, and I study electronics of brains," says Isacoff.

He remembers first becoming interested in neuroscience as a student at McGill University in Montreal, Canada. In his third year at the school, Isacoff enrolled in a lab course in which he took neural recordings of *Acatina*, a terrestrial slug whose cousin *Aplysia* is frequently used as a model organism by neuroscientists. Against instructions issued by the class teaching assistant, he and a few of his classmates used up reams of expensive chart paper one evening recording the slugs' neural responses. The next morning, Isacoff recalls, "I woke up in the early hours, and

when I opened my eyes, I saw the chart paper grid on the walls. No matter how much I rubbed my eyes, it was still there."

"Neuroscience had been indelibly imprinted on my brain," he adds.

Isacoff continued on to a Ph.D. program at McGill under the mentorship of neuroscientist Richard Birks. Birks had been working for years measuring the secretion of the neurotransmitter acetylcholine that occurs during the "fight or flight" response. He used a smoke-barrel kymograph, an instrument now seen only in museums, to measure the acetylcholine that leached out of neurons and into blood vessels by examining how the neurotransmitter lowered blood pressure in lab animals. Isacoff extended this work by determining how changes in acetylcholine altered the electrical activity of responding neurons.

Isacoff remembers the six years he spent working in Birks' lab as an idyllic time, full of hard work and regular periods of pursuing nonscientific interests. "I had the world's best Ph.D. experience," he says.

However, there were some drawbacks. He felt some frustration with the indirect methods he was using to measure what was taking place in the neurons that were releasing acetylcholine. That led Isacoff to take a different direction for his future research. "What I really wanted to start doing was working at the molecular level," he says.

He picked up some new molecular biology skills at a Cold Spring Harbor short course on *Drosophila* genetics two summers before he completed his doctoral degree in 1988. One of his instructors there was Lily Jan, a well-known neurobiologist at University of California at San Francisco. Impressed with Isacoff's interest and initiative, Jan invited him to join her lab as a post-doctoral fellow. Jan and her colleagues had recently cloned the first potassium channel,



Photo courtesy of Ehud Isacoff

an ion channel present in all neurons and many other cells throughout the body. Isacoff immediately chose two related projects on which to focus: determining whether the channel is made of more than one subunit, and exploring the mechanisms for how the channel's gates open and close.

By the time he finished his fellowship four years later, Isacoff had completed both projects. He accepted a job at University of California at Berkeley, where he started his own lab in February 1993.

Isacoff has continued to work on a number of different projects with the help of graduate students, postdoctoral fellows and collaborators. For example, he and then postdoctoral fellow Lidia Mannuzzu and Lawrence Berkeley National Laboratory staff scientist Mario Moronne developed a novel way to detect the movements within ion channels when they open and close. The three researchers monitored the movement of various amino acids in the channel protein by tagging them with fluorophores, such as rhodamine. By electrically stimulating the membranes of cells, which caused the ion channels to respond, Isacoff and his colleagues watched to see how the fluorescence of the probes changed. Their work provided the first real-time measure of protein motion in the channel's voltage sensor (2).

Recently, Isacoff's team devised a new way to investigate how a synapse forms between two neurons (3). He and his colleagues worked with Peter Scheiffele, a former postdoctoral fellow in a neighboring lab at Berkeley, who now runs his own lab at Columbia University in New York. The researchers wanted to identify the minimum protein signal during synapse formation that dendrites, the projections that conduct electricity into one neuron, send to axons, the projections that conduct electricity away from another neuron. Understanding what forms a connection between compatible dendrites and axons is tricky because both cells contain a host of proteins, says Isacoff,

any of which could have been the source of an unknown signal.

After Scheiffele identified a candidate signaling protein in dendrites, the researchers together purified the protein and incorporated it into an artificial bilayer. This protein-lipid complex was coated onto glass beads, which were then applied to axons. "It's incredible," says Isacoff. "The axon thinks it's just been contacted by a dendrite and develops a transmitter release site. It's not aware that it's been contacted only by a bead."

Isacoff continues to advance this project and others, including those on which he collaborates with Kramer and Trauner.

Richard Kramer

Born in 1956, Kramer grew up near New York City. He remembers a favorite teacher turning him on to biology in a



Photo courtesy of Richard Kramer

science class. The class encouraged his natural love for animals to grow, which in turn inspired the rest of his career.

"I came at this job from being interested in animal behavior, then how the nervous system controls behavior, then working my way down to getting interested in the nuts-and-bolts mechanisms of how neurons work," he says.

He began trying to understand these mechanisms as an undergraduate at the State University of New York in Albany. There, like Isacoff, Kramer began studying slugs. With the cast-off projects and equipment of a graduate student who was headed to medical school, Kramer spent a summer between his junior and senior year in a lab recording electrical feedback from *Aplysia's* neurons. "It was a unique opportunity for an undergraduate to be doing independent research," he says.

He also remembers the experience as lending insight into the academic lifestyle, which he was keen to join. "The professor I was working for would have extravagant

parties at his house now and then, and he traveled around the world. It seemed like an attractive lifestyle," Kramer notes.

When he finished his undergraduate degree in 1978, Kramer was intent on heading to the West Coast for graduate school. He applied, and was accepted, to the University of California at Berkeley. There, he worked under the mentorship of Robert Zucker, a neurophysiologist who also studied *Aplysia*. But Kramer had his own project in mind. He became interested in nerve cells that act as oscillators, firing rapidly for short bursts, then pausing before beginning another firing cycle.

"If you took apart your watch, you could figure out what makes the hand go around at constant intervals. But how do the cells do it? What makes them keep time?" Kramer wondered. He set out to answer this question using cultured neurons from *Aplysia*. He found that the neurons operated in a negative feedback loop that involves calcium acting on ion channels in cells (4, 5).

At the time, he says, few researchers were directly studying how ion channels were controlled. The techniques for measuring the behavior of single ion channels had just been invented, and not yet applied to real living neurons. Ion channels were "much more of a vague concept as opposed to an understood entity," he adds. Instead, scientists spoke mainly of "ion currents" because the only way to visualize a channel at the time was to monitor the flow of ions that passed through it.

Nonetheless, the project was his first intensive exposure to studying the mechanisms behind ion channels. By the time Kramer completed his Ph.D., he was intent on studying ion channels further. For his postdoctoral fellowship, he headed to Brandeis University in Waltham, Mass., where neuroscientist Irwin Levitan was taking some of the first recordings of the activity of how single ion channels were controlled by post-translational modifications such as phosphorylation.

“If you express these sensors in appropriate cells, in a way we’re creating artificial senses,” Trauner elaborates.

Soon after arriving at Levitan’s lab, Kramer read a paper showing evidence that olfactory neurons were directly activated by cyclic adenosine monophosphate (cAMP), an important biological second messenger. The finding represented a new mechanism for neuronal response.

Kramer wondered whether the membranes of olfactory neurons, which have ion channels that respond to cAMP, could function as sensors for cAMP levels in other cells. To test this hypothesis, he devised a new technique that involved taking a portion of olfactory neuron membrane and inserting it into other cell membranes. Once successful, he named the new technique “patch cramming,” a play on the “patch clamp” technique that neuroscientists use to stimulate and measure electrical responses from individual neurons (6).

After more than three years at Levitan’s lab, Kramer took a second postdoctoral fellowship in the lab of Steve Siegelbaum at Columbia University. He was soon invited to work with Richard Axel, one of Siegelbaum’s colleagues, to clone the ion channel in olfactory neurons that responds to cAMP. On successfully completing this project, Kramer started his own lab at the University of Miami School of Medicine.

There, he became interested in photoreceptor cells, the rod and cone cells in the eye that gather information on an image. Ion channels that respond to cyclic guanosine monophosphate (cGMP) also played a fundamental role in generating the photoreceptor’s response to light. Kramer wondered how these ion channels would respond to chemically dimerized cGMP. He collaborated with chemist Jeff Karpen to synthesize the new molecule.

“Individual cGMP molecules are diffusing around like flies, and the probability of them landing on their target is low,” he says. “But if you tie two flies together, when one binds to its target, the other is held in close proximity. The probability of finding a neighboring target is much higher.”

Kramer found that the dimeric cGMP stimulates ion channels thousands of times more potently than that of single cGMP molecule. The study, published in 1998 (7), was his first foray into chemical biology.

In 2000, Kramer took his knowledge of ion channels and the retina to start a new lab at the University of California at Berkeley. He continues to study how light information is received and processed in the retina, along with his continuing collaboration with Isacoff and Trauner.

Dirk Trauner

Trauner was born in Linz, Austria, in 1967. “I wasn’t one of those kids who grew up with a chemistry set,” he recalls.

Instead, Trauner says, he became interested in science through a biology class in high school. By the time he began earning his undergraduate degree at the University of Vienna, he had decided to study genetics. The field was a “hot topic,” Trauner says, when he was applying to schools in the late 1980s.

However, his interests abruptly shifted after taking an organic chemistry class during his sophomore year. He tried his hand at total chemical synthesis for the first time, making a natural product called ferulic acid that’s found in many plants. “At that point, I knew I wanted to switch to chemistry,” he remembers. He turned his studies first to biochemistry, then later to synthetic organic chemistry.

In 1994, Trauner completed his undergraduate degree and moved to Berlin, Germany, to begin a Ph.D. program under the mentorship of synthetic chemist Johann Mulzer. His thesis centered on the total synthesis of the opioid analgesic drug morphine. His aim was not to derive a method to produce more morphine, but to demonstrate novel chemical reactions that are important in crafting the drug from scratch (8). Working on the project was

Trauner’s first taste of studying a chemical important to neuroscience, a topic he had been interested in since dissecting human brains as an undergraduate.

After four years and two moves with Mulzer, first to Frankfurt, Germany, and then to Vienna, Trauner completed his doctoral degree. He chose to head to the United States for his postdoctoral fellowship. In 1998, he began working with bioorganic chemist Samuel J. Danishefsky at Memorial Sloan Kettering Hospital in New York.

There, he derived a method to synthesize halichlorine, an alkaloid molecule that interferes with communication between cells and has anti-cancer properties (9). Previously, the molecule was derived only from sponges, a method that’s an expensive and inefficient method, says Trauner. Thus, he adds, his successful synthesis “attracted attention from many labs. Halichlorin was a celebrity target for total synthesis in those days.”

At the end of his postdoctoral fellowship, Trauner accepted a faculty position at the University of California at Berkeley. “An offer at Berkeley you can’t refuse,” he quips.

In July of 2000, he moved to California and set up his lab. He took with him some ideas for performing the total synthesis of several organic compounds involved in immunology and cancer research. However, he says, he became intrigued in ion channels when their structures were first elucidated in the late 1990s. “Synthetic chemists always become interested when a structure is known,” he says. “We get really excited because structures are three-dimensional objects that you can rotate, think about their function, and contemplate ‘how can one reengineer this?’ As a chemist, you think that there should be opportunities to do something to soup up these molecular machines and manipulate them.”



Photo courtesy of Dirk Trauner

The Collaboration

Based on Trauner's interest in ion channels and his skill in synthesizing organic molecules, Isacoff and Kramer approached him soon after he arrived at Berkeley to propose combining their efforts. The three eventually hit on an ideal way to showcase all of their talents: designing an ion channel that responds simply to light, something no ion channel in animals does. "We all came up with this idea at the same time of trying to use particular kinds of chemical elements that change their shape when you shine light on them as a way of inducing functional change in an ion channel," says Isacoff.

To prove that such engineering would be possible, the researchers agreed that their first proof of principle experiment would focus on the potassium channel, a geometrically simple channel that's present on cells throughout the body. Inserting a specific ligand, or plug, inhibits the flow of ions through the doughnut-shaped channel.

Isacoff, Kramer, and Trauner devised a plan to control the plug's placement: Trauner designed an azobenzene molecule, which lengthens in long wavelengths of light but contracts with shorter wavelengths. He affixed one end of the molecule to the ligand and another to the channel.

"The azobenzene is like a fishing line that can be spooled back in or sent out with a bait. What's swallowing the bait is what it's attached to: the ion channel," says Kramer. "You've tethered the ligand onto the target it belongs to."

After the design was completed and fitted onto live neurons, Isacoff and Kramer tested it in their own labs, with Isacoff taking on the majority of the molecular biology work and Kramer taking on the bulk of the electrical recordings. In the end, the project was successful (10).

Ready for a fresh challenge, the researchers transferred their design to a new channel: one controlled by the neurotransmitter glutamate. Unlike the simple round

pore of a potassium channel, the three scientists needed a design compatible with the clam-shell shape of the glutamate receptor. Trauner created a string of molecules with azobenzene at its center to act as a tether tying glutamate to its receptor.

"The question became, can the design strategies that worked with the potassium channel be transferred to something more complex geometrically in the glutamate receptor? Yes, it works—and it's very satisfying," says Isacoff. He, Kramer, Trauner, and their colleagues published their results recently (1).

The three researchers have many ideas of what to pursue next. Trauner says that he and his collaborators are interested in working on new designs for ion channels that respond to other types of stimuli. "It's potentially limitless to think of signals to which ion channels could respond: light, magnetic fields, chemicals, or changes in temperatures," he elaborates. "If you express these sensors in appropriate cells, in a way we're creating artificial senses."

Kramer suggests that the new technique that he, Isacoff, and Trauner have created could make an ideal way to explore or enhance the senses that already exist. For example, the ability to stimulate individual light-sensitive neurons may aid neuroscientists in understanding how neurons are wired throughout the brain. It could also give researchers the means to replace light-sensitive cells that no longer work, such as those destroyed by macular degeneration, a leading cause of blindness.

Isacoff, Kramer, and Trauner continue to meet with each other and the graduate students of all three labs at least once a week to compare notes and bounce ideas around. Rather than shuttling compounds and cells between each other's labs, Trauner says, the three researchers simply send each other's students across Berkeley's small campus.

Though their collaboration is usually unhindered, Kramer notes that the three

have had their share of rough patches. "Like any ménage à trois, things get a little complicated and don't always run so smoothly," he says. "Since there's three of us, there are more ideas that we can possibly do, so we have to prioritize and wrestle with one another to see who is going to do what."

"It's sometimes heated," agrees Isacoff. But with the number of ideas and successful accomplishments growing between the three scientists, he says, "It's been a marvelous collaboration that will continue for some time."

"It's the classical example of a meaningful collaboration," adds Trauner. "We're bringing together so many techniques that are impossible to master in one lab. [Isacoff and Kramer] have learned a bit of chemistry, and I've learned a bit of biology. We're bridging the huge cultural divide between chemists and biologists."

—Christen Brownlee, Science Writer

REFERENCES

1. Volgraf, M., Gorostiza, P., Numano, R., Kramer, R.H., Isacoff, E.Y., and Trauner D. (2006) Allosteric control of an ionotropic glutamate receptor with an optical switch, *Nat Chem Biol.* 1, 47–52.
2. Mannuzzu, L.M., Moronne, M.M., and Isacoff, E.Y. (1996) Direct physical measure of conformational rearrangement underlying potassium channel gating, *Science* 271, 213–216.
3. Dean, C., Scholl, F.G., Choih, J., DeMaria, S., Berger, J., Isacoff, E., Scheiffele, P. (2003) Neurexin mediates the assembly of presynaptic terminals, *Nat Neurosci.* 6, 708–716.
4. Kramer, R.H. and Zucker, R.S. (1985) Calcium-dependent inward current in *Aplysia* bursting pacemaker neurons, *J Physiol.* 362, 107–130.
5. Kramer, R.H. and Zucker, R.S. (1985) Calcium-induced inactivation of calcium current causes the inter-burst hyperpolarization of *Aplysia* bursting neurons, *J Physiol.* 362, 131–160.
6. Kramer, R.H. (1990) Patch cramming: monitoring intracellular messengers in intact cells with membrane patches containing detector ion channels, *Neuron.* 4, 335–341.
7. Kramer, R.H. and Karpen, J.W. (1998) Spanning binding sites on allosteric proteins with polymer-linked ligand dimers, *Nature.* 395, 710–713.
8. Mulzer, J., Dümer, G., and Trauner, D. (1996) Formal Total Synthesis of (-)-Morphine by Cuprate Conjugate Addition, *Angew. Chem. Int. Ed. Engl.* 35, 2830–2832.
9. Trauner, D., Schwarz, J.B., and Danishefsky, S.J. (1999) Total Synthesis of (+)-Halichlorine, an Inhibitor of VCAM-1 Expression, *Angew. Chem. Int. Ed. Engl.* 38, 3542–3545.
10. Banghart, M., Borges, K., Isacoff, E., Trauner, D., and Kramer, R.H. (2004) Light-activated ion channels for remote control of neuronal firing, *Nat Neurosci.* 7, 1381–1386.

Building a Bridge to New Antibiotics

Nicola L. Pohl*

Department of Chemistry and the Plant Sciences Institute, 2756 Gilman Hall, Iowa State University, Ames, Iowa 50011

ABSTRACT Gram-positive bacteria modify their peptidoglycan layers with teichoic acid polymers via a highly conserved disaccharide bridge. Inhibition of the biosynthesis of this bridge is a potential antibiotic strategy that can be explored now that purified versions of the enzymes TagA and TagB and their substrates are accessible for the first time.

Gram-positive and Gram-negative bacteria have substantial differences in the structures surrounding the cells (Figure 1). However, both types of bacteria have a strong, protective peptidoglycan layer. Many powerful antibiotics inhibit the biosynthesis of this protective layer and thereby weaken or destroy the bacterial protective coat. For example, penicillin and vancomycin both inhibit the biosynthesis of the peptidoglycan outer layer of Gram-positive bacteria. Unfortunately, bacterial strains resistant to these antibiotics are becoming ever more common (1). Many of these Gram-positive bacteria modify their peptidoglycan layers with anionic wall teichoic acid (WTA) polymers (Figure 2), which are also crucial for survival of the organism. Ideally, compounds that specifically target the biosynthesis of these anionic polymers could be developed as new antibiotics. However, the biochemical functions of the enzymes involved in WTA biosynthesis have been difficult to verify because of a lack of availability of the complex substrates required. On page 25 of this issue (2), the first chemical and enzymatic syntheses of key substrates for the first two committed steps in WTA biosynthesis are reported along with characterization of the two enzymes, TagA and TagB, that carry out this chemistry. Access to these critical proteins and their substrates now finally opens the door to structural, enzymatic, and inhibitor studies to explore the exciting possibility of new antibiotics that inhibit a different stage of bacterial cell wall biosynthesis and thereby treat currently resistant strains.

Teichoic acid polymers are linked to bacterial peptidoglycans via a disaccharide bridge that is made of *N*-acetylmannosamine linked to *N*-acetylglucosamine (Figure 2) (3). Although the polymers themselves can be modified with alanine or glucose and be made from glycerol phosphate and ribitol phosphate monomers, the disaccharide bridge is highly conserved. This bridge is biosynthesized by the enzyme TagA from UDP-*N*-acetylmannosamine and undecaprenyl-diphospho-*N*-acetylglucosamine building blocks. TagB then forms a dimer or trimer of glycerol phosphate ester bonds attached to the mannosamine core before the polymer chain is extended by Tag F using CDP-glycerol made by TagD; the polymer can also then be modified by additional proteins that add glucose or alanine, for example (3). The final cell wall structures trade the lipid-linked disaccharide core for peptidoglycan (Figure 2). Originally, teichoic acid polymers were thought to be dispensable and therefore not a viable antibiotic target. However, recently these polymers have been shown to be crucial for the survival of *Bacillus subtilis*, a model Gram-positive organism, even under phosphate-limiting conditions in which bacteria can switch to synthesizing phosphate-free anionic teichuronic acid polymers (4, 5).

The biochemical functions of the enzymes involved in WTA biosynthesis have not been verified previously because of a lack of availability of the complex substrates. The 55 carbon long "carrier" lipid attached to many of the cell wall precursor substrates make substrate isolation and

*To whom correspondence should be addressed.
E-mail: npohl@iastate.edu.

Published online February 17, 2006

10.1021/cb0600003 CCC: \$33.50

© 2006 by American Chemical Society

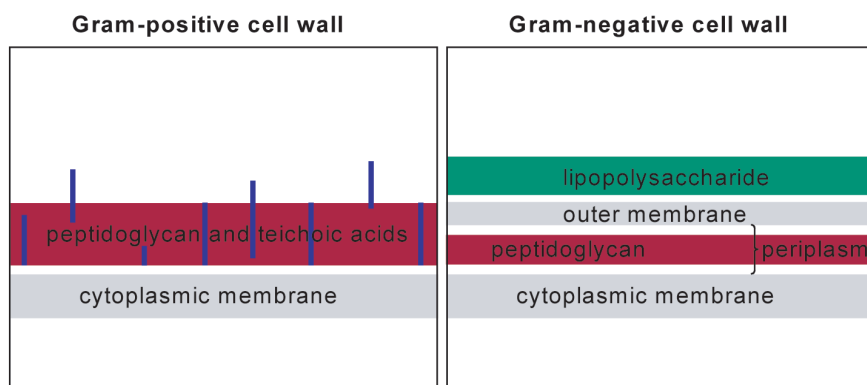


Figure 1. The cell walls of bacteria. Unlike those of Gram-negative bacteria, the peptidoglycan layers of Gram-positive bacteria are often modified by teichoic acids. Penicillin inhibits the biosynthesis of the peptidoglycan layer of many Gram-positive bacteria and some Gram-negative bacteria, but resistance to this and related antibiotics makes inhibition of alternate biosynthetic targets very attractive.

kinetic studies of this system particularly challenging. To circumvent this difficulty, a $C_{13}H_{27}$ hydrocarbon chain analogue was made by chemical synthesis and shown to be a substrate for the first enzyme, TagA, in the biosynthetic pathway. TagA forms a β -1,4-*N*-acetylmannosamine anomeric bond—one of the most problematic glycosidic linkages to make by chemical means alone. Both the first enzyme and the second enzyme in the pathway, TagB, are able to accept a substrate containing the shorter lipid. These *in vitro* studies demonstrate that long lipid chains are not necessary for the enzymatic reactions and, in addition, even a membrane interface is not crucial for the activity of TagA and TagB. These exciting discoveries suddenly make structural and molecular studies of these crucial WTA biosynthesis enzymes less daunting.

Although the basics of WTA biosynthesis have been elucidated by genetics experiments and crude studies using membrane preps (3), a molecular picture of the process necessary for drug develop-

ment is only starting to emerge. Structures of the cytidyltransferase enzyme that activates glycerol for polymerization are now available (6, 7), but the structures of other proteins in the WTA biosynthetic pathways remain to be solved. Recent progress on two fronts promises elucidation of the structures and catalytic mechanisms of two of the WTA proteins that are essentially black boxes now. The polymerization protein TagF has been produced in purified recombinant form (8). Initial studies with TagF assayed by incorporation of radioactive CDP-glycerol

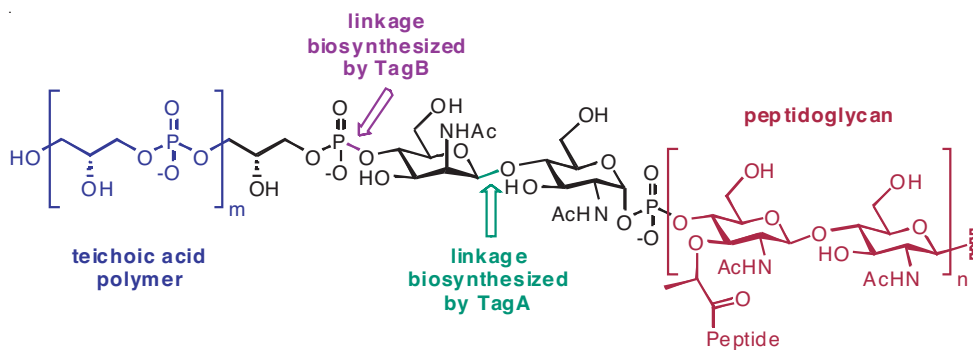


Figure 2. Teichoic acid polymer linked to peptidoglycan. TagA and TagB catalyze the biosynthesis of the highly conserved bridge between peptidoglycan and teichoic acid polymers. One potential antibiotic strategy that can be explored is to inhibit the reactions catalyzed by TagA and TagB.

into membranes suggest a mechanistic relationship between TagF and TagB, which shares about 30% pairwise sequence identity with the C-terminus of TagF (9). The work reported herein (2) finally nails down the role of TagB as a sort of primase for teichoic acid synthesis.

Anionic wall teichoic acid polymers attached to the peptidoglycans of Gram-positive bacteria not only protect the organism but also confer antigen specificity much like the lipopolysaccharides (LPS) that coat the surface of Gram-negative bacteria. Interestingly, patients with deep-seated *Staphylococcus aureus* infections all contained IgG antibodies to teichoic acid antigens whereas only 40% of those patients with superficial infections did; however, the level of IgG antibodies against peptidoglycan ranged from 60 to 72% in both patient groups (10). The anionic WTA coat of *S. aureus* helps host colonization and serves as a virulence factor by allowing binding of the bacteria to host endothelial cells (11–14). Access to defined WTA fragments, especially as more modification enzymes are added to the core teichoic acid biosynthesis enzymes, opens the possibility of understanding molecular level structure and function relationships required for host interactions with these essential Gram-positive bacterial cell wall components.

REFERENCES

1. Lambert, P. A. (2005) Bacterial resistance to antibiotics: modified target sites, *Adv. Drug Delivery Rev.* **57**, 1471–1485.
2. Ginsberg, C., Zhang, Y.-H., Yuan, Y., Walker, S. (2006) *In Vitro* reconstitution of two essential steps in wall teichoic acid biosynthesis, *ACS Chem. Biol.* **1**, 25–28.
3. Neuhaus, F. C., Baddiley, J. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria, *Microbiol. Mol. Biol. Rev.* **67**, 686–723.
4. Bhavsar, A. P., Erdman, L. K., Schertzer, J. W., Brown, E. D. (2004) Teichoic acid is an essential polymer in *Bacillus subtilis* that is functionally distinct from teichuronic acid, *J. Bacteriol.* **186**, 7865–7873.
5. Kobayashi, K., Ehrlich, S. D., Albertini, A., Amatic, G., Anderson, K. K., et al. (2003) Essential *Bacillus subtilis* genes, *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4678–4683.
6. Patridge, K. A., Weber, C. H., Friesen, J. A., Sanker, S., Kent, C., Ludwig, M. L. (2003) Glycerol-3-Phosphate Cytidylyltransferase. Structural changes induced by binding of CDP-glycerol and the role of lysine residues in catalysis, *J. Biol. Chem.* **278**, 51863–51871.
7. Fong, D. H., Yim, V. C., D'Elia, M. A., Brown, E. D., Berghuis, A. M. (2006) Crystal structure of CTP: glycerol-3-phosphate cytidylyltransferase from *Staphylococcus aureus*: examination of structural basis for kinetic mechanism, *Biochim. Biophys. Acta* **1764**, 63–69.
8. Schertzer, J. W., Brown, E. D. (2003) Purified, recombinant TagF protein from *Bacillus subtilis* 168 catalyzes the polymerization of glycerol phosphate onto a membrane acceptor *in vitro*, *J. Biol. Chem.* **278**, 18002–18007.
9. Schertzer, J. W., Bhavsar, A. P., Brown, E. D. (2005) Two conserved histidine residues are critical to the function of the TagF-like family of enzymes, *J. Biol. Chem.* **280**, 36683–36690.
10. Kumar, A.; Ray, P., Kanwar, M., Sharma, M., Varma, S. (2005) A comparative analysis of antibody repertoire against *Staphylococcus aureus* antigens in patients with deep-seated versus superficial staphylococcal infections, *Int. J. Med. Sci.* **2**, 129–136.
11. Weidenmaier, C., Peschel, A., Kempf, V. A., Lucindo, N., Yeaman, M. R., Bayer, A. S. (2005) DltABCD- and MprF-mediated cell envelope modifications of *Staphylococcus aureus* confer resistance to platelet microbicidal proteins and contribute to virulence in a rabbit endocarditis model, *Infect. Immun.* **73**, 8033–8088.
12. Kristian, S. A., Datta, C., Weidenmaier, C., Kansal, R., Fedtke, I., Peschel, A., Gallo, R. L., Nizet, V. (2005) D-Alanylation of teichoic acids promotes group A streptococcus antimicrobial peptide resistance, neutrophil survival, and epithelial cell invasion, *J. Bacteriol.* **187**, 6719–6725.
13. Weidenmaier, C., Peschel, A., Xiong, Y. Q., Kristian, S. A., Dietz, K., Yeaman, M. R., Bayer, A. S. (2005) Lack of wall teichoic acids in *Staphylococcus aureus* leads to reduced interactions with endothelial cells and to attenuated virulence in a rabbit model of endocarditis, *J. Infect. Dis.* **191**, 1771–1777.
14. Weidenmaier, C., Kokai-Kun, J. F., Kristian, S. A., Chanturiya, T., Kalbacher, H., Gross, M., Nicholson, G., Neumeister, B., Mond, J. J., Peschel, A. (2004) Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections, *Nat. Med.* **10**, 243–245.

Stopping Trouble before It Starts

Jon Clardy*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115

The translation of mRNA into proteins is a fundamental, tightly controlled, and complex process involving an elaborate molecular machine with both ribonucleic acids and proteins (Figure 1). Our understanding of translation has benefited from several small molecules that stall the process, many of which have become drugs. Many common antibiotics, such as tetracycline and erythromycin, stall prokaryotic translation in the processive stages where new amino acids are added to a growing chain, while cycloheximide, a fungicide, and anisomycin, an antiprotozoal, stall eukaryotic translation in the processive stages. A very recent publication (1) reports the first selective inhibition of prokaryotic initiation by the curious naturally occurring peptide GE81112 (Figure 2, 2) from *Streptomyces* sp. Eukaryotic initiation (Figure 1) involves several initiation factors such as the eIF4 complex in which eIF4A, eIF4E, and eIF4G play critical roles. The precise role of eIF4A, which is both an RNA-dependent ATPase and an ATP-dependent RNA helicase, is not known. Now, Liu *et al.* (2) show that pateamine A (Figure 2, 1) interferes with the interaction between eIF4A and eIF4G (Figure 1). These findings open the door for mechanistic studies of eIF4A and once again show the utility of natural products in understanding complex processes. A brief review of the discovery and development of Taxol (Figure 2, 3), a natural product that became an important anticancer drug, provides a historical perspective for this latest revelation.

Taxol (Figure 2, 3) illustrates both the enormous potential of natural products,

that is, identifying new targets and providing the basis for new therapies, along with their frustrating liabilities, that is, identifying their mechanism of action, assuring an adequate supply, and dealing with a molecular template ill-suited to traditional medicinal chemistry (3). Work on the molecule that was to become Taxol (Figure 2, 3) began in 1965 when a large-scale screening effort by the National Cancer Institute (NCI) found that a crude extract of the Pacific yew (*Taxus brevifolia*) had *in vivo* activity in a mouse leukemia model (L1210 and P388; 4). The active compound, Taxol, was isolated and identified in 1969, although it was not published until 1971 (4). At that time, Taxol was just one of many possible leads that had come out of the NCI program, but with the 1975 discovery of its pronounced activity against B16 melanoma, it became a development candidate. When the Horwitz laboratory defined Taxol's mechanism of action as promoting tubulin polymerization and stabilizing microtubules against depolymerization in 1979, drug development began in earnest. Although Taxol entered phase I clinical trials in 1983, it did not receive approval from the Food and Drug Administration (FDA) until 1992, a delay attributable both to supply and formulation issues (4). Taxol was first marketed in 1993; in the next 10 years, it had sales of \$9 billion, and in 2005, its sales will be a little over \$1 billion.

Pateamine A has now taken a significant step along this path. In 1991, the Munro and Blunt laboratory reported its isolation from a sponge (*Mycale* sp.) collected off

ABSTRACT A recent publication revealing that the cytotoxic marine natural product pateamine A targets eukaryotic initiation factor eIF4A continues a story with lessons for both chemists and biologists, that is, the significance of natural products, the importance of synthetic organic chemistry, the small molecule regulation of eukaryotic translation machinery, and possibly a new approach to cancer chemotherapy.

*To whom correspondence should be addressed.
E-mail: jon_clardy@hms.harvard.edu

Published online February 17, 2006

10.1021/cb0600029 CCC: \$33.50

© 2006 American Chemical Society

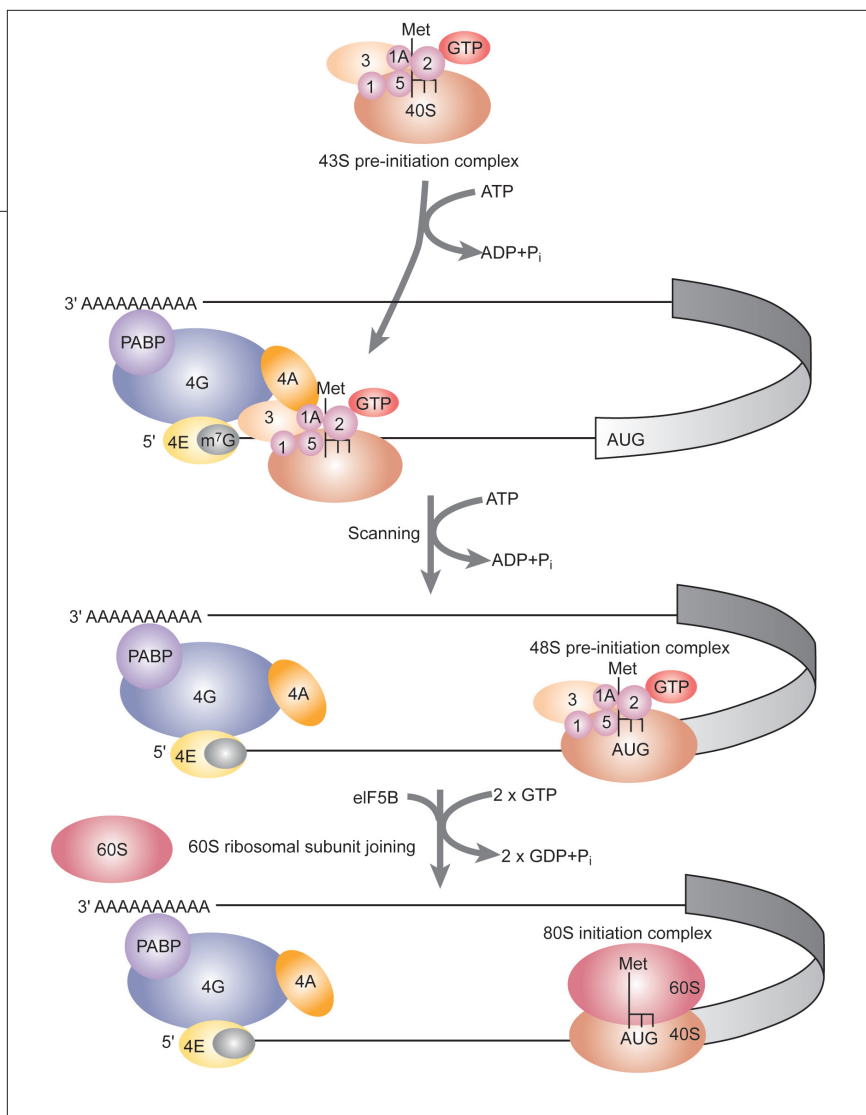


Figure 1. The role of the eIF4 complex in eukaryotic translation. The 5'-end of eukaryotic mRNAs contains a 7-methyl guanosine (m⁷G) cap structure that serves as a scaffold for the formation of a large protein complex containing various eukaryotic initiation factors (eIFs) including eIF4A, eIF4E, eIF4G, and PolyA binding protein (PABP). This complex brings the 5'- and 3'-ends of the mRNA in proximity and marks the mRNA for translation. The role of eIF4A in this complex is not clear, but it is known to bind to eIF4G. The 43S pre-initiation complex containing the 40S ribosomal subunit, initiator tRNA (stick model), eIF1, eIF1A, eIF2, eIF3, and eIF5 joins the cap complex, and scanning for the start codon (AUG) begins. Upon encountering the start codon, the complex is remodeled, the 60S large ribosomal subunit joins the process, and translation of the mRNA into protein begins.

the coast of New Zealand (5). The isolation scheme followed pateamine A's activity against P388 leukemia (using cell lines, not mice) along with additional studies that highlighted its selective cytotoxicity for rapidly growing cells (5). At that time, pateamine A was just one of many sponge metabolites with selective cytotoxicity against cancer cell lines, an unknown target, and an uncertain source of supply.

In 1995 the Romo laboratory set about exploring the chemistry of pateamine A

(Figure 2, 1) with two related goals: (1) find simpler active analogues and (2) identify pateamine A's cellular target. Key discoveries in this decade-long effort were an active compound lacking both the C-3 amino and C-5 methyl groups and close analogues with no activity, and the preparation of these analogues with deep-seated structural changes depended on efficient total syntheses. The ability to acylate the C-3 amino group without loss of activity led to the synthesis of a biotin-linked affinity

reagent (Figure 2, 4). In the Liu laboratory, this affinity reagent (Figure 2, 4) was used to identify two potential targets: STRAP, a serine-threonine kinase receptor-associated protein, and eukaryotic initiation factor eIF4A (2). The inability of the inactive pateamine A analogues to displace pateamine in the STRAP and eIF4A complexes confirmed the specificity of the binding. A study using HeLa-derived cell lines overexpressing each protein showed increased resistance to pateamine A for the eIF4A line, and no increased resistance for the STRAP line indicated that the physiologically relevant target was eIF4A. An independent study (6) using a high-throughput screen for protein synthesis inhibitors and secondary assays by the Pelletier laboratory reached the same conclusion a few months earlier. While agreeing on the target, the two laboratories differ on the detailed mechanism by which pateamine A disrupts protein synthesis. These studies suggest the following two things: (1) our ability to identify cellular targets for small molecules identified in simple cytotoxicity screens has not improved, as identifying the targets for Taxol and pateamine A both took roughly a decade and (2) including natural products in high-content screens more efficiently links natural products, biological activities, and targets.

Pateamine A is the first published (7) ligand to specifically modify eIF4A-based activity, and its discovery will undoubtedly lead to an improved understanding of eIF4A's role in protein translation. Somewhat paradoxically, pateamine A enhances both catalytic activities of eIF4A but inhibits protein translation. Low *et al.* resolved the paradox by using standard molecular biology techniques to show that pateamine A disrupts the formation of the eIF4 complex by decreasing the interaction between eIF4A and eIF4G (2). Further studies, which will undoubtedly employ pateamine A and its analogues, will be needed to precisely define eIF4A's role in

initiation and its regulation by small molecules. After the initial discovery of Taxol as a tubulin-interacting agent, many other tubulin-interacting molecules with very different molecular structures were discovered, and it is likely that many small molecules that target eIF4A will now emerge.

Pateamine A and its analogues could conceivably be drug development candidates, but development will have to overcome the current move towards targets specific to cancer cells (think Gleevec) and away from targets with widespread and essential activities such as protein synthesis. One could argue the case for pateamine A by noting that Taxol targets the widespread and essential cellular micro-

tubules, but trends are changing as more specific drugs show encouraging clinical efficacy.

If development of an anticancer drug based on pateamine A, or possibly pateamine A itself, proceeds, supply will be a major concern. Taxol supply was a major focus of the organic chemistry community with total synthesis, semisynthesis, and extraction from Pacific yew all contributing. Eventually, both extraction and semisynthesis starting with more readily available plant metabolites provided the most economical solution (4). Pateamine A represents a much bigger challenge. Marine sponges have proved to be notoriously difficult sources to recollect, and even the original

paper noted the sporadic occurrence of pateamine A in *Mycale* sponges. However, recent work suggests that many so-called sponge metabolites are biosynthesized by bacterial symbionts, and either these bacterial producers can be cultured in the laboratory or their biosynthetic genes can be placed into alternative hosts for heterologous expression. Two papers last year on the sponge-derived anticancer agent patellamide (Figure 2, 5) illustrate the power of the latter approach (8, 9). The one sure prediction is that the pateamine A story will continue.

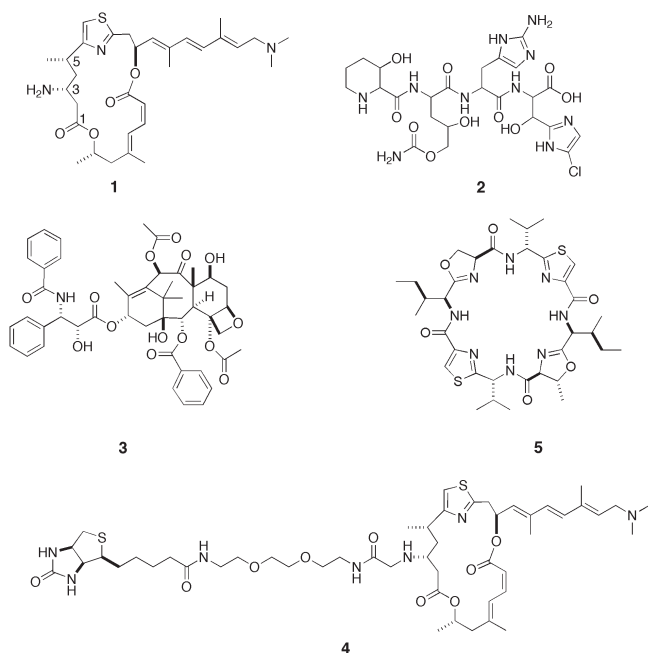


Figure 2. Pateamine A (1), which disrupts translation at the initiation stage by binding to eIF4A, was isolated from a marine sponge *Mycale* sp. GE81112 (2), which disrupts prokaryotic translations, was isolated from a *Streptomyces* sp. Taxol (3), which binds to tubulin, was isolated from the Pacific yew (*Taxus brevifolia*). The affinity reagent (4) was used to identify the cellular target of pateamine A. Patellamide (5), which has shown some promise as an anticancer agent, was originally isolated from a marine sponge, but later work showed that it was produced by a cyanobacterial symbiont.

REFERENCES

- Brandi, L., Fabbretti, A., La Teana, A., Abbondi, M., Losi, D., Donadio, S., and Gualerzi, C. O. (2006) Specific, efficient, and selective inhibition of prokaryotic translation initiation by a novel peptide antibiotic, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 39–44.
- Low, W. K., Dang, Y., Schneider-Poetsch, T., Shi, Z., Choi, N. S., Merrick, W. C., Romo, D., and Liu, J. O. (2005) Inhibition of eukaryotic translation initiation by the marine natural product pateamine A, *Mol. Cell* **20**, 709–722.
- Clardy, J., and Walsh, C. (2004) Lessons from natural molecules, *Nature* **432**, 829–837.
- Cragg, G. M. (1998) Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development, *Med. Res. Rev.* **18**, 315–331.
- Northcote, P. T., Blunt, J. W., and Munro, M. H. G. (1991) Pateamine: a potent cytotoxin from the New Zealand marine sponge, *Mycale* sp., *Tetrahedron Lett.* **32**, 6411–6414.
- Bordeleau, M. E., Matthews, J., Wojnar, J. M., Lindqvist, L., Novac, O., Jankowsky, E., Sonenberg, N., Northcote, P., Teesdale-Spittle, P., and Pelletier, J. (2005) Stimulation of mammalian translation initiation factor eIF4A activity by a small molecule inhibitor of eukaryotic translation, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10460–10465.
- Justman, C. J. (1999) Torreyanic acid: affinity chromatographic identification of receptors and biochemical analysis of the torreyanic acid-eIF4A complex, Ph.D. Thesis, Harvard University, Cambridge, MA.
- Schmidt, E. W., Nelson, J. T., Rasko, D. A., Sudek, S., Eisen, J. A., Haygood, M. G., and Ravel, J. (2005) Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinium patella*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7315–7320.
- Long, P. F., Dunlap, W. C., Battershill, C. N., and Jaspars, M. (2005) Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production, *ChemBioChem* **6**, 1760–1765.

From Loops to Chains: Unraveling the Mysteries of Polyubiquitin Chain Specificity and Processivity

Mathew E. Sowa and J. Wade Harper*

Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115

ABSTRACT Regulated protein degradation via polyubiquitination controls almost every aspect of eukaryotic cellular biology; however, the precise mechanism by which specifically linked polyubiquitin chains are formed on target proteins as well as how the processivity of chain elongation is achieved remains a mystery. Recent work using the yeast ubiquitin ligase SCF^{Cdc4} and the ubiquitin conjugating enzyme, Cdc34, has helped to answer these questions by identifying the determinants of lysine-48 specific ubiquitin chain polymerization.

Protein turnover through the ubiquitin–proteasome pathway controls a vast assortment of signaling pathways, and genetic inactivation of individual components in the system is responsible for numerous human diseases, ranging from neurodegenerative diseases to cancer (1, 2). In response to particular signals, specific regulatory proteins are tagged with a chain of ubiquitin molecules through the action of an enzymatic cascade composed of an E1 ubiquitin activating enzyme, an E2 ubiquitin conjugating enzyme (Ubc), and an E3 ubiquitin ligase. E3s are responsible for binding substrates and for bringing substrates into the proximity of a ubiquitin-charged E2, which then transfers ubiquitin to the substrate and to the ensuing ubiquitin chain in a processive reaction. Formation of lysine-48 linked polyubiquitin chains facilitates recognition and degradation by the 26S proteasome (1). During the last decade, our understanding of the genes that contribute to the ubiquitin pathway and many of the rules that control substrate selection by E3s have been uncovered. In addition, structural information is available for each of the two main classes of E3s: HECT and RING domain containing proteins (3, 4). These structures reveal the major role of proximity in facilitating ubiquitin transfer to substrates but tell us very little about the dynamics and chemical mechanisms involved in ubiquitin transfer and chain elongation. Indeed, a central unanswered question concerns how particular lysine residues in either the substrate or the growing ubiquitin chain are selected for chain elongation. Two

major classes of ubiquitin–ubiquitin linkages, lysine-48 (K48) and lysine-63 (K63), are found *in vivo*, but precisely how E2s select one of the seven lysines in ubiquitin (Figure 1, panel a) over another is poorly defined. Moreover, ubiquitination reactions can be highly processive, yet it is unclear how processivity is achieved given the structural constraints of a growing ubiquitin chain. Recent work published by Petroski and Deshaies (5) has provided novel insight into how the yeast E2, Cdc34, functions together with the RING E3, SCF^{Cdc4}, to promote processive polyubiquitination of its substrate, Sic1, and how selective chain elongation at K48 in ubiquitin is achieved. Unexpectedly, acceleration of the ubiquitin–ubiquitin conjugation step appears to reflect a specialized motif in Cdc34 that serves to orient K48 in the acceptor ubiquitin molecule for attack on the donor ubiquitin thiol-ester bond.

Conjugating Ubiquitin. Ubiquitin is a 76 amino acid protein (Figure 1, panel a) that becomes activated for conjugation through the action of E1 ubiquitin activating enzyme. E1 uses ATP to form a thiol-ester between its active site cysteine residue and the C-terminal carboxylate of ubiquitin. The E1-S~Ub intermediate is then recognized by one of more than 30 E2 ubiquitin conjugating enzymes encoded by the human genome, which use an active site cysteine to perform a trans thiolesterification reaction giving E2-S~Ub. At steady state *in vivo*, individual E2s exist as a mixture of ubiquitin-charged and -uncharged forms. E2 enzymes possess a highly conserved core domain, contain-

*To whom correspondence should be addressed.
E-mail: wade_harper@hms.harvard.edu.

Published online February 17, 2006

10.1021/cb0600020 CCC: \$33.50

© 2006 by American Chemical Society

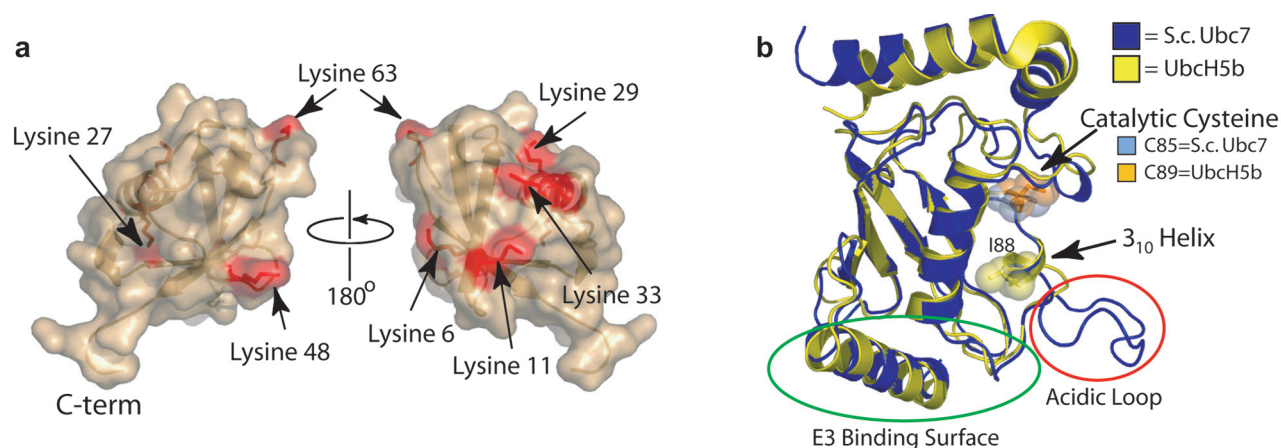


Figure 1. Structure of ubiquitin and the E2. a) Ubiquitin is small protein comprised of only 76 amino acids and forms a compact structure with the C-terminal tail exposed. A glycine residue is found at the C-terminus, which is used to form the thiol-ester bond between ubiquitin and the active site cysteine in both E1 and E2 proteins and which ultimately makes an isopeptide linkage with either the substrate or with a second ubiquitin molecule. Ubiquitin has seven surface-exposed lysine residues, the most well studied of which are K48 and K63. b) Two different E2 proteins, *S. cerevisiae* Ubc7 (2UCV) and *H. sapiens* UbcH5 (1ESK), are shown aligned by the C α backbones using Pymol. The acidic loop insertion in Ubc7 is clearly visible and is located near the E3 binding site. I88 from UbcH5 is found on the 3₁₀ helix common to all E2s.

ing the active site cysteine, as well as sequences that are used to recognize RING or HECT motifs in E3s (Figure 1, panel b). Many E2s have yet to be studied in detail, but it is clear that there is some specificity among E2s for particular subclasses of E3s. However, the rules dictating the functional complementarity of E2s and E3s are largely unknown.

Once a particular ubiquitin-charged E2 associates with a RING or HECT domain present in an E3, one of two processes can occur. If the E3 lacks a bound substrate, the E2 can discharge its ubiquitin to form an isopeptide bond with the ϵ -amino group from one of seven lysine residues in a second ubiquitin molecule through a poorly understood process. Indeed, for some isolated RING domain E3s in the absence of substrate, binding to the E2 stimulates the formation of polyubiquitin chains initiating from a single ubiquitin molecule (6). In contrast, if a substrate is bound to the E3, ubiquitin may be preferentially discharged to a lysine residue in the substrate. At least part of the preference for transfer to sub-

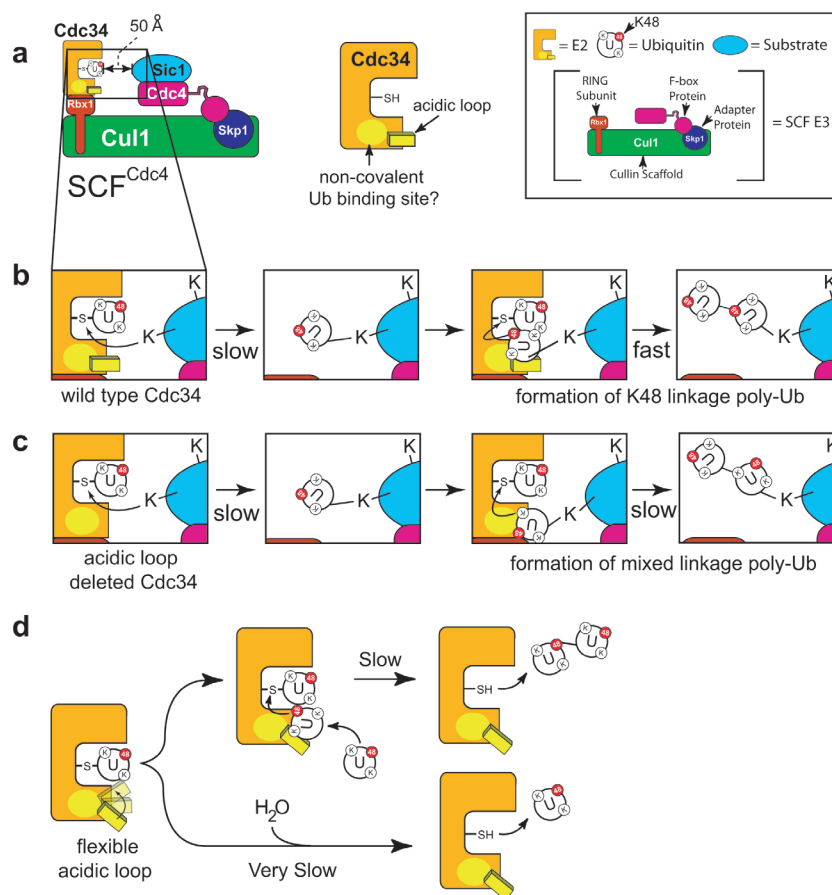
strate, as opposed to free ubiquitin, comes from the fact that the bound substrate is present at an effective concentration in the millimoles per liter range while free ubiquitin in the cell (or in most *in vitro* experiments) is present at ~ 2 orders of magnitude lower effective concentration.

Understanding the biochemical basis for E3-accelerated ubiquitin discharge is a central problem in the ubiquitin field, and several questions have emerged: First, how are the appropriate lysine residues in the substrate selected for conjugation. For some substrates, it has been demonstrated that different turnover rates are achieved by ubiquitination of different lysine residues (7), so in biological systems, it may be important that one particular lysine out of many serves as the recipient for the first ubiquitin. There is little evidence that E2s for ubiquitin can impart this type of specificity, but structural data suggest a major role for the E3 in orienting the substrate for appropriate conjugation (8, 9). Second, once the initial lysine has been selected, polyubiquitination can occur in a highly

processive manner. It is unclear precisely what biochemical features of the system drive processivity nor is it clear how the structural constraints of a growing ubiquitin chain are reconciled with the requirement that the substrate and E2 likely remain physically associated with the E3 for chain elongation even after many ubiquitin molecules have been attached to a substrate.

Mechanism of Ubiquitin Release from the E2. To begin to answer these questions, Petroski and Deshaies (5) examined the kinetics of ubiquitin transfer using the prototypical RING E3, SCF^{Cdc4}. SCF complexes are composed of the scaffold Cul1, the adaptor protein Skp1, the substrate receptor Cdc4, and the RING protein Rbx1 (2) (Figure 2, panel a). The best characterized SCF^{Cdc4} substrate, Sic1 (2), binds to Cdc4 in a phosphorylation-dependent manner and is ubiquitinated by the E2 enzyme Cdc34, which interacts with the RING domain of Rbx1 (10, 11). Previous work had demonstrated that the conjugation activity of Cdc34 (either on itself as a surrogate substrate or on free ubiquitin)

Figure 2. Model for polyubiquitin chain specificity and processivity. a) The SCF^{Cdc4} ubiquitin ligase is a large multisubunit protein complex which binds the E2 through the RING domain protein Rbx1, substrate (in this example, Sic1) through the F-box protein (in this example, Cdc4), and Cul1 which acts as a scaffold. The *S. cerevisiae* E2, Cdc34, contains an “acidic loop” located near its E3 binding site and also includes a low affinity noncovalent ubiquitin binding site, the location of which is not known but could be contiguous with the “acidic loop”. b) Initial monoubiquitination of Sic1 is relatively slow, but subsequent rounds of ubiquitin conjugation to the initial ubiquitin are relatively rapid and specific for K48 due to proper orientation of the attacking ubiquitin’s K48 by the Cdc34 “acidic loop”. c) In the Cdc34 “acidic loop” mutant, the formation of polyubiquitin chains is slow and the linkage is nonspecific due to the lack of appropriate positioning of the attacking ubiquitin. d) In the absence of E3, Cdc34 is still able to catalyze the formation of ubiquitin conjugates through attack of the E2~Ub thiol-ester bond by another ubiquitin, but this reaction is much slower than in the presence of E3. In the absence of free ubiquitin (or absence of available lysines in free ubiquitin), the discharge of ubiquitin from the E2 is very slow, relying solely on the hydrolysis of the thiol-ester bond.



is greatly stimulated by Rbx1, a function that was attributed to association of the RING domain with Cdc34 (6, 10, 11), and that Cdc34 is specific for the generation of K48-linked polyubiquitin chains.

In order to dissect individual steps, Petroski and Deshaies performed single-turnover assays, initially focusing on ubiquitin discharge from Cdc34 (Figure 2, panel d). Such single-turnover reactions have the ability to reveal changes in rate constants associated with engagement of Cdc34-S~Ub with SCF^{Cdc4} or substrate-SCF^{Cdc4} complexes. As expected, discharge of ubiquitin was enhanced by the SCF complex, independently of substrate, but when a version of ubiquitin lacking lysine residues (UbK0) was used, the rate of discharge was greatly decreased and

was not stimulated by the SCF. This result indicated a role for lysines in ubiquitin in promoting the discharge reaction, but because the Cdc34-S~UbK0 intermediate is relatively stable, it simultaneously set the stage for an examination of how different molecules attack the thiol-ester.

Experiments using wild-type and lysine-free Sic1 (Sic1-K0) demonstrated a requirement for lysine residues in the substrate to promote the UbK0 discharge reaction (Figure 2, panel b). Moreover, SCF-driven discharge using a second ubiquitin molecule to form diubiquitin was dependent upon the presence of Lys-48 in ubiquitin, consistent with the specificity of Cdc34 for building K48-linked chains (Figure 2, panel d). The effect of SCF on the discharge rate was due exclusively

to a 40-fold enhancement in V_{max} , as the presence of the SCF had little impact on the K_m for ubiquitin in the diubiquitin synthesis reaction (~600 μ M).

Chain Specificity and Processivity.

Detailed analysis of ubiquitination products on Sic1 proteins containing multiple lysine recipients or a single lysine recipient using wild-type versus K0 ubiquitin revealed that the primary reaction products were Sic1 molecules containing a ubiquitin chain, as opposed to multiple monoubiquitination events (5). The implication of this result is that attachment of the first ubiquitin molecule to the substrate is the rate-limiting step in the conjugation process (Figure 2, panel b). To test this, Petroski and Deshaies generated Sic1 containing a single ubiquitin conjugated to lysine 36 (Sic1-Ub₁) and

compared the rate of its polyubiquitination by SCF–Cdc34 with Sic1 lacking ubiquitin. Surprisingly, the rate of ubiquitination was up to 10-fold higher for Sic1-Ub1 than for Sic1, indicating that the rate of ubiquitin chain extension on a substrate already containing the first ubiquitin is faster than the initial conjugation step (Figure 2, panel b).

The molecular details of this effect are likely to underlie the processivity of polyubiquitination. One possible explanation is that the presence of the first ubiquitin places the accepting lysine much closer to the Cdc34-S-Ub bound to the RING domain. Indeed, crystallographic data of substrates bound to SCF^{Cdc4} indicate that the position of the bound E2 could be as much as 50 Å away from the substrate lysine, and this gap could be significantly bridged by the first ubiquitin conjugation step (Figure 2, panel a). Alternatively, Petroski and Deshaies surmised, Cdc34 may have an intrinsic affinity for K48 in ubiquitin relative to lysine residues in substrates. The implication of this idea is that Cdc34 would be capable of more rapidly forming a polyubiquitin chain because of the enhanced molecular complementarities between Cdc34 and its elongating substrate.

To test whether there is something special about K48 in recipient ubiquitin, the authors tested whether the K48R ubiquitin mutant is conjugated faster to Sic1-Ub1 than to Sic1 and found that unlike the case with wild-type Ub there was no rate enhancement (5). Minimally, this indicated that acceleration of the second conjugation event required K48 in the recipient ubiquitin. Through a further series of experiments, the authors demonstrated that the rate increase seen for K48 ubiquitin chain extension is an intrinsic feature of Cdc34 and can be traced to a feature which, among E2s, is only found in Cdc34 and Ubc7 orthologs. Alignment of more than 30 human E2s for ubiquitin reveals that Cdc34 and Ubc7 orthologs

all contain a unique insertion near the active site cysteine residue, which contains several acidic residues. The crystal structure of Ubc7 superimposed on UbcH5b, lacking the “acidic loop”, shows that the loop is on the surface of the enzyme and could potentially influence ubiquitin conjugation (Figure 1, panel b). In fact, mutations in the acidic loop of Cdc34 are largely nonfunctional *in vivo* (12). Kinetic analysis of acidic loop mutations in Cdc34 revealed a complex relationship between the presence of this structural feature and the rates and types of conjugates that can form (5). First, Cdc34 acidic loop mutants were highly active in transferring lysine-free ubiquitin to Sic1 as well as in promoting polyubiquitination of Sic1, indicating that the acidic loop is not crucial to polymerization of ubiquitin per se (Figure 2, panel c). However, analysis of the types of chains formed indicated that the acidic loop mutants of Cdc34 were very inefficient in the formation of K48-linked ubiquitin chains. In this sense, the acidic loop mutants can be thought of as maintaining intrinsic conjugation activity while losing selectivity in the formation of degradation-competent K48-linked chains (Figure 2, panel c).

Models for Processive Ubiquitin

Transfer: Orientation versus Allostery. The data presented by Petroski and Deshaies suggest a model wherein processivity in ubiquitin chain formation by Cdc34 reflects an inherent specificity of Cdc34 for utilization of K48 in the recipient ubiquitin in the conjugation reaction (Figure 2, panel b). Presumably, the acidic loop helps to bind ubiquitin and orient K48 in an optimal position for catalysis, an effect that is revealed in V_{\max} . The effect of the acidic loop in this model is reminiscent of how the heterodimeric E2, Ubc13–Mms2, functions to orient ubiquitin to achieve selective conjugation at K63 (13). In this case, a catalytically inactive Mms2 E2 binds the recipient ubiquitin in a manner that orients K63 in the direction of the active

site cysteine of the associated Ubc13-S-Ub molecule. The location of the acidic loop in Ubc7, and presumably in Cdc34, is located in a structurally distinct position relative to Mms2 in the Ubc13 complex but could nevertheless perform the analogous orientation function for K48 chain extension. Of additional note is that the Mms2 subunit in the Ubc13–Mms2 heterodimer likely acts to decrease K_m for the attacking ubiquitin, which is in contrast to the acidic loop of Cdc34 which has little to no effect on K_m ; however, in both cases these structural features can increase V_{\max} .

An important question concerns the generality of this model. Indeed, the vast majority of E2s, including UbcH4 and UbcH5, which are capable of functioning with the SCF *in vitro*, lack the acidic loop feature (Figure 1, panel b). Thus, it is unclear precisely how these enzymes achieve processivity and selectivity. UbcH4 and UbcH5 can be activated by several RING E3s *in vitro*, including the SCF, although in some cases it appears that they are not as processive as Cdc34. Interestingly, Özkan *et al.* (14) have recently investigated the possibility of an allosteric pathway within the E2 UbcH5b, which may provide some insight into the mechanism of ubiquitin discharge during both the initial monoubiquitination of substrate and the proceeding polyubiquitination reactions. In this work, a computational method called statistical coupling analysis (SCA) (15) was used to determine residues in the E2 family of proteins (345 individual sequences) that have similar patterns of variation over evolution and are hence considered to have coevolved. In addition to the expected cluster of “coupled” hydrophobic core residues, there was a second cluster located near the E2 active site.

Using a ubiquitin discharge assay, similar to that used by Petroski and Deshaies, Özkan *et al.* found that mutation of I88 to alanine, among several mutants examined, had the greatest effect

Sequences proximal to the active site of the E2 can govern the specificity and rate of the polyubiquitin chain synthesis on substrates.

on discharge stimulated by two RING E3s (the anaphase promoting complex (APC) subunit Apc11 and the RING domain of CNOT4) (14). The authors suggested that I88 functions in an allosteric network that links the RING binding site with the catalytic apparatus of UbcH5b. However, I88 in UbcH5b is located at the point where the acidic loop of Ubc7 emerges from the conserved 3_{10} helix located adjacent to the E2 active site (Figure 1, panel b). Therefore, I88 may not be part of an allosteric network coupling RING binding to the catalytic cysteine but instead may be playing a role analogous to that of the acidic loop in Cdc34–Ubc7, to optimally orient an attacking lysine toward the thiol-ester bond of the E2~Ub. Petroski and Deshaies found that the acidic loop mutants in Cdc34 formed mixed linkage polyubiquitin chains, indicative of a loss of K48 specificity (5).

If the UbcH5b I88A mutant is analogous to the Cdc34 acidic loop mutants, then two effects would be predicted. First, rates of substrate-independent discharge of the UbcH5b mutant should be slower, both with and without E3 binding. Second, it should be able to transfer the initial ubiquitin to the E3-bound substrate faster than the wild type yet should form random lysine linkages of lower molecular weight than wild-type UbcH5b. From the limited data available, it is clear that the UbcH5b I88A mutant discharges ubiquitin at a slower rate than wild-type UbcH5b in the presence of Apc11 and CNOT4, consistent with the first prediction. Moreover, when UbcH5b I88A is used in conjunction with the full APC complex and its substrate cyclin B, a drastic reduction in polyubiquitin chain length was observed, instead larger amounts of shorter ubiquitin chains are observed as compared with wild type.

Although an allosteric mechanism for facilitating ubiquitin discharge from the E2 is still possible, the phenotype described for the UbcH5b I88A mutant is also consistent with a loss in processivity. Although

these data do not definitively show that the I88A mutant is behaving in a manner that is similar to the Cdc34 acidic loop mutants, they are consistent with the idea that sequences proximal to the active site of the E2 can govern the specificity and rate of the polyubiquitin chain synthesis on substrates. To distinguish between the possibilities of a coupled allosteric pathway within the E2 or an orientational role of the 3_{10} helix (including surrounding residues), it will be necessary to investigate the effect of the UbcH5b I88A mutant, as well as similar mutations in other E2s, using the comprehensive battery of assay methods analogous to those employed to study Sic1 but within the context of different E3-substrate pairs. Although the jury is still out on the generality of this mechanism with respect to controlling processivity and specificity in ubiquitination reactions, this work has uncovered a central feature of Cdc34 that facilitates K48 specific chain elongation in the context of the SCF ubiquitin ligase.

Acknowledgment: Our work on the ubiquitin pathway is supported by grants from the National Institute of General Medical Sciences and the National Aging Institute to J.W.H. M.S. is the recipient of an American Cancer Society Postdoctoral Fellowship.

REFERENCES

1. Pickart, C. M. (2004) Back to the future with ubiquitin, *Cell* 116, 181–190.
2. Petroski, M. D., and Deshaies, R. J. (2005) Function and regulation of cullin-RING ubiquitin ligases, *Nat. Rev. Mol. Cell Biol.* 6, 9–20.
3. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W., and Pavletich, N. P. (2002) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex, *Nature* 416, 703–709.
4. Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P.M., Huijbregtse, J. M., and Pavletich, N. P. (1999) Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade, *Science* 286, 1321–1326.
5. Petroski, M. D., and Deshaies, R. J. (2005) Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34, *Cell* 123, 1107–1120.
6. Ohta, T., Michel, J. J., Schottelius, A. J., and Xiong, Y. (1999) ROC1, a homolog of APC11, represents a family of cullin partners with an associated ubiquitin ligase activity, *Mol. Cell.* 3, 535–541.
7. Petroski, M. D., and Deshaies, R. J. (2003) Context of multiubiquitin chain attachment influences the rate of Sic1 degradation, *Mol. Cell.* 11, 1435–1444.
8. Wu, G., Xu, G., Schulman, B. A., Jeffrey, P. D., Harper, J. W., and Pavletich, N. P. (2003) Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase, *Mol. Cell.* 11, 1445–1456.
9. Orlicky, S., Tang, X., Willems, A., Tyers, M., and Sicheri, F. (2003) Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase, *Cell* 112, 243–256.
10. Skowyra, D., Koepp, D. M., Kamura, T., Conrad, M. N., Conaway, R. C., Conaway, J. W., Elledge, S. J., and Harper, J. W. (1999) Reconstitution of G1 cyclin ubiquitination with complexes containing SCFgr1 and Rbx1, *Science* 284, 662–665.
11. Seol, J. H., Feldman, R. M., Zachariae, W., Shevchenko, A., Correll, C. C., Lyapina, S., Chi, Y., Galova, M., Claypool, J., Sandmeyer, S., Nasmyth, K., Shevchenko, A., and Deshaies, R. J. (1999) Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34, *Genes Dev.* 13, 1614–1626.
12. Pitluk, Z. W., McDonough, M., Sangan, P., and Gonda, D. K. (1995) Novel CDC34 (UBC3) ubiquitin-conjugating enzyme mutants obtained by charge-to-alanine scanning mutagenesis, *Mol. Cell. Biol.* 15, 1210–1219.
13. Chan, N. L., and Hill, C. P. (2001) Defining polyubiquitin chain topology. *Nat. Struct. Biol.* 8, 650–652.
14. Özkan, E., Yu, H., and Deisenhofer, J. (2005) Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18890–18895.
15. Lockless, S. W., and Ranganathan, R. (1999) Evolutionarily conserved pathways of energetic connectivity in protein families, *Science* 286, 295–299.

In Vitro Reconstitution of Two Essential Steps in Wall Teichoic Acid Biosynthesis

Cynthia Ginsberg, Yu-Hui Zhang, Yanqiu Yuan, and Suzanne Walker*

Department of Microbiology and Molecular Genetics, Harvard Medical School, and Department of Chemistry & Chemical Biology, Harvard University, Boston, Massachusetts 02115

The emergence of antibiotic resistance poses a major threat to human health, prompting interest in the exploration of new antibiotic targets (1). Many antibiotics inhibit the biosynthesis of peptidoglycan, the cross-linked carbohydrate polymer that comprises the major structural component of the cell wall and prevents the cell from bursting under high internal osmotic pressure. Peptidoglycan is not the only important cell wall component, however. The peptidoglycan layers in many Gram-positive organisms are functionalized with wall teichoic acids (WTAs), anionic polymers that are attached to peptidoglycan via a phosphodisaccharide core (Figure 1; 2). Although their exact functions are unknown, wall teichoic acids play important biological roles. They have been shown to be essential for survival in *Bacillus subtilis* (3, 4) and to function as virulence factors that promote *Staphylococcus aureus* infections (5, 6). The enzymes involved in WTA synthesis are, therefore, potential targets for the development of new antibiotics to treat Gram-positive bacterial infections. Here we report the preparation of synthetic substrate analogues and their use in reconstituting the activities of TagA and TagB, two essential enzymes in the pathway for wall teichoic acid biosynthesis in *B. subtilis*, which is the major Gram-positive model organism. This work verifies the proposed functions of the enzymes and lays the foundation for detailed mechanistic and structural studies.

WTAs are synthesized as lipid-linked precursors on the cytoplasmic surface of the bacterial membrane (Figure 1; 2).

The first membrane-bound intermediate, undecaprenyl-diphospho-*N*-acetylglucosamine (GlcNAc-pp-und) (Figure 1), which is formed by the enzyme TagO, is utilized in the synthesis of several different anionic cell wall polymers in *B. subtilis* (7). Therefore, the first committed step in the biosynthesis of the major WTA in *B. subtilis* 168 involves the transfer of *N*-acetylmannosamine (ManNAc) from UDP to the C4 hydroxyl of GlcNAc-pp-und to form the ManNAc- β -(1,4)-GlcNAc disaccharide core of WTA (Figure 1). This disaccharide core is elaborated to a charged polymer by a series of enzymatic modifications that are proposed to begin with the addition of a glycerol phosphate ester to the C4 position of the ManNAc sugar. The enzymes proposed to catalyze these first two committed steps in teichoic acid biosynthesis, TagA and TagB, were identified based on data from transcriptional fusions (8), thermosensitive mutations (9, 10), and sequence homology (11). However, the putative functions of TagA and TagB have not been demonstrated *in vitro* because the lipid-anchored substrates needed to monitor activity cannot be isolated in useful quantities from bacterial cells. Moreover, these substrates contain a 55-carbon 'carrier' lipid that is expected to hamper both substrate isolation and kinetic investigations.

Problems related to substrate availability and lipid chain length can potentially be addressed by synthesizing alternative substrates containing shorter lipid chains. However, chemical synthesis cannot provide ready access to the substrates down-

ABSTRACT Wall teichoic acids (WTAs) are anionic polymers that decorate the cell walls of many Gram-positive bacteria. These structures are essential for survival or virulence in many organisms, which makes the enzymes involved in their biosynthesis attractive targets for the development of new antibacterial agents. We present a strategy to obtain WTA biosynthetic intermediates that involves a combination of chemical and enzymatic transformations. Using these intermediates, we have reconstituted the first two committed steps in the biosynthetic pathway. This work enables the exploration of WTA-synthesizing enzymes as antibiotic targets.

*To whom correspondence should be addressed.
E-mail: suzanne_walker@hms.harvard.edu

Received for review October 2, 2005
and accepted October 28, 2005

Published online January 24, 2006
10.1021/cb0500041 CCC: \$33.50

© 2006 by American Chemical Society

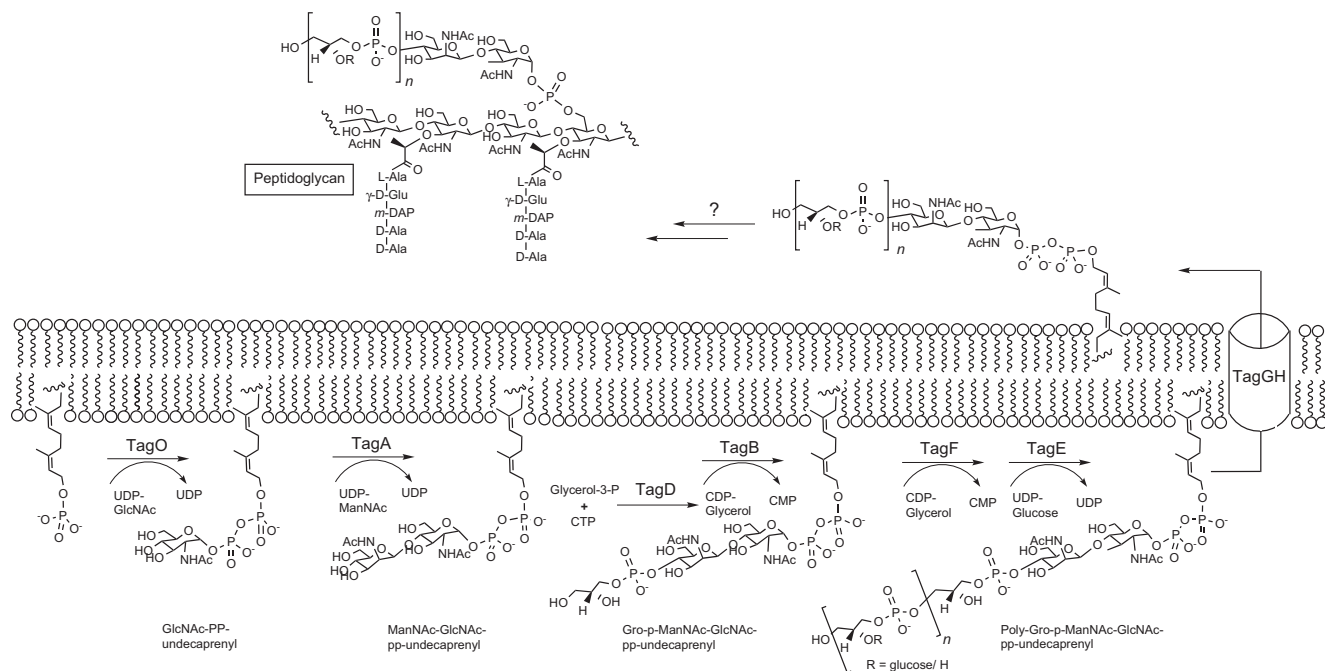


Figure 1. Polyglycerol phosphate WTA biosynthesis in *B. subtilis* 168.

stream of TagA, which contain β linkages to *N*-acetyl mannosamine. 1,2-*cis*-*O*-Glycosidic linkages to mannosamine are notoriously challenging to construct, and the ManNAc- β -(1,4)-GlcNAc linkage found in WTA has never been made chemically. Furthermore, synthetic routes to related compounds are lengthy and involve cumbersome functional group interconversions (12, 13; see Supporting Information.) Therefore, we decided to explore a chemo-enzymatic strategy in which TagA itself is used to make the ManNAc- β -(1,4)-GlcNAc linkage found in all downstream substrates in the WTA pathway.

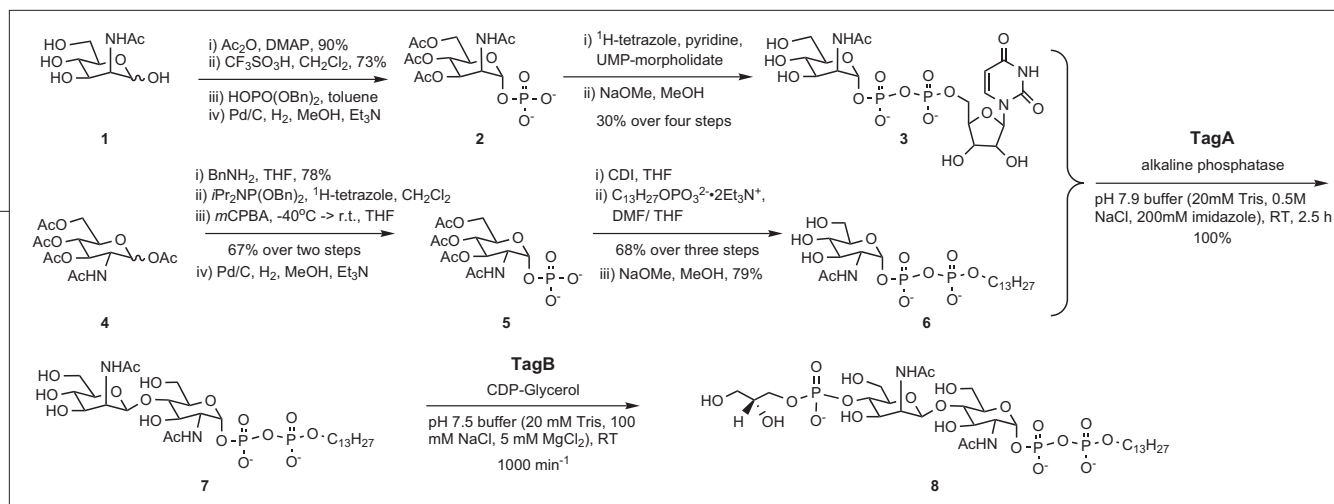
The natural TagA acceptor contains a 55-carbon undecaprenyl chain that is embedded in the bacterial membrane. Sequence analysis suggests that TagA does not contain any membrane spanning regions, and we surmised that the undecaprenyl chain is not specifically recognized by TagA. To simplify the synthesis and subsequent kinetic studies, we prepared an acceptor analogue containing a 13-carbon saturated lipid chain (Scheme 1, **6**; 14–17).

Following a known route, we also prepared the required UDP-ManNAc donor sugar **3**, which is not commercially available (Scheme 1; 18, 19).

The ability of TagA to utilize acceptor analogue **6** was evaluated by incubating the recombinant enzyme with synthetic substrates **3** and **6** overnight and analyzing the reaction mixture by liquid chromatography/mass spectrometry (LC/MS). A large peak at $m/z = 765$ was observed, consistent with the formation of the ManNAc-GlcNAc-pp-lipid product. To assess the efficiency with which TagA uses **6**, we measured the reaction rate using a continuous coupled enzyme assay that links the production of UDP to the oxidation of NADH (15, 17, 20, 21). The apparent K_m of **6** at a UDP-ManNAc concentration of 1.5 mM was found to be $190 \pm 30 \mu\text{M}$, and the turnover number was calculated to be $410 \pm 50 \text{ min}^{-1}$, consistent with the role of this glycosyltransferase in the biosynthesis of a primary metabolite. The kinetic measurements confirm the proposed function of TagA as a ManNAc transferase and

demonstrate that this synthetic acceptor analogue is a reasonable substrate to use in further characterization of TagA.

The high turnover number of TagA suggested that it would have utility for preparing the substrate for the next enzyme in the pathway, TagB. Unlike TagA, TagB has previously been overexpressed and efforts to reconstitute activity have been reported (22). Because the disaccharide acceptor of TagB was not available, however, activity was monitored by following the incorporation of radioactivity from CDP- ^{32}P -glycerol into bacterial membranes thought to contain the TagB substrate. The reported turnover number, 0.004 min^{-1} , is far too low to keep pace with cell wall biosynthesis, indicating a problem with either the assay conditions or the enzyme itself. To obtain pure substrate to monitor TagB activity, we converted 2 mg of compound **3** to compound **7** using TagA and UDP-ManNAc (Scheme 1). In the presence of alkaline phosphatase to consume UDP, quantitative conversion to product was achieved within 2.5 h. The identity of the product was con-



Scheme 1. Preparation of substrates for TagA and TagB.

firmed by NMR, with TOCSY spectra used to make resonance assignments and the expected β stereochemistry of the linkage confirmed by large ROESY cross peaks between H¹, H³, and H⁵ of the ManNAc sugar (See Supporting Information). Because the chemical synthesis of **3** and its subsequent enzymatic transformation to **7** are straightforward, it should be possible to obtain more material simply by scaling up the enzymatic reaction; however, we note that milligram quantities of substrates are typically sufficient for thousands of kinetic assays.

The purified recombinant enzyme was incubated with substrate **7** and commercially available CDP-glycerol, and the reaction mixture was analyzed by LC/MS. After a 3 h incubation, the starting material **7** was virtually gone and new peaks were observed at $m/z = 459$ and 919 , consistent with the formation of the Gro-p-ManNAc- β -(1,4)-GlcNAc-pp-lipid product **8** (Scheme 1). The appearance of this product was found to depend on the presence of active TagB and both substrates, thus verifying the proposed function of the enzyme.

The activity of TagB was evaluated by separating quenched reactions on a high-performance liquid chromatography column and quantitating CMP formation by UV absorbance. The reaction rate was found to be $1 \mu\text{M product min}^{-1} \text{ nM}^{-1} \text{ enzyme}$, providing an estimated turnover number of 1000 min^{-1} under the reaction conditions. This turnover number is more than 5 orders of magnitude greater than that reported previously, highlighting the advantages

of using discrete synthetic substrates rather than crude bacterial membrane preparations in studying WTA synthesizing enzymes. Since both TagA and TagB are functional in the absence of detergents and utilize substrates containing short lipid chains, we conclude that neither enzyme requires a membrane interface for activity even though their natural substrates are membrane anchored.

TagA and TagB are potential antibiotic targets, but they are also of interest for other reasons. For example, TagA catalyzes a transformation—attachment of a sugar to a lipid-anchored monosaccharide—which is analogous to that performed by the glycosyltransferase MurG in the biosynthesis of bacterial peptidoglycan (23–24) and also by the second enzyme in the dolichol pathway for N-linked glycosylation (25). Sequence alignments suggest that TagA is unrelated to these or other glycosyltransferases (Gtfs) for which structural information exists, and TagA may thus represent an uncharacterized Gtf superfamily. TagB belongs to a class of glycerophosphotransferases that contains both processive and nonprocessive enzymes. There is no detailed structural or mechanistic information for any member of this class of enzymes, and information on TagB may shed light on the broader family. Furthermore, we anticipate that TagB can be used to make acceptor substrates for TagF, an unusual processive glycerophosphotransferase belonging the same enzyme superfamily as TagB (26, 27). Detailed structural and mechanistic studies of TagA and TagB are now underway.

METHODS

General Methods. Chemicals and solvents were from Sigma-Aldrich. Silica gel (60 Å, 32–63 μm) was from Sorbent Technologies. NMR spectra were recorded on Varian Inova 400 or 500 MHz spectrometers. Mass spectra (ESI) were recorded using an Agilent 1100 series LC/MSD instrument with an electrospray ionization (ESI) source in negative ion mode. LC/MS analysis of enzymatic reactions was also performed on this instrument, using a Zorbax 300-SB-C18 column for LC separation.

Synthesis of Compound 3. The triethylammonium salt of **2** (Scheme 1; 0.15 g, 0.239 mmol), prepared following published methods (18, 19), and uridine 5'-monophosphomorpholidate (0.534 g, 0.774 mmol) were dissolved in 6 mL of dry pyridine. ¹H-Tetrazole (0.050 g, 0.717 mmol) was added, the reaction was stirred for 48 h at room temperature (RT), the solvent was removed, and the residue was purified over a C-18 column (0–10% EtOH in 0.1% aqueous NH_4HCO_3) to give the peracylated precursor of **3**. (Spectral details are provided in Supporting Information.) This compound (0.06 g, 0.077 mmol) was dissolved in 4 mL of dry MeOH, and 0.46 mL of NaOMe (0.228 mmol) was added. The reaction was stirred for 1 h at RT and quenched with $\text{CH}_3\text{COONH}_4$ (1 mL, 0.5 M in H_2O), the solvent was removed, and the residue was purified over a C-18 column (0–10% EtOH in 0.1% aqueous NH_4HCO_3) to give **3** in 30% total yield from **8**.

Synthesis of Compound 6. The peracylated precursor of **6** was prepared following a published route (Scheme 1; 14–17). This compound was dissolved in 6 mL of dry MeOH, and 1.4 mL of NaOMe (0.705 mmol) was added. After 1 h at RT, the reaction was quenched with $\text{CH}_3\text{COONH}_4$ (2.8 mL, 0.5 M in H_2O), the solvent was removed, and the residue was purified over a C-18 column using a gradient of 0–45% EtOH in 0.1% aqueous NH_4HCO_3 to give **6** in 79% yield.

Preparation of Compound 7. To a solution of TagA (1.5 mL, 1 mg mL^{-1}) in 20 mM Tris buffer (pH 7.9, 0.5 M NaCl, 200 mM imidazole) was added compound **6** (1.5 mg, 2.5 mmol), UDP-ManNAc (3.3 mg, 5.0 mmol), and 40 μL of alkaline phosphatase (20 U μL^{-1}). After 2.5 h at RT, the reaction was purified over a C-18 column 0–50% EtOH in 0.1% aqueous NH_4HCO_3 to give **7**.

Cloning, Expression, and Purification of TagA. The *tagA* gene was amplified from *Bacillus subtilis* strain PY79 genomic DNA using two rounds of PCR. Primers 5'-TGATTTTCTTTAGAACTCTCG-3'

and 5'-TTTCAGTTCACCCATTGC-3' were used to amplify a fragment containing *tagA*. The gene was then amplified using primers 5'-GTGAGAATTCGATGCAACAGAGACTATTAC-3' and 5'-GTGACTCGAGAATCTGTTTGTATGATCTTTTC-3' and cloned into the *EcoRI* and *XhoI* sites (underlined) of pET-24b(+) (Novagen). TagA was expressed in mid-log phase cultures of *E. coli* strain Rosetta(DE3)pLysS (Novagen), induced with 0.5 mM IPTG for 3 h at 37 °C. Cells were lysed by freeze-thaw and resuspended in 40 mL of buffer (20 mM Tris-HCl, pH 7.9, 5 mM imidazole, 0.5 M NaCl) containing 20 μ L of Benzonase (25 U μ L⁻¹), 40 μ L of protease inhibitor cocktail, and 40 μ L of Triton X-100. The clarified lysate was purified over a Ni²⁺ affinity column (Novagen His-Bind resin) yielding 4 mg L⁻¹ TagA.

Cloning, Expression, and Purification of TagB.

The *tagB* gene was amplified from *Bacillus subtilis* PY79 genomic DNA using primers 5'-GAGCAT-GTCGCTAGCATGAAAATAAGATCACTACTGG-3' and 5'-CACTGCAGTCTCGAGGCTTATAAATTTTC-GATGAAATT-3'. TagB was cloned into the *NheI* and *XhoI* sites pET-24b(+) (Novagen). TagB was expressed in mid-log phase cultures of the *E. coli* strain BL21(DE3) (Novagen), induced with 0.5 mM IPTG for 22 h at 16 °C. Cells were lysed with Bugbuster (Novagen) supplemented with Benzonase (Novagen), protease inhibitor cocktail, and lysozyme following the Novagen protocol. Clarified lysate was purified over a Ni²⁺ column (Novagen IDA His-Bin resin) yielding 12 mg L⁻¹ of TagB.

Continuous Coupled Enzyme Assay for TagA.

This assay couples the production of UDP to the oxidation of NADH, leading to a drop in NADH fluorescence (15, 17, 20, 21). Reactions were carried out in 96-well plate microplates (Costar 3603) in a volume of 100 μ L. NADH fluorescence was monitored at 465 nM using a Perkin-Elmer HTS 7000 Plus BioAssay reader. Reactions contained buffer (20 mM Tris, pH 7.9, 5 mM MgCl₂, 450 mM NaCl), 0.4 U μ L⁻¹ PK, 0.2 U μ L⁻¹ NDPK, 0.5 mM PEP, 0.2 U μ L⁻¹ LDH, 0.25 mM NADH, 0.5 mM ATP, and 0.1 mg mL⁻¹ BSA. Reaction mixtures containing substrates were incubated at RT for 15 min prior to adding TagA (100 nM). Kinetic parameters for TagA with substrate analogue **6** were calculated as described in the Supporting Information.

HPLC Assay for TagB. TagB (10 nM) was incubated at RT with 200 mM **7** and 400 mM CDP-glycerol in buffer (20 mM Tris, pH 7.5, 100 mM NaCl, and 5 mM MgCl₂). Reactions were quenched with an equal volume of cold methanol, loaded onto an analytical anion exchange HPLC column (Phenosphere SAX), and eluted with a gradient of 0–56% buffer B over 17 min (buffer A, 5 mM NH₄H₂PO₄ pH 2.8; buffer B, 750 mM NH₄PO₄ pH 3.7). CMP was monitored at 260 nm.

Acknowledgment: We thank D. Rudner for providing purified *B. subtilis* PY79 genomic DNA. This work was supported in part by NIH Grant A144854.

Supporting Information Available: Complete synthetic and spectral data for compounds **3**, **6**, and **7**, and enzyme assay details. This material is available free of charge via the Internet.

REFERENCES

- Walsh, C. T. (2003) *Antibiotics: Actions, Origins, Resistance*, American Society of Microbiology Press, Washington, DC.
- Neuhaus, F. C., and Baddiley, J. (2003) A Continuum of Anionic Charge: Structures and Functions of D-Alanyl-Teichoic Acids in Gram-Positive Bacteria, *Microbiol. Mol. Biol. Rev.* 67, 686–723.
- Bhavsar, A. P., Erdman, L. K., Schertzer, J. W., and Brown, E. D. (2004) Teichoic Acid is an Essential Polymer in *Bacillus subtilis* that is Functionally Distinct from Teichuronic Acid, *J. Bacteriol.* 186, 7865–7873.
- Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., et al. (2003) Essential *Bacillus subtilis* Genes, *Proc. Natl. Acad. Sci. U.S.A.* 100, 4678–4683.
- Weidenmaier, C., Kokai-Kun, J. F., Kristian, S. A., Chanturiya, T., Kalbacher, H., Gross, M., et al. (2004) Role of Teichoic Acids in *Staphylococcus aureus* Nasal Colonization, a Major Risk Factor in Nosocomial Infections, *Nat. Med.* 10, 243–245.
- Weidenmaier, C., Peschel, A., Xiong, Y. Q., Kristian, S. A., Dietz, K., Yeaman, M. R., et al. (2005) Lack of Wall Teichoic Acids in *Staphylococcus aureus* Leads to Reduced Interactions with Endothelial Cells and to Attenuated Virulence in a Rabbit Model of Endocarditis, *J. Infect. Dis.* 191, 1771–1777.
- Soldo, B., Lazarevic, V., and Karamata, D. (2002) TagO is Involved in the Synthesis of All Anionic Cell-Wall Polymers in *Bacillus subtilis* 168, *Microbiology* 148, 2079–2087.
- Mauel, C., Young, M., Monsutti-Grecescu, A., Marriott, S. A., and Karamata, D. (1994) Analysis of *Bacillus subtilis* tag Gene Expression Using Transcriptional Fusions, *Microbiology* 140, 2279–2288.
- Pooley, H. M., Abellan, F. X., and Karamata, D. (1992) CDP-Glycerol:Poly(Glycerophosphate) Glycerophosphotransferase, which is Involved in the Synthesis of the Major Wall Teichoic Acid in *Bacillus subtilis* 168, is Encoded by *tagF* (*rodC*), *J. Bacteriol.* 174, 646–649.
- Pooley, H. M., Abellan, F. X., and Karamata, D. (1991) A Conditional-Lethal Mutant of *Bacillus subtilis* 168 with a Thermosensitive Glycerol-3-Phosphate Cytidyltransferase, an Enzyme Specific for the Synthesis of the Major Wall Teichoic Acid, *J. Gen. Microbiol.* 137, 921–928.
- Lazarevic, V., Abellan, F. X., Moller, S. B., Karamata, D., and Mauel, C. (2002) Comparison of Ribitol and Glycerol Teichoic Acid Genes in *Bacillus subtilis* W23 and 168: Identical Function, Similar Divergent Organization, but Different Regulation, *Microbiology* 148, 815–824.
- El Ashry, E. S. H., Rashed, N., and Ibrahim, E. S. I. (2005) Strategies of Synthetic Methodologies for Constructing Beta-Mannosidic Linkages, *Curr. Org. Syn.* 2, 175–213.
- Gridley, J. J., and Osborn, H. M. I. (2000) Recent Advances in the Construction of Beta-D-Mannose and Beta-D-Mannosamine Linkages, *J. Chem. Soc., Perkin Trans. 1* 10, 1471–1491.
- Sim, M. M., Kondo, H., and Wong, C.-H. (1993) Synthesis and Use of Glycosyl Phosphites—An Effective Route to Glycosyl Phosphates, Sugar Nucleotides, and Glycosides, *J. Am. Chem. Soc.* 115, 2260–2267.
- Chen, L., Men, H., Ha, S., Ye, X. Y., Brunner, L., Hu, Y., et al. (2002) Intrinsic Lipid Preferences and Kinetic Mechanism of *Escherichia coli* MurG, *Biochemistry* 41, 6824–6833.
- Ye, X. Y., Lo, M. C., Brunner, L., Walker, D., Kahne, D., and Walker, S. (2001) Better Substrates for Bacterial Transglycosylases, *J. Am. Chem. Soc.* 123, 3155–3156.
- Liu, H., Ritter, T. K., Sadamoto, R., Sears, P. S., Wu, M., and Wong, C.-H. (2003) Acceptor Specificity and Inhibition of the Bacterial Cell-Wall Glycosyltransferase MurG, *ChemBioChem* 4, 603–609.
- Yamazaki, T., Warren, C. D., Herscovics, A., and Jeanloz, R. W. (1980) Convenient Synthesis of Uridine 5'-(2-Acetamido-2-Deoxy-Alpha-D-Mannopyranosyluronic Acid Pyrophosphate), *Carbohydr. Res.* 79, C9–C12.
- Freese, S. J., and Vann, W. F. (1996) Synthesis of 2-Acetamido-3-O-Acetyl-2-Deoxy-D-Mannose Phosphoramidites, *Carbohydr. Res.* 281, 313–319.
- Huang, Y. H., Picha, D. H., and Killili, A. W. (1999) A Continuous Method for Enzymatic Assay of Sucrose Synthase in the Synthetic Direction, *J. Agric. Food Chem.* 47, 2746–2750.
- Gosselin, S., Alhussaini, M., Streiff, M. B., Takabayashi, K., and Palcic, M. M. (1994) A Continuous Spectrophotometric Assay for Glycosyltransferases, *Anal. Biochem.* 220, 92–97.
- Bhavsar, A. P., Truant, R., and Brown, E. D. (2005) The TagB Protein in *Bacillus subtilis* 168 is an Intracellular Peripheral Membrane Protein that can Incorporate Glycerol-Phosphate onto a Membrane Bound Acceptor *In Vitro*, *J. Biol. Chem.* 280, 36691–36700.
- Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) The 1.9 Å Crystal Structure of *Escherichia coli* MurG, a Membrane-Associated Glycosyltransferase Involved in Peptidoglycan Biosynthesis, *Protein Sci.* 9, 1045–1052.
- Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., et al. (2003) Crystal Structure of the MurG:UDP-GlcNAc Complex Reveals Common Structural Principles of a Superfamily of Glycosyltransferases, *Proc. Natl. Acad. Sci. U.S.A.* 100, 845–849.
- Chantret, I., Dancourt, J., Barbat, A., and Moore, S. E. (2005) Two Proteins Homologous to the N- and C-terminal Domains of the Bacterial Glycosyltransferase MurG are Required for the Second Step of Dolichyl-Linked Oligosaccharide Synthesis in *Saccharomyces cerevisiae*, *J. Biol. Chem.* 280, 9236–9242.
- Schertzer, J. W., Bhavsar, A. P., and Brown, E. D. (2005) Two Conserved Histidine Residues are Critical to the Function of the TagF-Like Family of Enzymes, *J. Biol. Chem.* 280, 36683–36690.
- Schertzer, J. W., and Brown, E. D. (2003) Purified, Recombinant TagF Protein from *Bacillus subtilis* 168 Catalyzes the Polymerization of Glycerol Phosphate onto a Membrane Acceptor *In Vitro*, *J. Biol. Chem.* 278, 18002–18007.

Enzymatic Tailoring of Enterobactin Alters Membrane Partitioning and Iron Acquisition

Minkui Luo[†], Hening Lin[‡], Michael A. Fischbach^{‡,§}, David R. Liu[§], Christopher T. Walsh^{*,*}, and John T. Groves^{†,*}

[†]Department of Chemistry, Princeton University, Princeton, New Jersey 08544, [‡]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and

[§]Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Iron is essential for growth in nearly all bacteria. In microenvironments where iron (largely present as Fe³⁺ in aerobic settings) is limiting, bacteria respond by activating genes required for synthesis and export of iron chelators known as siderophores, as well as membrane receptors and other proteins for the subsequent import of iron–siderophore complexes (1, 2). Many siderophores are small, nonribosomal peptide scaffolds containing catechols, hydroxamates, α -hydroxy acids, and similar bidentate functional groups to chelate ferric iron (1, 2).

A prime example of a catechol-containing siderophore is enterobactin (Ent; Figure 1), produced by Gram-negative enteric bacteria such as *Escherichia coli* and *Salmonella typhimurium*. With a K_d of 10⁻⁴⁹ M for the hexadentate coordination of Fe³⁺ (3), Ent, the cyclic trilactone of *N*-2,3-dihydroxybenzoyl-L-serine, appears admirably engineered for removing ferric iron from vertebrate proteins such as transferrin during infection. However, the mammalian proteins serum albumin (4) and siderocalin (5) bind apo- and ferric (Fe)-Ent, respectively, thereby suppressing bacterial growth in various mammalian microenvironments. Siderocalin can also bind several other siderophores in the same site (6).

Various pathogenic strains of *E. coli* and *Salmonella* that harbor the *iroA* gene cluster can overcome the antimicrobial effects of serum proteins by enzymatic tailoring of Ent (7, 8). Modification of the Ent scaffold is effected by the *C*-glucosyl-

transferase IroB, which transfers a glucosyl moiety from UDP-glucose to C5 of each of the 2,3-dihydroxybenzoyl rings of Ent. This yields monoglucosyl enterobactin (MGE) and the corresponding diglucosyl (DGE) and triglucosyl (TGE) forms of the siderophore (Figure 1) (9). DGE, but not MGE or TGE, has been detected in the culture broth of *iroA*-harboring *Salmonella* (7, 8).

There are at least three consequences of siderophore *C*-glucosylation. First, the unusual *C*-glycosidic linkage is stable to hydrolysis in contrast to an *O*-glycosidic linkage. Second, the hydrophobicity of the scaffold is decreased by the hydrophilic glucosyl moieties in MGE and DGE. Third, and perhaps most importantly, glucosylation is likely to block binding and sequestration of Fe-MGE and Fe-DGE by siderocalin (5, 10), leaving these siderophores available for import by bacterial cells. We have suggested that the IroB-mediated tailoring of the periphery of enterobactin may be a bacterial counterattack against the host's innate immune system (9).

The five-gene *iroA* cluster also consists of *iroC*, *iroN*, *iroD*, and *iroE*. IroC and IroN are thought to be involved in export of apo-MGE/DGE and/or uptake of Fe-MGE/DGE. We have demonstrated that IroD is a cytoplasmic esterase that hydrolyzes both apo- and Fe-MGE/DGE to fragments with lower affinity for Fe³⁺ (11). Likewise, we have shown that IroE is a hydrolase, but unlike IroD, IroE is periplasmic and cleaves apo-MGE/DGE only once to produce linearized versions of these

ABSTRACT Enterobactin (Ent), a prototypic bacterial siderophore, is modified by both the *C*-glucosyltransferase IroB and the macrolactone hydrolase IroE in pathogenic bacteria that contain the *iroA* cluster. To investigate the possible effects of glucosylation and macrolactone hydrolysis on the physical properties of Ent, the membrane affinities and iron acquisition rates of Ent and Ent-derived siderophores were measured. The data obtained indicate that Ent has a high membrane affinity ($K_x = 1.5 \times 10^4$) similar to that of ferric acinetoferrin, an amphiphile containing two eight-carbon hydrophobic chains. Glucosylation and macrolactone hydrolysis decrease the membrane affinity of Ent by 5–25-fold. Furthermore, in the presence of phospholipid vesicles, the iron acquisition rate is significantly increased by glucosylation and macrolactone hydrolysis, due to the resultant decrease in membrane sequestration of the siderophore. These results suggest that IroB and IroE enhance the ability of Ent-producing pathogens to acquire iron in membrane-rich microenvironments.

*To whom correspondence should be addressed.

E-mail: christopher_walsh@hms.harvard.edu; jtgroves@princeton.edu.

Received for review September 28, 2005 and accepted January 3, 2006

Published online January 27, 2006

10.1021/cb0500034 CCC: \$33.50

© 2006 by American Chemical Society

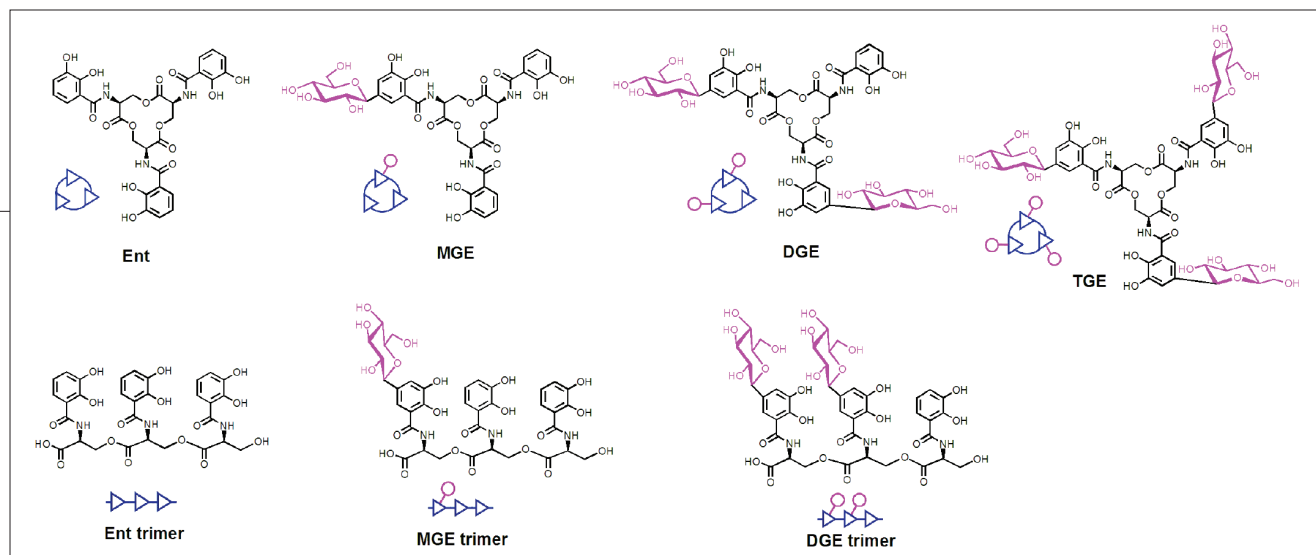


Figure 1. Chemical structures of Ent family siderophores used in this study. IroB-catalyzed glycosylation of Ent produces mono-, di-, and tri-glucosylated Ent (MGE, DGE, and TGE, respectively). In turn, IroD-catalyzed hydrolysis of Ent, MGE, and DGE produces their linear trimers.

siderophores (Figure 2; 11). Indeed, a suite of siderophore species is found in the culture broth of *iroA*-harboring *Salmonella*; these include macrocyclic and linearized Ent and DGE and smaller hydrolytic fragments of each (8, 12). The linear trimer of Ent also has high affinity for ferric iron ($\Delta G = -8.9 \text{ kcal mol}^{-1}$), although with a higher entropic barrier for hexadentate coordination (13, 14). The biological significance of IroE's enzymatic action (namely, why the bacteria would want to linearize the siderophores prior to their release) is thus very interesting. We hypothesized that IroB/IroE-catalyzed tailoring of Ent by producing pathogens could affect Fe^{3+} binding kinetics and the hydrophobicity of this siderophore. In this study, we used

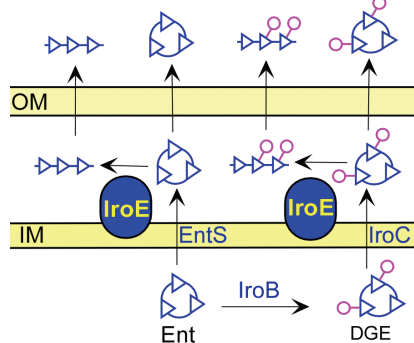


Figure 2. The tandem action of IroB and IroE creates a suite of siderophores that are secreted into the culture medium. IroB glycosylates Ent in the cytoplasm, forming DGE. Ent is transported to the periplasm by EntS, while IroC is proposed to transport DGE to the same compartment. IroE then hydrolyzes these trilactones, generating their linearized derivatives.

a membrane partitioning assay we have recently developed to investigate both possibilities (15, 16).

The membrane partition coefficients (K_x) of the apo-siderophores were determined from observed changes in the rates of iron acquisition (k_{obs}) as the phospholipid membrane concentration of the medium was varied. The k_{obs} values for all siderophore species tested are similar ($0.033\text{--}0.058 \text{ mM}^{-1} \text{ s}^{-1}$) (Table 1). The rates of iron acquisition by Ent and its derivatives were determined to be due to a direct interaction of the siderophores with iron citrate. An alternative mechanism involving rate-limiting dissociation of iron from the iron citrate cluster can be ruled out on several grounds. The series of experiments were carried out with a 20-fold excess of iron. Although rate-limiting iron release from iron citrate has been reported, such a process was observed only in the presence of high siderophore to iron ratios (17). The iron-chelation kinetics of Ent and its derivatives reported here show excellent pseudo-first-order fits of concentration of chelator versus time with a half-time around 20 s (Supplementary Figure 1). By contrast, an iron-dissociation-controlled process would be zero-order in concentration of Ent and would show a linear relationship with a half-life of several hours (18). Consequently, we conclude that glycosylation, linearization, and the attendant introduction of a carboxylate functionality do not hinder the facility of Fe^{3+} ligation by Ent.

In contrast, the membrane partition coefficients (K_x) showed substantial variation among the siderophore species. Ent shows a remarkable affinity for membranes with a K_x value of 15000, similar to the value we have found for Fe-acinetoferrin, a siderophore amphiphile with two eight-carbon side chains (16). The addition of one (MGE) or two (DGE) C-glycosyl units decreased K_x only by ~5-fold, and addition of three (TGE) C-glycosyl units decreased K_x to about 1/10 that of Ent, consistent with the hydrophilic nature of the sugars. Correspondingly, monohydrolysis of the Ent trilactone to a linear trimer with one carboxylate decreased its K_x 25-fold, and hydrolysis of DGE decreased its K_x more than 10-fold. It has been observed that Ent, after being synthesized, is not secreted efficiently because it accumulates in the periplasm (19). The high K_x value of Ent, which was unrecognized prior to this report, could partially explain this observation and suggests that glycosylation by IroB and macrolactone linearization by IroE could be strategies to increase secretion efficiency.

We further show that Ent binds ferric iron significantly more slowly in the presence of lipid membranes than it does in membrane-free aqueous solution, due to the high partition ratio into the membrane phase and out of the aqueous medium (Table 1 and Figure 3). This decrease in iron acquisition rate is a consequence of the lower mole fraction of Ent in aqueous solution and, thereby, inefficient access to the iron source. This result is consistent with

TABLE 1. Ferric iron acquisition rate constants and membrane partitioning coefficients

	Ent	Ent trimer	MGE	MGE trimer ^d	DGE	DGE trimer ^d	TGE
Iron acquisition rate constant (k_{obs} , $\text{mM}^{-1} \text{s}^{-1}$) ^b	0.041	0.058	0.042	0.048	0.045	0.048	0.033
Membrane partitioning coefficient (K_x)	15000	490	3400	640	3100	230	1400
Relative iron acquisition rate with 10 mM lipid ^c	0.27	0.92	0.62	0.90	0.64	0.96	0.81

^aThe MGE and DGE linear trimers were obtained by IroD-catalyzed regioselective hydrolysis of MGE and DGE and are different from the linear trimers that can be obtained with IroE, which is not regioselective. However, we obtained similar results using the mixture of linear trimers generated by IroE-catalyzed MGE or DGE hydrolysis. ^bThis is the iron acquisition rate measured in the absence of phospholipids vesicles. ^cThese data were obtained by dividing the iron acquisition rate in the presence of lipid by that in the absence of lipid. The iron acquisition rates in the absence of lipid are defined as 1.

the reported observation that *E. coli* strains producing Ent but not aerobactin scavenge transferrin-bound (extracellular) iron more efficiently than cellular iron (20). Hydrolytic linearization and glucosylation suppress partitioning of Ent into the membrane phase and result in higher rates of iron acquisition. Therefore, it may be advantageous for siderophore-producing bacteria to secrete a suite of iron chelators that cover a range of membrane affinities and hydrophobicities.

The tandem action of IroB and IroE creates just such a suite of siderophores. These glucosylated and linearized Ent derivatives partition more efficiently into the aqueous phase and thus may forage more effectively for ferric iron in a mammalian host. For the marinobactin (15, 21), aquachelin (21), and amphibactin (22) siderophores, a convergent tailoring strategy is employed. These hydrophilic tetra- to hexapeptide scaffolds are enzymatically acylated at their amino termini, thus converting them to lipophilic amphiphiles through the introduction of fatty acyl substituents to the iron-binding peptide core (23). The acyltransferases responsible for these modifications evidently show

promiscuity for the acyl-ACP substrate, again to create a suite of siderophores with a range of hydrophobicities. Similarly, mycobacteria produce siderophores with a range of hydrophobicities (24). The most hydrophobic of these, mycobactin, has been shown recently to permeate cell

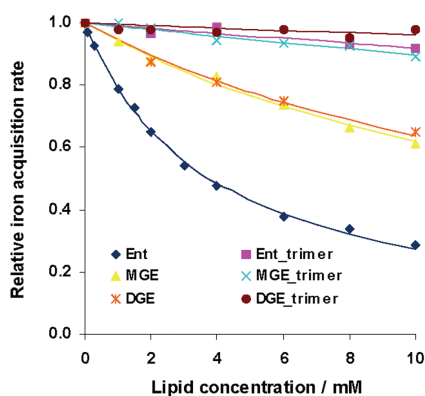


Figure 3. The iron-acquisition kinetics and membrane partitioning of Ent-derived siderophores. Relative Fe^{3+} binding rates of Ent-derived siderophores (50 μM) from 1 mM ferric ammonium citrate at 37 °C and pH 7.5 were determined in the presence of varied lipid vesicle concentration by monitoring the appearance of the ferric catecholate chromophore at 490 nm.

membranes to extract iron from target cells such as macrophages (25). It is also clear that profound changes in the conformation of siderophores such as acinetoferrin (26) and rhizobactin (27) upon binding iron transform the properties of these molecules such that while the apo-forms are tuned for iron prospecting, the iron-bound forms become homing devices for the bacterial membrane receptor. Taken together with the *iroA* system, bacteria have elaborated a series of enzymatic tailoring strategies to control the hydrophobic/hydrophilic balance of siderophores. The capacity to add either hydrophilic (glucosyl) or hydrophobic (fatty acyl) groups to nonribosomal peptide scaffolds represents two convergent strategies to titrate the physical properties of a set of siderophores that scavenge the essential ferric iron nutrient.

METHODS

Preparation of Siderophores. Ent, MGE, DGE, TGE, and the linearized trimers from IroD-catalyzed MGE/DGE hydrolysis were prepared as described previously (11). IroD was used because it catalyzes the hydrolysis of MGE/DGE regioselectively and gives only one linear trimer isomer as the major hydrolysis product. Ent linear trimer was prepared using IroE N-30 catalyzed hydrolysis because there is no regioselectivity issue and IroE affords almost exclusively the linear trimer product. Since IroD and IroE give different regioisomers of linear trimer products for MGE/DGE hydrolysis, we also prepared MGE/DGE linear trimers using IroE-catalyzed hydrolysis to make sure that different regioisomers have similar membrane affinities and iron acquisition rates. The hydrolysis of Ent/MGE/DGE with IroE was carried out with 128 mM Ent/MGE/DGE, 40 nM IroE N-30, in 50 mL of 75 mM HEPES buffer pH 7.5 for 1 h. The reaction mixture was quenched with 25 mL of 2.5 N HCl in methanol (prepared by mixing 10 mL of concentrated HCl with 40 mL of methanol), and the hydrolysis products were purified by reverse phase HPLC using a gradient of 0–40% acetonitrile, with the aqueous phase containing 0.1% (v/v) trifluoroacetic acid. The HPLC fractions were lyophilized, and the linear trimer products were dissolved in DMSO. The concentrations of the resulting solutions were determined using HPLC by co-injecting equal volumes of the solutions with a known concentration Ent solution and comparing the areas of absorption at 316 nm.

Preparation of Lipid Vesicles. Unilamellar vesicles were prepared as previously described (15, 16) by sonication for small unilamellar vesicles (SUV) with a diameter of 30–40 nm. Briefly, weighed 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC, from Avanti Polar

Lipids) was dissolved in chloroform and then transferred into 5-mL test tubes. The thin films of lipid were deposited on the walls of the test tubes after evaporating the solvent with a stream of argon and then were subjected to high vacuum overnight. The dried lipid films were hydrated in HEPES buffer (100 mM HEPES, 150 mM NaCl) at 40 °C for 30 min. The pretreated lipid buffers were sonicated with a probe tip sonicator in an ice–water bath until the ice was melted completely and the bath temperature had reached room temperature (taking about 30 min). After centrifugation of the vesicle suspension at 12,000 rpm for 5 min to remove the sonicator tip debris, a translucent SUV suspension was obtained.

Measurement of Iron-Acquisition Rates and Membrane Affinities. A Hi-Tech SF-61 DX2 stopped-flow spectrophotometer with the photomultiplier absorbance mode was used for kinetic measurements of siderophore-mediated iron acquisition from ferric ammonium citrate (FAC, from Aldrich-Sigma Inc.) at 37 °C. Briefly, one mixing syringe was filled with FAC stock solution and the other one was loaded with the siderophore solution containing various concentrations of lipid vesicles (1 mM FAC, 50 μ M siderophore, and 0–10 mM DMPC SUV for final concentrations). Upon mixing, the changes of the absorbance at 490 nm were measured versus time. The pseudo-first-order rate constants were obtained upon fitting a general exponential equation (eq 2) using Scientist software. The second-order rate constants (k_{obs} in Table 1) were derived by dividing the pseudo-first-order rate constants by the concentration of FAC (1 mM).

The partition coefficients (K_x) of all compounds tested were obtained on the basis of the kinetic relationship (eq 1.1) between the relative rates for siderophore-mediated iron acquisition from FAC (k_{obs}/k_w) and the DMPC SUV concentration (Vesicle). Here α is the ratio (k_x/k_w) of the lipid-phase iron-mobilization rate (k_x) to the aqueous-phase rate (k_w); k_w and k_{obs} correspond to the rates of siderophore-mediated iron acquisition from FAC in the absence or presence of the various concentrations of DMPC SUV, respectively. k_w and k_{obs} were derived from eq 2, where A_t is the absorbance changes at 490 nm versus time t , A_{obs} is the maximal absorption change, and C is the initial absorption. Equation 1.1 is derived from eq 1.2, which we have reported previously (16). The difference between eqs 1.1 and 1.2 lies in the value of α , which is set as zero for the former and a variable parameter for the latter. For Ent, the two equations give results almost identical with $K_x = 1.5 \times 10^4$ for eq 1.1 (Figure 3), and $K_x = 1.6 \times 10^4$ and near-zero α ($\alpha = 0.03$, data not shown) for eq 1.2. For DGE, MGE, DGE trimer, and MGE trimer, only eq 1.1 can produce a reasonable fit because of the relatively low membrane affinities of these siderophores. Consequently, all partition coefficients were obtained on the basis of eq 1.1, in which $\alpha = 0$ was arbitrarily used. This lipid-concentration-dependent hyperbolic relationship arises from siderophore-mediated membrane partitioning, since the rate of iron mobilization

from FAC is close to zero in the lipid phase in contrast to non-zero k_w in the aqueous phase.

$$\frac{k_{\text{obs}}}{k_w} = \frac{[\text{Water}]}{K_x \times [\text{Vesicle}] + [\text{Water}]} \quad (1.1)$$

$$\frac{k_{\text{obs}}}{k_w} = (\alpha - 1) \frac{K_x \times [\text{Vesicle}]}{K_x \times [\text{Vesicle}] + [\text{Water}]} \quad (1.2)$$

$$A_t = A_{\text{obs}} [1 - \exp(-k_{\text{obs}} \times t)] + C \quad (2)$$

Acknowledgment: This work was supported in part by the National Science Foundation through the Environmental Molecular Science Institute, CEBIC at Princeton University, CHE-0221978 (J.T.G.) and CHE-0316301 (J.T.G.), NIH AI 47238 (C.T.W.), NIH R01GM065400 (D.R.L.), a post-doctoral fellowship from the Jane Coffin Childs Memorial Fund (H.L.), and a graduate fellowship from the Hertz Foundation (M.A.F.).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Neilands, J. B. (1995) Siderophores: Structure and function of microbial iron transport compounds, *J. Biol. Chem.* **270**, 26723–26726.
- Crosa, J. H., and Walsh, C. T. (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria, *Microbiol. Mol. Biol. Rev.* **66**, 223–249.
- Loomis, L. D., and Raymond, K. N. (1991) Solution equilibria of enterobactin and metal-enterobactin complexes, *Inorg. Chem.* **30**, 906–911.
- Konopka, K., and Neilands, J. B. (1984) Effect of serum albumin on siderophore-mediated utilization of transferrin iron, *Biochemistry* **23**, 2122–2127.
- Goetz, D. H., Holmes, M. A., Borregaard, N., Bluhm, M. E., Raymond, K. N., and Strong, R. K. (2002) The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition, *Mol. Cell.* **10**, 1033–1043.
- Holmes, M. A., Paulsens, W., Jide, X., Ratledge, C., and Strong, R. K. (2005) Siderocalin (Lcn 2) also binds carboxymycobactins, potentially defending against mycobacterial infections through iron sequestration, *Structure* **13**, 29–41.
- Hantke, K., Nicholson, G., Rabusch, W., and Winkelmann, G. (2003) Salmochelins, siderophores of *Salmonella enterica* and uropathogenic *Escherichia coli* strains, are recognized by the outer membrane receptor IroN, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3677–3682.
- Bister, B., Bischoff, D., Nicholson, G. J., Valdebenito, M., Schneider, K., Winkelmann, G., Hantke, K., and Sussmuth, R. D. (2004) The structure of salmochelins: C-glycosylated enterobactins of *Salmonella enterica*, *Biometals* **17**, 471–481.
- Fischbach, M. A., Lin, H., Liu, D. R., and Walsh, C. T. (2005) In vitro characterization of IroB, a pathogen-associated C-glycosyltransferase, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 571–576.
- Flo, T. H., Smith, K. D., Sato, S., Rodriguez, D. J., Holmes, M. A., Strong, R. K., Akira, S., and Aderem, A. (2004) Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron, *Nature* **432**, 917–921.
- Lin, H., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2005) In vitro characterization of salmochelin and enterobactin trilactone hydrolases IroD, IroE, and Fes, *J. Am. Chem. Soc.* **127**, 11075–11084.
- Zhu, M., Valdebenito, M., Winkelmann, G., and Hantke, K. (2005) Functions of the siderophore esterases IroD and IroE in iron-salmochelin utilization, *Microbiology* **151**, 2363–2372.
- Scarrow, R. C., Ecker, D. J., Ng, C., Liu, S., and Raymond, K. N. (1991) Iron(III) coordination chemistry of linear dihydroxyserine compounds derived from enterobactin, *Inorg. Chem.* **30**, 900–906.
- O'Brien, I. G., Cox, G. B., and Gibson, F. (1971) Enterobactin hydrolysis and iron metabolism in *Escherichia coli*, *Biochim. Biophys. Acta* **237**, 537–549.
- Xu, G., Martinez, J. S., Groves, J. T., and Butler, A. (2002) Membrane affinity of the amphiphilic marinobactin siderophores, *J. Am. Chem. Soc.* **124**, 13408–13415.
- Luo, M., Fadeev, E. A., and Groves, J. T. (2005) Membrane dynamics of the amphiphilic siderophore acinetoferrin, *J. Am. Chem. Soc.* **127**, 1726–1736.
- Faller, B., and Nick, H. (1994) Kinetics and mechanism of iron(III) removal from citrate by desferrioxamine B and 3-hydroxy-1,2-dimethyl-4-pyridone, *J. Am. Chem. Soc.* **116**, 3860–3865.
- Bates, G. W., Billups, C., and Saltman, P. (1967) The kinetics and mechanism of iron(III) exchange between chelates and transferrin, *J. Biol. Chem.* **242**, 2810–2815.
- Vartanian, M. D. (1988) Differences in excretion and efficiency of the aerobactin and enterochelin siderophores in a bovine pathogenic strain of *Escherichia coli*, *Infect. Immun.* **56**, 413–418.
- Brock, J. H., Williams, P. H., Liceaga, J., and Wooldridge, K. G. (1991) Relative availability of transferrin-bound iron and cell-derived iron to aerobactin-producing and enterochelin-producing strains of *Escherichia coli* and to other microorganisms, *Infect. Immun.* **59**, 3185–3190.
- Martinez, J. S., Zhang, G. P., Holt, P. D., Jung, H.-T., Carrano, C. J., Haygood, M. G., and Butler, A. (2000) Self-assembling amphiphilic siderophores from marine bacteria, *Science* **287**, 1245–1247.
- Martinez, J. S., Carter-Franklin, J. N., Mann, E. L., Martin, J. D., Haygood, M. G., and Butler, A. (2003) Bioinorganic Chemistry Special Feature: Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3754–3759.
- Groves, J. T. (2003) Bioinorganic Chemistry Special Feature: The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3569–3574.
- Ratledge, C. (2004) Iron, mycobacteria and tuberculosis, *Tuberculosis* **84**, 110–130.
- Luo, M., Fadeev, E. A., and Groves, J. T. (2005) Mycobactin-mediated iron acquisition within macrophages, *Nat. Chem. Biol.* **1**, 149–153.
- Fadeev, E. A., Luo, M., and Groves, J. T. (2004) Synthesis, structure, and molecular dynamics of gallium complexes of schizokinen and the amphiphilic siderophore acinetoferrin, *J. Am. Chem. Soc.* **126**, 12065–12075.
- Fadeev, E. A., Luo, M., and Groves, J. T. (2005) Synthesis and structural modeling of the amphiphilic siderophore rhizobactin-1021 and its analogs, *Bioorg. Med. Chem. Lett.* **15**, 3771–3774.

Redefining the Requisite Lipopolysaccharide Structure in *Escherichia coli*

Timothy C. Meredith[†], Parag Aggarwal[†], Uwe Mamat[‡], Buko Lindner[‡], and Ronald W. Woodard^{†,*}

[†]Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, 428 Church Street, Ann Arbor, Michigan 48109-1065, and [‡]Department of Immunochemistry and Biochemical Microbiology, Research Center Borstel, Leibniz Center for Medicine and Biosciences, Parkallee 10, D-23845 Borstel, Germany

Gram-negative bacteria are defined by the presence of the unique cellular structure enclosing the peptidoglycan, the outer membrane (OM) (1). The OM is an asymmetric lipid bilayer that is composed of various glycerophospholipids in the inner leaflet and nearly exclusively of the glycolipid lipopolysaccharide (LPS) in the outer leaflet. In *Escherichia coli* and other closely related enteric bacteria, there are $\sim 10^6$ LPS molecules per cell covering nearly three quarters of the total outer cell surface area, accounting for 30% of the total OM gross weight (1–3). LPS is the main OM surface-associated antigen and consequently is involved in a diverse spectrum of pathological and physiological activities associated with the human host immune response and, for historical reasons, is also referred to as endotoxin (4, 5). The LPS layer is essential to both the form and function of the OM of Gram-negative bacteria. Thus, in addition to being a key player in Gram-negative pathogenesis, LPS is also a critical determinant of the survival of the bacterium.

LPS is a complex glycolipid conceptually divided into three regions: the OM-embedded lipid A, the oligosaccharide core, and, in enteric bacteria, an O-specific hydrophilic polysaccharide chain that determines the antigenic specificity of the strain (6). Lipid A is the most conserved LPS domain among Gram-negative bacterial genera and, being the structural component responsible for the biological activities within the host, represents the endotoxic principle of LPS. The oligosaccharide core connects lipid A to the hypervariable polysaccharide chain, and itself is further divided into the inner and outer oligosaccharide core regions. Whereas the outer core is less well conserved, varying both in saccharide composition and glycosidic linkages, the majority of Gram-negative bacteria elaborate

ABSTRACT Gram-negative bacteria possess an asymmetric lipid bilayer surrounding the cell wall, the outer membrane (OM). The OM inner leaflet is primarily composed of various glycerophospholipids, whereas the outer leaflet predominantly contains the unique amphiphilic macromolecule, lipopolysaccharide (LPS or endotoxin). The majority of all Gram-negative bacteria elaborate LPS containing at least one 2-keto 3-deoxy-D-manno-octulosonate (Kdo) molecule. The minimal LPS structure required for growth of *Escherichia coli* has long been recognized as two Kdo residues attached to lipid A, inextricably linking viability to toxicity. Here we report the construction and characterization of the nonconditional *E. coli* K-12 suppressor strain KPM22 that lacks Kdo and is viable despite predominantly elaborating the endotoxically inactive LPS precursor lipid IV_A. Our results challenge the established *E. coli* Kdo₂-lipid A dogma, indicating that the previously observed and well-documented dependence of cell viability on the synthesis of Kdo stems from a lethal pleiotropy precipitated after the depletion of the carbohydrate, rather than an inherent need for the Kdo molecule itself as an indispensable structural component of the OM LPS layer. Inclusion of the inner membrane LPS transporter MsbA on a multicopy plasmid partially suppresses the lethal Δ Kdo phenotype directly in the auxotrophic parent strain, suggesting increased rates of nonglycosylated lipid A transport can, in part, compensate for Kdo depletion. The unprecedented nature of a lipid IV_A OM redefines the requisite LPS structure for viability in *E. coli*.

*To whom correspondence should be addressed.
E-mail: rww@umich.edu.

Received for review November 29, 2005
and accepted January 3, 2006

Published online January 24, 2006

10.1021/cb0500015 CCC: \$33.50

© 2006 by American Chemical Society

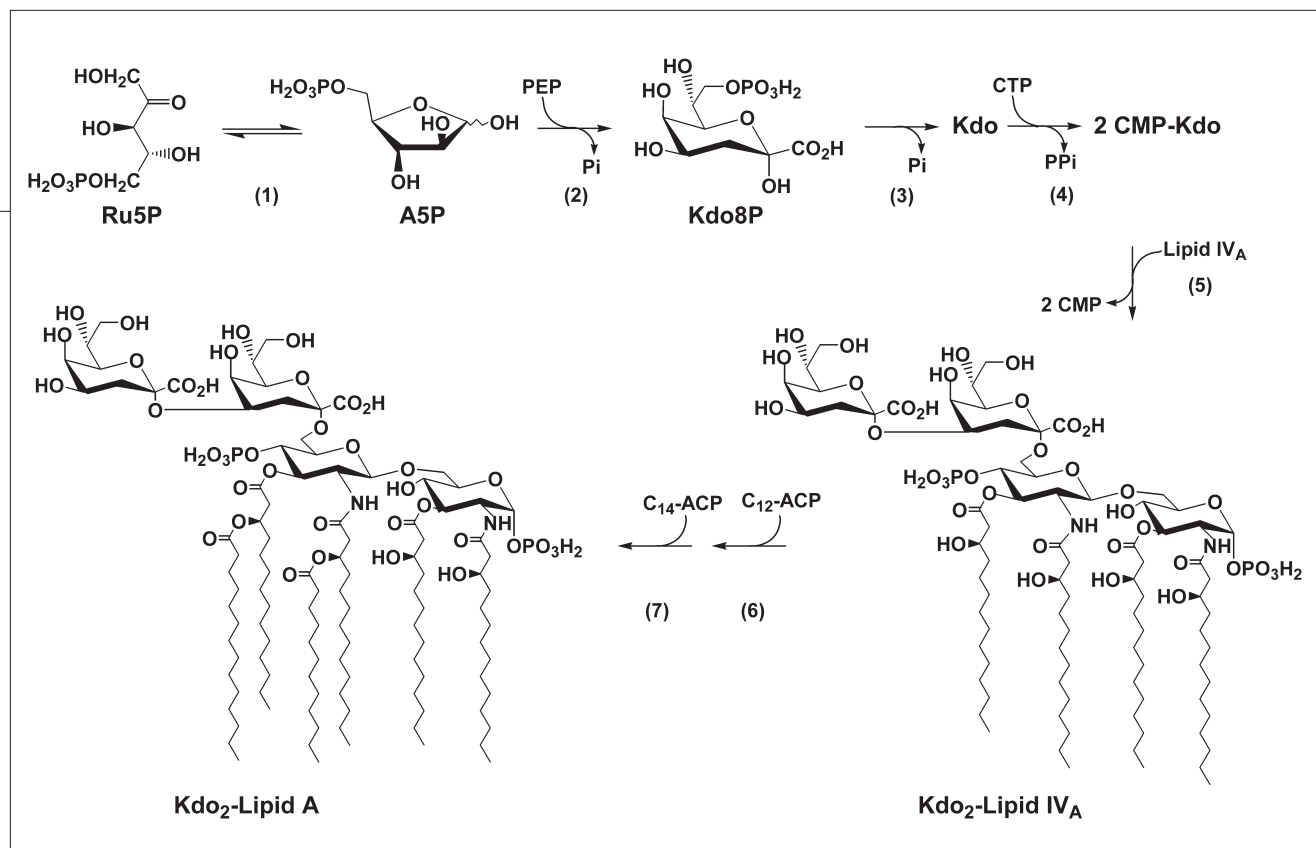


Figure 1. Biosynthesis and incorporation of Kdo into LPS. Enzymes involved are (1) D-arabinose 5-phosphate isomerase (KdsD/GutQ), (2) Kdo8P synthase (KdsA), (3) Kdo8P phosphatase (KdsC), and (4) CMP-Kdo synthetase (KdsB). In *E. coli*, two molecules of activated Kdo are then sequentially transferred to lipid IV_A by (5) Kdo transferase (WaaA) before the stepwise addition of the secondary acyl chains (6) laurate (LpxL) and (7) myristate (LpxM).

an inner core containing at least one 2-keto 3-deoxy-D-manno-octulosonate (Kdo) molecule. Kdo is an essential component of LPS that is thus far the only conserved residue found in nearly all LPS structures investigated to date (7). The minimal LPS structure required for growth of *E. coli* is two Kdo residues attached to lipid A (Kdo₂-lipid A or Re LPS) (6, 8), underscoring the fundamental role of Kdo in maintaining OM integrity and cell viability.

The ubiquitous nature of Kdo within LPS structures has prompted detailed investigation into its biosynthesis. The pathway is initiated by the enzyme D-arabinose 5-phosphate (A5P) isomerase (API), which converts the pentose pathway intermediate D-ribulose 5-phosphate into A5P. Subsequently, A5P is condensed with phosphoenolpyruvate to form Kdo 8-phosphate (Kdo8P) (KdsA), hydrolyzed to Kdo (KdsC), activated as the sugar nucleotide CMP-Kdo (KdsB), before finally being transferred from CMP-Kdo to the acceptor lipid IV_A (WaaA) (Figure 1). The late acyltransferases LpxL and LpxM next transfer the fatty acids laurate and myristate, respectively, to Kdo₂-lipid IV_A to form the characteristic acyloxyacyl units of hexaacetylated Kdo₂-

lipid A. In *E. coli* K-12, there are two API genes (*kdsD* and *gutQ*) which we have previously cloned, characterized, and shown to have nearly identical biochemical properties (9, 10). While virtually all sequenced genomes of Gram-negative bacteria encode KdsD, only a subset of *Enterobacteriaceae* also encode GutQ. The biological significance of this apparent API redundancy in *E. coli* is presently unknown, though it has been shown that either gene alone can support LPS synthesis at near wildtype levels (10). The auxotrophic strain *E. coli* K-12 TCM15 with both API genes deleted becomes dependent on exogenous A5P for growth (10), in accord with the established *E. coli* Kdo₂-lipid A (Re LPS) dogma. We now report the successful isolation of a suppressor strain derived from TCM15 that is viable despite lacking Kdo. This strain, KPM22, has been characterized and shown to be viable with an OM composed predominantly of lipid IV_A, an LPS pathway precursor that lacks any glycosylation. In addition to having potentially useful practical applications, KPM22 redefines the minimal LPS structure that is necessary to support OM biogenesis in the model Gram-negative organism *E. coli*.

RESULTS AND DISCUSSION

Our initial attempts to achieve a Kdo-negative mutant in *E. coli* by directly deleting the *kdsD* gene were unsuccessful due to full complementation by a second API gene (*gutQ*) (10). Simply deleting *gutQ* after deleting *kdsD* was also unsuccessful, as no positive double

TABLE 1. Generation times in LB media

Strain	Media Temperature		
	30 °C (min)	37 °C (min)	42 °C (min)
BW30270	39	24	22
KPM22	55	38	N/A ^a
KPM25	40	25	23

^aAfter two to three generations, growth rate was nonexponential.

API deletion clones were obtained. Only upon the inclusion of A5P in the media was a positive, though conditional, Δ API construct achieved. The auxotrophic strain *E. coli* TCM15 is conditional for A5P and is incapable of forming colonies on solid agar, regardless of incubation time or temperature. In liquid medium, however, an extended initial lag of more than 24 h is routinely followed by resumption of growth. Coincidentally, cells regain the ability to form colonies on solid agar plates, suggesting the stepwise gene deletion procedure used here facilitates the development of a compensatory suppressor mutation that can apparently subvert the A5P auxotrophic phenotype. The suppressor strain KPM22 is a nonconditional Δ API mutant capable of sustained growth in LB media (or minimal media) at 37 °C without an initial lag, though the generation time has increased by ~50% in comparison to the parent wildtype strain (Table 1; Supplementary Figure 1). However, exponential growth rates are not maintained at temperatures above 42 °C. Growth can be restored to KPM22 at 42 °C by complementation with a plasmid containing *kdsD* (i.e., KPM25), suggesting a partially defective cell envelope due to the block in Kdo synthesis.

To further investigate the nature of the defect, the saccharide composition of purified LPS extracts (Supplementary Table 2) was estimated for the inner core constituents (Kdo and heptose) and lipid A (glucosamine (GlcN)) (Figure 2, panel a). The ratios for both wildtype BW30270 (1 GlcN:0.9 Kdo:2.2 heptose) and

KPM25 (1.0 GlcN:1.0 Kdo:2.5 heptose) were consistent with the ratio for the predominant LPS species (glycoform I) of *E. coli* K-12 (1.0 GlcN:1.0 Kdo:2.0 heptose) (11). Only traces of Kdo or heptose were detected in comparison for KPM22, though GlcN was still present indicating the lipid A backbone was intact.

Silver-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of samples detected no LPS bands for KPM22 (Figure 2, panel b, top). Blotted membranes were treated with acid to cleave the saccharide core before being probed with the mAb A6 antibody, which recognizes the non-glycosylated 1,4'-bisphosphorylated β -1,6-linked GlcN disaccharide backbone of lipid A. A single band from KPM22 that migrated faster than the Kdo₂-lipid A standard but at the same level as the synthetic LPS pathway intermediate lipid IV_A was recognized by the antibody (Figure 2, panel b, middle, lanes 7 and 16, respectively). Only LPS samples prepared from KPM22 and the lipid IV_A standard were recognized by mAb A6 when the acid hydrolysis step was omitted, confirming the native structure was non-glycosylated (Figure 2, panel b, bottom). Similar results were obtained in two other constructs derived from KPM22 by deleting the first committed step (*kdsA*, KPM31) and the last step

(*waaA*, KPM40) of the Kdo pathway (Figure 2, panel b, lanes 10–15). In contrast to KPM22

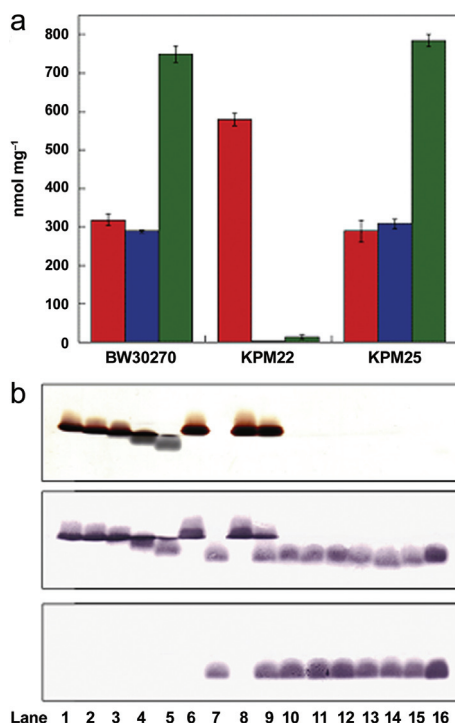
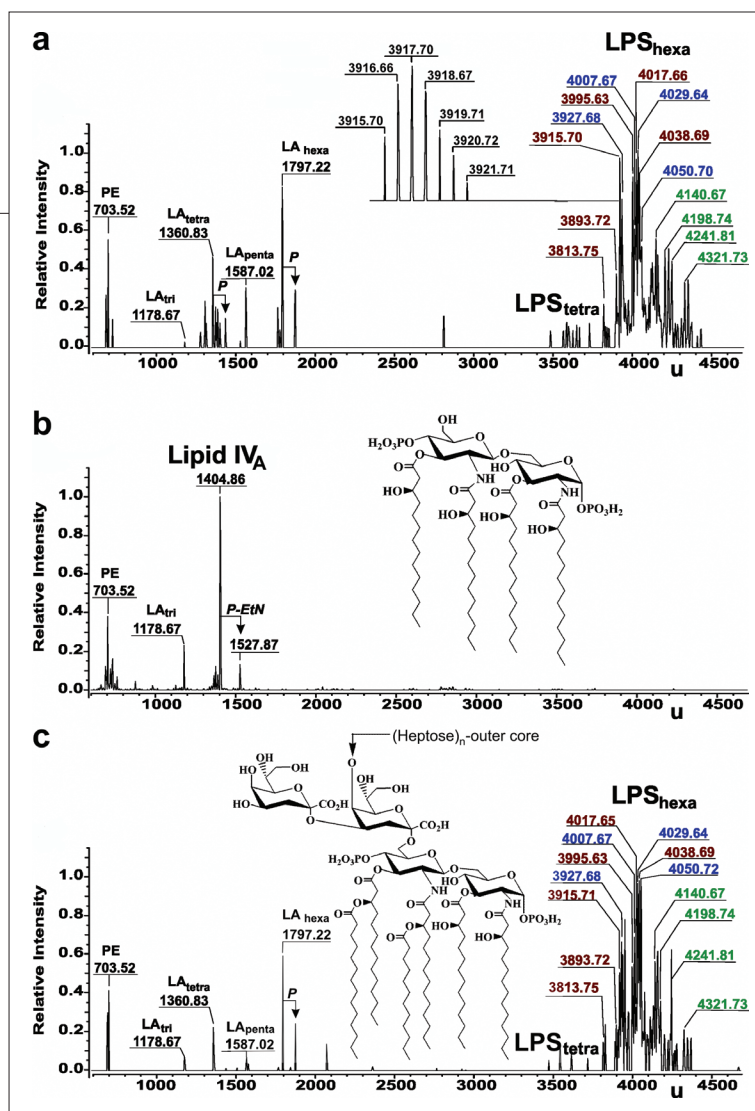


Figure 2. LPS saccharide composition and SDS-PAGE analysis. a) Inner core-lipid A saccharide composition (red bars, GlcN; blue bars, Kdo; green bars, heptose). b) Silver-stained SDS-PAGE gels (top panel) and corresponding immunoblots developed using mAb A6 (middle and bottom panels). The middle panel was treated with acetic acid to release lipid A prior to immunological reactions. Lanes 1–5 are *Salmonella enterica* serovar Typhimurium reference strains of different LPS chemotypes (Lane 1. 3749 (Ra); Lane 2. 3750 (Rb2); Lane 3. 3748 (Rb3); Lane 4. 3769 (Rd1), Lane 5. 1102 (Re)); Lane 6. wildtype BW30270; Lane 7. KPM22; Lane 8. KPM25; Lane 9. KPM22 with A5P in the growth media; Lane 10. KPM31; Lane 11. KPM34; Lane 12. KPM31 with A5P in the growth media; Lane 13. KPM40; Lane 14. KPM42; Lane 15. KPM40 with A5P in the growth media; Lane 16. 200 ng of chemically synthesized lipid IV_A (compound 406).

Figure 3. Charge deconvoluted ESI FT-ICR mass spectra in negative ion mode. a) BW30270 (*inset* isotopic distribution of glycoform I; 3915.71 u). b) KPM22 (*inset* structure of lipid IV_A; 1404.86 u). c) KPM25 (*inset* wildtype LPS with Kdo₂-lipid A (Re LPS) structure depicted and heptose attachment point indicated by arrow. Red, blue, and green peak labels correspond to peak families for glycoforms I, IV, and II, respectively (17). Mass numbers given refer to the monoisotopic masses of neutral LPS molecules and peak assignments are listed in Supplementary Table 4. Peaks corresponding to LA_{tri}, LA_{tetra}, LA_{penta}, and LA_{hexa} are presumably artifacts produced during LPS purification and/or ionization that result in cleavage of the relatively labile Kdo-lipid A linkage as they are not consistent with any known pathway intermediates. PE, phosphatidyl-ethanolamine; P, phosphate; P-EtN, phosphoethanolamine; LA_{tri}, LA_{tetra}, LA_{penta}, LA_{hexa}, acylation state of lipid A.

(Figure 2, panel b, lanes 8 and 9), neither exogenous A5P in KPM31/KPM40 (lanes 12 and 15) nor plasmid borne API in KPM34/KPM42 (lanes 11 and 14), respectively, restored LPS synthesis, consistent with the ability of KPM22 to survive without the entire Kdo pathway.

The chemotype of the LPS precursor in KPM22 was determined by electrospray ionization Fourier transform ion cyclotron (ESI FT-ICR) mass spectrometry in negative ion mode (Figure 3; Supplementary Table 3). Spectra of both wildtype BW30270 and KPM25 displayed similar peak patterns and heterogeneity within the characteristic mass range (~3900 to ~4300 u) of the various hexaacylated *E. coli* K-12 LPS glycoforms (Figure 3, panels a and c). The most prominent and nearly exclusive LPS related peak in KPM22 had a molecular mass of 1404.86 u, consistent with the structure of lipid IV_A (1,4'-bisphosphorylated tetraacylated lipid A, calculated mass 1404.854 u) (Figure 3, panel b). Lipid IV_A is an intermediate of the LPS pathway that serves as the acceptor for the sequential addition of two Kdo residues to form Kdo₂-lipid IV_A (Figure 1). The secondary acyl chain fatty acids laurate and myristate are next sequentially attached, forming hexaacylated Kdo₂-lipid A. Raetz and co-workers have shown that both acyltransferases from *E. coli* require Kdo in the lipid substrate for activity (12), explaining the lack of secondary acyl chains in KPM22.



In order to address the subcellular location of lipid IV_A and determine whether it is transported to the OM of KPM22, discontinuous sucrose gradient centrifugation was used to separate the OM from the inner membrane (IM) (Figure 4). Both membranes were well resolved, though the OM for KPM22 did not migrate as far as the wildtype OM, suggesting a decrease in buoyant density. Aside from an increase in the amount of OM porin (OMP) proteins (~35 kDa) remaining localized in the IM at the expense of accumulating in the OM, the overall total protein content and constitution as analyzed by SDS-PAGE was remarkably similar. As it has been shown that many OM proteins depend on the molecular chaperone properties of LPS for both their folding and function (13–15), the decrease in OMPs may reflect a decrease in protein transport rates and/or insertion efficiency into the OM of KPM22. Isolated OM fractions were assayed for the presence of 3-hydroxy myristate (3-OH C14:0), a characteristic LPS/lipid IV_A fatty acid marker. The OM of wildtype and KPM22 contained 11.7 and 31.1 μg of 3-OH C14:0

per mg of dried membrane, respectively, suggesting substantial quantities of lipid IV_A at least equal to the amount of LPS in wildtype are in fact present in the OM of KPM22. Further, ESI FT-ICR mass spectrometry revealed peaks for lipid IV_A in both the OM and IM of KPM22, whereas no peaks attributable to lipid IV_A were detected in either membrane fraction from wildtype (data not shown). Collectively, this indicates that while lipid IV_A is transported to the OM of KPM22, the rate of lipid IV_A transport has become uncoupled to its rate of synthesis.

Unlike hexaacylated LPS from wildtype, lipid IV_A lacks secondary acyl chains. Secondary acyl chains have been implicated in maintaining a low degree of fluidity (16), a condition that is critical to OM function. While it was previously shown that secondary acyl chains are not absolutely necessary for viability, the degree of LPS underacylation often correlates with the extent of growth defects (17–19). The tight packing of saturated acyl chains induces a network of hydrophobic interactions that help in part to maintain the integrity of the OM outer leaflet through van der Waals forces. It has been suggested that Kdo (and possibly other sugars attached distal to Kdo) further stabilize the lipid bilayer by participating in divalent cation (Mg²⁺ and/or Ca²⁺) bridges formed between negative charges contributed by both the phosphorylated lipid A backbone and the carboxylate of Kdo (16). These ionic bridges minimize electrostatic repulsion while fostering strong lateral interactions between neighboring LPS molecules. Lipid IV_A is apparently sufficient for supporting OM biogenesis despite containing only four acyl chains and no Kdo. The unprecedented nature of a lipid IV_A layer in the OM of KPM22 redefines the requisite LPS structure for viability in *E. coli*.

The cell morphology of KPM22 was examined by transmission electron microscopy (TEM). Overall, the structure of KPM22 was quite similar to the parent wildtype strain (Figure 5). There were no observable division defects, cells maintained the normal coliform rod shape, and two clearly distinct membranes can be discerned enclosing the periplasm. The contracted periplasmic volume may be an artifact induced during the cell dehydration process caused by the destabilized OM. Substantial OM instability is further suggested by the appearance of membrane vesicles at the surface of KPM22 (Figure 5, panel c). Outer membrane vesicle (OMV) formation may be caused by electrostatic charge

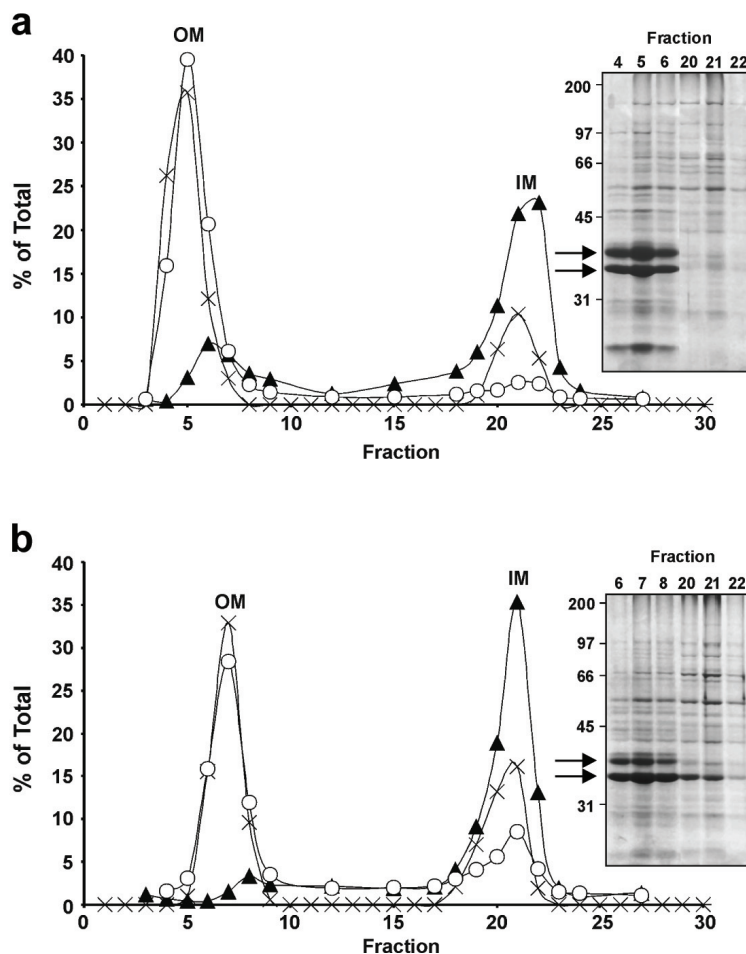


Figure 4. Sucrose gradient separation of the inner and outer membranes of wildtype BW30270 (a) and KPM22 (b). Fractions were assayed for total protein content (×), outer membrane phospholipase A (OMPLA) (O), and inner membrane NADH oxidase (Δ). SDS-PAGE gels (12%) of protein samples were run under reducing conditions. Proteins were resolved under denaturing conditions using SDS-PAGE (12% w/v). Molecular mass protein markers (kDa) are listed on the left side of each gel. Arrows indicate the position of OMP proteins (~35 kDa).

repulsion between the 1,4'-GlcN phosphates of neighboring lipid IV_A molecules that are no longer stabilized by interactions of the inner oligosaccharide core. This in turn increases the membrane curvature, possibly resulting in OMV extrusion at the bacterial surface.

The main function of the LPS layer is to act as a permeability barrier toward the diffusion of large, hydrophobic molecules and defensins (polycationic peptides) as well as to provide a measure of defense against nonspecific host cellular responses. The network of lateral interactions between adjacent LPS molecules

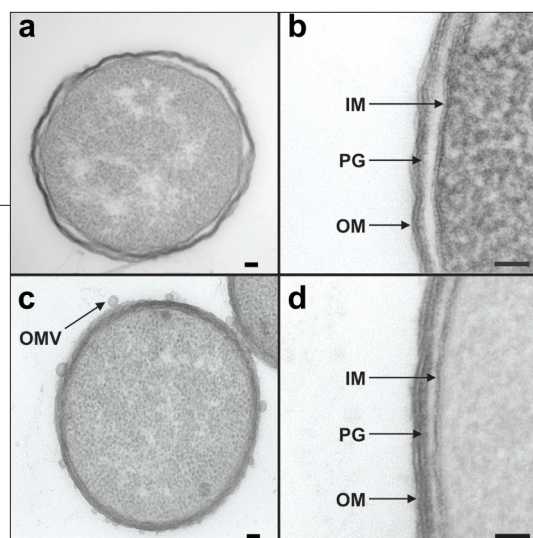


Figure 5. Transmission electron microscopy images of wildtype BW30270 (a and b) and KPM22 (c and d). Arrows indicate outer membrane vesicles at the surface of KPM22 (c). IM, inner membrane; OM, outer membrane; PG, peptidoglycan. Scale bars = 50 nm.

is particularly well suited for such a role. A panel of antibiotics and detergents were screened against KPM22 to gauge the effectiveness of an OM containing lipid IV_A as a permeability barrier (Table 2). KPM22 was supersusceptible to large, hydrophobic antibiotics that usually have reasonable efficacy against only Gram-positive bacteria, which naturally do not have an OM. Normally denied access to their cellular targets by the OM, the minimum inhibitory concentrations (MICs) were substantially reduced in KPM22. The degree of increase in potency for a particular compound corresponded well with the partition coefficient (Table 2, column XlogP). The large increase in susceptibility is likely both a direct consequence of the decreased barrier properties of lipid IV_A combined with glycerophospholipid, or perhaps even lipid IV_A, bilayer patches. In comparison, the MICs of small (<500 Da) compounds that are afforded passage primarily through water-filled OMP protein channels were only modestly decreased. KPM22 was particularly sensitive to detergents that directly target membrane integrity. Since the concentration of biological relevant detergent-like molecules such as bile salts (cholesterol metabolites) in the human intestinal tract ranges from

4 to 16 mM (~1650–6650 μg mL⁻¹) (20), *E. coli* lacking Kdo would clearly no longer be able to thrive within its native niche.

Bacterial LPS endotoxins are potent pro-inflammatory molecules that elicit an innate immune response in humans even when present in only trace amounts (21). Septic shock results from an imbalanced, dysregulated immune response. In part, this pathophysiological cascade is triggered by the LPS-induced activation of macrophages that in response secrete an array of inflammatory mediators including the pleiotropic cytokine TNF-α (tumor necrosis factor). Previous work has shown that *E. coli* mutants with underacylated lipid A have reduced endotoxicity (22), and so the immunogenic potential of samples isolated from KPM22 was explored (Figure 6). KPM22 did not elicit hTNF-α secretion from human mononuclear cells at concentrations up to 1 μg mL⁻¹, whereas restoration of Kdo synthesis with an API plasmid correlated with an increase in cytokine secretion. The disconnection in KPM22 between the dependence of cell viability on having a Kdo-containing agonistic LPS structure is a potentially useful property. Currently, the removal of LPS from biopharmaceuticals such as recombinant proteins,

TABLE 2. Permeability properties of KPM22

Compound	MW ^a (g mol ⁻¹)	XlogP ^a	MIC Wildtype (μg mL ⁻¹)	MIC KPM22 (μg mL ⁻¹)	Fold difference
Rifampin	822.9	3.72	16	0.03	512
Fusidic acid	516.7	3.7	512	2	256
Novobiocin	612.6	2.74	256	1	256
Erythromycin	733.9	1.98	128	1	128
Vancomycin	1449.3	-0.47	256	32	8
Bacitracin ^b	1422.7	-1.03	4096	512	8
Chloramphenicol	323.1	1.476	8	2	4
Ampicillin	349.4	0.25	4	2	2
Cephaloridine	416.5	1.73	4	4	1
Sodium dodecyl sulfate (SDS)			>32000	8	>4000
Bile salts ^c			16000	128	125

^aData from <http://pubchem.ncbi.nlm.nih.gov/>. ^b74000 units g⁻¹. ^cSodium cholate and deoxycholate.

DNA, and whole cell or OMV-based vaccines produced in a Gram-negative bacterial host poses a challenging problem without a universal solution (23). As the OM of KPM22 is naturally detoxified, it is envisioned to be a suitable host for the production of macromolecules in an endotoxin-free environment using the prototypical Gram-negative organism *E. coli* K-12.

Previous studies have shown that lipid A is not essential to the viability of other Gram-negative bacteria, including *Neisseria meningitidis* (24) and *Moraxella catarrhalis* (25), though OM biogenesis and cell division in *E. coli* is still considered to be absolutely dependent upon both lipid A and Kdo biosynthesis (6, 8). Targeting the API *kdsD* gene to attain Kdo-negative constructs was originally accomplished in *N. meningitidis* (26) and then later repeated in *Yersinia pestis* (27) (Various gene names have previously been used for *kdsD* including *kpsF* (26) and *yrbH* (9,27)). These mutant strains produced non-glycosylated lipid A precursors containing six acyl chains in *N. meningitidis* while the acylation state was not determined for the latter. Although both *E. coli* and *Y. pestis* are enteric bacteria, fundamental differences exist in their LPS pathways which uniquely suit their respective biology. Wildtype *Y. pestis* naturally predominately makes and transports tetraacylated LPS at 37 °C (28, 29), hinting

at a potential mechanism for evading the host immune response during infection by lowering its innate immunogenic potential. In contrast, *E. coli* and *Salmonella* do not efficiently transport tetraacylated LPS (30, 31). This may reflect the demand for a highly acylated LPS layer within the confines of the intestinal tract. In turn, evolution has restricted the substrate promiscuity of the molecular machinery responsible for LPS recognition and transport in *E. coli* to ensure under-acylated (and non-glycosylated) LPS precursors are not prematurely transported, and thus the OM permeability barrier is kept uncompromised.

Consistent with this theory, the inner membrane ABC (ATP binding cassette) transporter that flips LPS from the cytoplasm to the periplasmic face of the IM is highly selective for hexaacylated LPS/lipid A substrates *in vitro* (17, 32). MsbA was originally identified as a multicopy suppressor of LpxL (HtrB) temperature-sensitive phenotypes (33). Complementation of the auxotrophic TCM15 strain with a cosmid library of KPM22 genomic DNA revealed that MsbA was a multicopy suppressor of the Δ Kdo phenotype. Seventeen separate cosmid clones were isolated containing the *msbA* locus. A cosmid subclone (pMMW52), containing a 3.5 kb insert with only an intact wildtype *msbA* sequence identical to the wildtype, was apparently able to directly rescue TCM15 without the need to develop the presumed suppressor mutation(s), as judged by loss of A5P auxotrophy and restoration of colony-forming ability on solid agar (Table 3). The growth rate of TCM15(pMMW52) is strikingly similar to KPM22 (Tables 1 and 3). These results suggest that while lipid IV_A is a poor substrate *in vitro* (32), lipid IV_A likely becomes a substrate for MsbA *in vivo* when present in high concentrations by simple mass action. An *E. coli* construct with a mutant *kdsA* allele producing reduced levels of Kdo, resulting in a temperature-sensitive growth defect, has also been shown to be partially suppressed by increased levels of MsbA (34). Taken together, the mechanism of suppression in TCM15(pMMW52) poses two main, though not altogether unrelated, possibilities: (1) the actual minimal OM LPS structural requirement is lipid IV_A, which needs to be transported at least at a basal rate sufficient to supply enough lipid for OM biogenesis, and/or (2) lipid IV_A needs to be removed from the IM in order to relieve toxic side effects stemming from lipid IV_A induced IM perturbation.

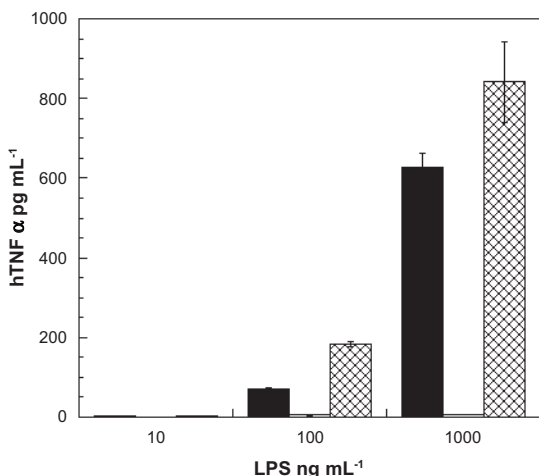


Figure 6. Endotoxic activity of LPS preparations. Human mononuclear cells (MNC) were challenged with various concentrations of LPS and hTNF- α release was measured in duplicate (shaded bars, wildtype BW30270; empty bars, KPM22; hatched bars, KPM25).

TABLE 3. Multicopy suppression of TCM15 auxotrophy by MsbA

Strain	Colony forming units (cfu) mL ⁻¹ ^a		Growth in liquid LB media ^b	
	LB only	LB + A5P/G6P ^{c,d}	LB ^d	LB + A5P/G6P ^{c,d}
TCM15	0	8.7 × 10 ⁷		+++ (23)
TCM15 (pMBL19) ^e	0	2.1 × 10 ⁶		+++ (22)
TCM15 (pMMW52) ^f	4.4 × 10 ³	3.1 × 10 ⁵	++ (33)	+++ (23)

^aCfu values correspond to either direct plating (TCM15) or post-electrotransformation as described in Methods. ^bWhere measurable, generation times (min) at 37 °C are listed in parentheses. ^c15 μM A5P, 10 μM G6P. ^dAmp (100 μg mL⁻¹) was included for strains carrying plasmid. ^eCloning vector. ^fSubclone containing *msbA*.

The derivation of KPM22 affords the unique opportunity to both challenge the existing Kdo₂-lipid A LPS dogma and to further explore the role of LPS in cell growth. Kdo is evidently not obligatory for the transport of the lipid A backbone to, and then assembly of, the OM. The discovery of MsbA as a multicopy suppressor of the lethal ΔKdo genotype presents the possibility for the *de novo* engineering of naturally detoxified ΔKdo *E. coli* strains, either through overexpression of

MsbA or perhaps eventually through the utilization of naturally more promiscuous MsbA homologues from other bacteria that can accept tetraacylated lipid IV_A as a substrate. The permeability properties of KPM22 directly and dramatically demonstrate the potential of API biosynthesis inhibition as a means to broaden the spectrum of activity of existing antibiotics, for which numerous analogues already have been made and studied, by lowering the intrinsic

resistance conferred by the OM barrier. In *E. coli* and related bacteria, API inhibition not only increases the susceptibility of the bacteria to otherwise OM-impermeable antibiotics but also decreases the risk of sepsis by lowering the endotoxin burden. The addition of these potent, but nonpenetrating, compounds to the Gram-negative antibiotic arsenal holds promise for the development of new chemotherapeutic approaches to treat Gram-negative bacterial infections.

METHODS

Strains and Media. All strains are listed in Supplementary Table 1 and were grown in standard Luria–Bertani media or MOPS-minimal media with 0.2% glycerol as the sole carbon source. KPM22 was derived by diluting (1:200 v/v) exponentially dividing cultures of TCM15 at 37 °C in MOPS-minimal media supplemented with 10 μM D-glucose 6-phosphate (G6P) and 15 μM D-arabinose 5-phosphate (A5P) into fresh media lacking sugar phosphate supplements. G6P is included to induce the *uhp* hexose phosphate transporter for which A5P is a noninducing surrogate substrate (35). After an initial lag lasting from 24 to 32 h, growth resumed and cultures were colony purified on LB agar plates. *E. coli* strain KPM22 was used as the host for chromosomal *kdsA* and *waaA* gene disruptions using the phage λ Red recombinase system according to the procedure of Datsenko and Wanner (36). Kanamycin and ampicillin were used at 15 and 100 μg mL⁻¹, respectively. Primer pairs P1/P2 with pKD13(*kan*) or P3/P4 with pKD4(*kan*) as templates were used to construct insert cassettes for KPM31 and KPM40, respectively. Antibiotic resistance markers were excised using the FLP recombinase system as described (36), except all plasmids were cured at 37 °C. Growth rates were determined using cultures in early log phase in LB media by monitoring density at 600 nm (Tables 1 and 3; Supplementary Figure 1). Colony forming unit (cfu) values were determined after electrotransformation (50 μL, ~3.8 × 10⁹ cells; 50 ng of plasmid) by plating serial dilutions.

LPS Purification. Samples were routinely prepared by growing 500 mL of each strain in LB media (37 °C, 250 rpm). Stationary phase cultures were dried and extracted according to the phenol–chloroform–petroleum ether procedure (37). Samples were further purified by sequential treatment with DNase I/RNase A, proteinase K, pelleted by ultracentrifugation, washed three times with distilled water, before being extensively

dialyzed. Representative purification yields are listed in Supplementary Table 2.

Carbohydrate Composition Analysis. The hexosamine (38), 3-deoxy sugar (39), and heptose (40) colorimetric chemical assays were used to estimate the glucosamine (GlcN), Kdo, and heptose content of LPS samples, respectively. Hexosamine content was determined posthydrolysis (500 μL of 4 M HCl, 100 °C, 18 h).

SDS-PAGE and Lipid A Immunoblots. LPS profiles of whole-cell lysates were analyzed on 13% SDS-PAGE according to the method of Hitchcock and Brown (41). Gels were silver stained for LPS analysis (41), or electrotransferred to PVDF membranes using standard procedures. Prior to incubation of the blots with mAb A6, which recognizes the nonglycosylated 1,4'-bisphosphorylated β-1,6-linked GlcN disaccharide backbone of lipid A (42), the membranes were boiled for 1 h in 1% acetic acid to cleave the α-2,6-Kdo-GlcN linkage before being developed by the usual immuno-procedure (43).

ESI FT-ICR Mass Spectrometry. Electrospray ionization Fourier transform ion cyclotron (ESI FT-ICR) mass spectrometry was performed in the negative ion mode using an APEX II Instrument (Bruker Daltonics) equipped with a 7 T actively shielded magnet and an Apollo ion source as described previously (44). Peaks were assigned on the basis of the previously published detailed structural analysis of LPS from *E. coli* K-12 strain W3100 (11). Structure assignments of the most abundant ions are summarized in Supplementary Table 3.

Sucrose Gradient Separation of IM and OM Fractions. Cells were grown in 2 L of LB medium at 37 °C to mid-log phase ($A_{600} = 0.8$), harvested by centrifugation, washed twice with buffer (50 mM Tris-HCl, pH 7.8, 1 mM EDTA), and then resuspended in 20 mL of the same buffer containing DNase I/RNase A (0.1 mg mL⁻¹ each). After cell disruption in a French pressure cell (three passes, 20000 psi), the cellular debris was

removed by centrifugation (8000g; 30 min) and the membranes were sedimented (15000g, 60 min) and then washed three times (10 mM HEPES, pH 7.4). The membrane pellet was resuspended and layered onto a discontinuous sucrose gradient and centrifuged as described (45, 46). Fractions were assayed for protein content (Bio-Rad Protein Assay Reagent), for the IM marker NADH oxidase (47) and the OM marker phospholipase A (OMPLA) (48). OM and IM fractions were pooled, sequentially diluted with buffer (10 mM HEPES, pH 7.4), and then sedimented by ultracentrifugation (150000g, 2 h) to remove sucrose before finally being resuspended in water and lyophilized.

After separation, aliquots of isolated membrane fractions were analyzed by two separate methods. The total content of the LPS/lipid IV_A marker fatty acid 3-hydroxy myristate in the membrane fractions was measured using gas chromatography-mass spectrometry (GC-MS) as described (49). To isolate lipid IV_A, samples were extracted using a single-phase Bligh-Dyer (50) mixture and then analyzed by ESI-MS.

Transmission Electron Microscopy (TEM). Cultures of cells growing in early log phase in LB media at 37 °C were fixed (2% osmium tetroxide, 90 min, room temperature, 22 °C), washed with distilled water, incubated with 2% uranyl acetate contrast solution (1 h), and then rewashed. Cells were dehydrated by a series of ethanol washes (30%, 50%, 70%, 90% ethanol, 15 min each), twice bathed in propylene oxide (15 min each), and impregnated in a propylene oxide/Epon mixture (1:1 v/v, overnight at 4 °C). Samples were polymerized (overnight at 60 °C), and the block was sliced into ultrathin sections (80–100 nm), placed on grids, and contrasted in a lead citrate solution. Images were acquired on a Phillips CM-100 transmission electron microscope equipped with an automated compustage and Kodak 1.6 Megaplug high-resolution digital camera.

Minimum Inhibitory Concentration (MIC) Determinations. The MICs of all antibiotics and drugs studied were measured in LB media using the standard serial microdilution method as described (51). Cultures were incubated with shaking (~200 rpm) at 37 °C for 18 h, and growth was scored by visual inspection. MIC values were interpreted as the lowest concentration of a drug that completely inhibited growth.

Human TNF- α Cytokine Assay. The release of tumor necrosis factor (TNF)- α cytokine from human mononuclear cells (MNCs) was measured using an enzyme-linked immunosorbent assay (52). LPS samples were resuspended in Hanks' balanced salt (HBS) solution by vortexing, aged overnight at 4 °C, and then vortexed immediately prior to use. Heparinized blood was mixed with an equal volume of HBS solution, and MNCs were isolated by differential gradient centrifugation using the Leucosep system with Lymphoprep media (Greiner Bio-One). MNCs were washed twice with RPMI 1640 media, transferred to 96-well culture plates (7.5×10^5 cells/well), and challenged with LPS. Data were collected in duplicate in three separate experiments. A representative data set is depicted in Figure 6.

Construction of KPM22 Cosmid Library. A cosmid library was constructed from KPM22 genomic DNA by partial digestion with *Sau3A*, ligation into SuperCos1, and packaged using the Gigapack III XL packaging extract as described by the manufacturer (Stratagene). TCM15 was prepared for phage infection by growth in LB media containing 0.2% (w/v) maltose and 10 mM MgSO₄ as well as additionally supplemented with A5P and G6P. Transformants were selected for growth on LB plates lacking supplemental sugar phosphates, along with the cosmid vector antibiotic resistance marker (100 μ g mL⁻¹ Amp). Cosmids were subcloned by partial *Sau3A* digestion followed by ligation into the *Bam*HI site of the medium-copy number pMBL19 cloning vector (53).

Acknowledgment: We thank Oliver Scheibe, Brigitte Kunz, Heike Kuehl, and Hermann Moll (Research Center Borstel) for

technical assistance and Professor Daniel Remick's (University of Michigan) laboratory for assistance with hTNF- α cytokine assays. The mAb A6 was provided by Dr. Lore Brade (Research Center Borstel), chemically synthesized lipid IV_A (compound 406) by Professor Koichi Fukase (Osaka University), and 2-hexadecanoyl-thioethane-1-phosphocholine by Professor Maarten R. Egmond (Utrecht University).

Supporting Information Available: A figure showing growth curves of wildtype *E. coli* K-12, KPM22, and KPM25 and tables listing bacterial strains/plasmids/primers, LPS purification summary, and ESI FT-ICR MS peak list. This material is free of charge via the Internet.

REFERENCES

1. Nikaïdo, H. (1996) In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology* (Neidhardt, F. C., Ed.), pp 29–47, American Society for Microbiology, Washington, DC.
2. Galloway, S. M., and Raetz, C. R. (1990) A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis, *J. Biol. Chem.* 265, 6394–6402.
3. Leive, L. (1974) The barrier function of the gram-negative envelope, *Ann. N.Y. Acad. Sci.* 235, 109–129.
4. Wiese, A., Brandenburg, K., Ulmer, A. J., Seydel, U., and Muller-Loennies, S. (1999) The dual role of lipopolysaccharide as effector and target molecule, *Biol. Chem.* 380, 767–784.
5. Heine, H., Rietschel, E. T., and Ulmer, A. J. (2001) The biology of endotoxin, *Mol. Biotechnol.* 19, 279–296.
6. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins, *Annu. Rev. Biochem.* 71, 635–700.
7. Holst, O. (2002) Chemical structure of the core region of lipopolysaccharides — an update, *Trends Glycosci. Glycotechnol.* 14, 87–103.
8. Gronow, S., and Brade, H. (2001) Lipopolysaccharide biosynthesis: which steps do bacteria need to survive?, *J. Endotoxin Res.* 7, 3–23.
9. Meredith, T. C., and Woodard, R. W. (2003) *Escherichia coli* YrbH is a D-arabinose 5-phosphate isomerase, *J. Biol. Chem.* 278, 32771–32777.
10. Meredith, T. C., and Woodard, R. W. (2005) Identification of GutQ from *Escherichia coli* as a D-arabinose 5-phosphate isomerase, *J. Bacteriol.* 187, 6936–6942.
11. Muller-Loennies, S., Lindner, B., and Brade, H. (2003) Structural analysis of oligosaccharides from lipopolysaccharide (LPS) of *Escherichia coli* K12 strain W3100 reveals a link between inner and outer core LPS biosynthesis, *J. Biol. Chem.* 278, 34090–34101.
12. Brozek, K. A., and Raetz, C. R. (1990) Biosynthesis of lipid A in *Escherichia coli*. Acyl carrier protein-dependent incorporation of laurate and myristate, *J. Biol. Chem.* 265, 15410–15417.
13. de Cock, H., and Tommassen, J. (1996) Lipopolysaccharides and divalent cations are involved in the formation of an assembly-competent intermediate of outer-membrane protein PhoE of *E. coli*, *Embo. J.* 15, 5567–5573.
14. Bulieris, P. V., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2003) Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide, *J. Biol. Chem.* 278, 9092–9099.
15. Sen, K., and Nikaïdo, H. (1991) Lipopolysaccharide structure required for in vitro trimerization of *Escherichia coli* OmpF porin, *J. Bacteriol.* 173, 926–928.
16. Nikaïdo, H. (2003) Molecular basis of bacterial outer membrane permeability revisited, *Microbiol. Mol. Biol. Rev.* 67, 593–656.
17. Zhou, Z., White, K. A., Polissi, A., Georgopoulos, C., and Raetz, C. R. (1998) Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis, *J. Biol. Chem.* 273, 12466–12475.
18. Vorachek-Warren, M. K., Ramirez, S., Cotter, R. J., and Raetz, C. R. (2002) A triple mutant of *Escherichia coli* lacking secondary acyl chains on lipid A, *J. Biol. Chem.* 277, 14194–14205.

19. Clementz, T., Zhou, Z., and Raetz, C. R. (1997) Function of the *Escherichia coli* msbB gene, a multicopy suppressor of htrB knock-outs, in the acylation of lipid A. Acylation by MsbB follows laurate incorporation by HtrB, *J. Biol. Chem.* **272**, 10353–10360.
20. Borgstrom, B. (1974) Bile salts—their physiological functions in the gastrointestinal tract, *Acta. Med. Scand.* **196**, 1–10.
21. Van Amersfoort, E. S., Van Berkel, T. J., and Kuiper, J. (2003) Receptors, mediators, and mechanisms involved in bacterial sepsis and septic shock, *Clin. Microbiol. Rev.* **16**, 379–414.
22. Somerville, J. E., Jr., Cassiano, L., Bainbridge, B., Cunningham, M. D., and Darveau, R. P. (1996) A novel *Escherichia coli* lipid A mutant that produces an antiinflammatory lipopolysaccharide, *J. Clin. Invest.* **97**, 359–365.
23. Petsch, D., and Anspach, F. B. (2000) Endotoxin removal from protein solutions, *J. Biotechnol.* **76**, 97–119.
24. Steeghs, L., den Hartog, R., den Boer, A., Zomer, B., Roholl, P., and van der Ley, P. (1998) Meningitis bacterium is viable without endotoxin, *Nature* **392**, 449–450.
25. Peng, D., Hong, W., Choudhury, B. P., Carlson, R. W., and Gu, X. X. (2005) *Moraxella catarrhalis* bacterium without endotoxin, a potential vaccine candidate, *Infect. Immun.* **73**, 7569–7577.
26. Tzeng, Y. L., Datta, A., Strole, C., Kolli, V. S., Birc, M. R., Taylor, W. P., Carlson, R. W., Woodard, R. W., and Stephens, D. S. (2002) KpsF Is the Arabinose-5-phosphate Isomerase Required for 3-Deoxy-D-manno-octulosonic Acid Biosynthesis and for Both Lipooligosaccharide Assembly and Capsular Polysaccharide Expression in *Neisseria meningitidis*, *J. Biol. Chem.* **277**, 24103–24113.
27. Tan, L., and Darby, C. (2005) *Yersinia pestis* is viable with endotoxin composed of only lipid A, *J. Bacteriol.* **187**, 6599–6600.
28. Kawahara, K., Tsukano, H., Watanabe, H., Lindner, B., and Mat-suura, M. (2002) Modification of the structure and activity of lipid A in *Yersinia pestis* lipopolysaccharide by growth temperature, *Infect. Immun.* **70**, 4092–4098.
29. Knirel, Y. A., Lindner, B., Vinogradov, E. V., Kocharova, N. A., Senchenkova, S. N., Shaikhutdinova, R. Z., Dentovskaya, S. V., Fursova, N. K., Bakhteeva, I. V., Titareva, G. M., Balakhonov, S. V., Holst, O., Gremyakova, T. A., Pier, G. B., and Anisimov, A. P. (2005) Temperature-dependent variations and intraspecies diversity of the structure of the lipopolysaccharide of *Yersinia pestis*, *Biochemistry* **44**, 1731–1743.
30. Nishijima, M., and Raetz, C. R. (1981) Characterization of two membrane-associated glycolipids from an *Escherichia coli* mutant deficient in phosphatidylglycerol, *J. Biol. Chem.* **256**, 10690–10696.
31. Osborn, M. J., Rick, P. D., and Rasmussen, N. S. (1980) Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Translocation and integration of an incomplete mutant lipid A into the outer membrane, *J. Biol. Chem.* **255**, 4246–4251.
32. Doerfler, W. T., and Raetz, C. R. (2002) ATPase activity of the MsbA lipid flippase of *Escherichia coli*, *J. Biol. Chem.* **277**, 36697–36705.
33. Polissi, A., and Georgopoulos, C. (1996) Mutational analysis and properties of the msbA gene of *Escherichia coli*, coding for an essential ABC family transporter, *Mol. Microbiol.* **20**, 1221–1233.
34. Fujishima, H., Nishimura, A., Wachi, M., Takagi, H., Hirasawa, T., Teraoka, H., Nishimori, K., Kawabata, T., Nishikawa, K., and Nagai, K. (2002) kdsA mutations affect FtsZ-ring formation in *Escherichia coli* K-12, *Microbiology* **148**, 103–112.
35. Eidels, L., Rick, P. D., Stimler, N. P., and Osborn, M. J. (1974) Transport of D-arabinose-5-phosphate and D-sedoheptulose-7-phosphate by the hexose phosphate transport system of *Salmonella typhimurium*, *J. Bacteriol.* **119**, 138–143.
36. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640–6645.
37. Galanos, C., Luderitz, O., and Westphal, O. (1969) A new method for the extraction of R lipopolysaccharides, *Eur. J. Biochem.* **9**, 245–249.
38. Strominger, J. L., Park, J. T., and Thompson, R. E. (1959) Composition of the cell wall of *Staphylococcus aureus*: its relation to the mechanism of action of penicillin, *J. Biol. Chem.* **234**, 3263–3268.
39. Karkhanis, Y. D., Zeltner, J. Y., Jackson, J. J., and Carlo, D. J. (1978) A new and improved microassay to determine 2-keto-3-deoxyoctonate in lipopolysaccharide of Gram-negative bacteria, *Anal. Biochem.* **85**, 595–601.
40. Osborn, M. J. (1963) Studies on the Gram-Negative Cell Wall. I. Evidence for the Role of 2-Keto-3-Deoxyoctonate in the Lipopolysaccharide of *Salmonella Typhimurium*, *Proc. Natl. Acad. Sci. U.S.A.* **50**, 499–506.
41. Hitchcock, P. J., and Brown, T. M. (1983) Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels, *J. Bacteriol.* **154**, 269–277.
42. Brade, L., Holst, O., and Brade, H. (1993) An artificial glycoconjugate containing the bisphosphorylated glucosamine disaccharide backbone of lipid A binds lipid A monoclonal antibodies, *Infect. Immun.* **61**, 4514–4517.
43. Pantophlet, R., Brade, L., and Brade, H. (1997) Detection of lipid A by monoclonal antibodies in S-form lipopoly-saccharide after acidic treatment of immobilized LPS on western blot, *J. Endotoxin Res.* **4**, 89–95.
44. Kondakova, A., and Lindner, B. (2005) Structural characterization of complex bacterial glycolipids by Fourier transform ion cyclotron mass spectrometry, *Eur. J. Mass Spectrom.* **11**, 535–546.
45. Schnaitman, C. A. (1970) Protein composition of the cell wall and cytoplasmic membrane of *Escherichia coli*, *J. Bacteriol.* **104**, 890–901.
46. Koplów, J., and Goldfine, H. (1974) Alterations in the outer membrane of the cell envelope of heptose-deficient mutants of *Escherichia coli*, *J. Bacteriol.* **117**, 527–543.
47. Osborn, M. J., Gander, J. E., Parisi, E., and Carlson, J. (1972) Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Isolation and characterization of cytoplasmic and outer membrane, *J. Biol. Chem.* **247**, 3962–3972.
48. Dekker, N., Tommassen, J., Lustig, A., Rosenbusch, J. P., and Verheij, H. M. (1997) Dimerization regulates the enzymatic activity of *Escherichia coli* outer membrane phospholipase A, *J. Biol. Chem.* **272**, 3179–3184.
49. Wollenweber, H. W., and Rietschel, E. T. (1990) Analysis of Lipopolysaccharide (Lipid-a) Fatty-Acids, *J. Microbiol. Methods* **11**, 195–211.
50. Bligh, E. G., and Dyer, W. J. (1959) A rapid method of total lipid extraction and purification, *Can. J. Biochem. Physiol.* **37**, 911–917.
51. Vuorio, R., and Vaara, M. (1992) The lipid A biosynthesis mutation lpxA2 of *Escherichia coli* results in drastic antibiotic supersusceptibility, *Antimicrob. Agents Chemother.* **36**, 826–829.
52. Copeland, S., Warren, H. S., Lowry, S. F., Calvano, S. E., and Remick, D. (2005) Acute inflammatory response to endotoxin in mice and humans, *Clin. Diagn. Lab. Immunol.* **12**, 60–67.
53. Nakano, Y., Yoshida, Y., Yamashita, Y., and Koga, T. (1995) Construction of a series of pACYC-derived plasmid vectors, *Gene* **162**, 157–158.

High Affinity InhA Inhibitors with Activity against Drug-Resistant Strains of *Mycobacterium tuberculosis*

Todd J. Sullivan[¶], James J. Truglio^{†,‡,#}, Melissa E. Boyne[§], Polina Novichenok[¶], Xujie Zhang[¶], Christopher F. Stratton[¶], Huei-jiun Li[¶], Tejinder Kaur^{¶,§}, Amol Amin[§], Francis Johnson^{¶,†}, Richard A. Slayden^{§,*}, Caroline Kisker^{†,‡,#,*}, and Peter J. Tonge^{¶,*}

[¶]Department of Chemistry, SUNY Stony Brook, Stony Brook, New York 11794-3400, [†]Department of Pharmacology and [‡]Center for Structural Biology, Stony Brook University, Stony Brook, New York 11794-5115, and [§]Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado 80523-1682. [#]Current address: Rudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, Versbacher Str. 9, 97078 Würzburg, Germany, and ⁵Albert Einstein College of Medicine, Departments of Ophthalmology and Neuroscience, Kennedy Center, Bronx, New York 10461

Tuberculosis is the leading cause of mortality from a single infectious agent and is responsible for more than 2 million deaths worldwide every year (1). Current control efforts are severely hampered by the fact that *Mycobacterium tuberculosis* (MTB) is a leading opportunistic infection in patients with AIDS and by the spread of multidrug-resistant strains of MTB (MDRTB) (2–5). Since no effective vaccine is available, the major strategy for combating the spread of this disease is through chemotherapy. The current front line treatment regimen relies on isoniazid (INH; Scheme 1), one of the most effective and widely used drugs for the treatment of tuberculosis. INH compromises the integrity of the mycobacterial cell wall by inhibiting the biosynthesis of mycolic acids (6). Although the molecular basis for the action of INH is undoubtedly complex (7–10), there is clear evidence that InhA, the enoyl reductase in the type II fatty acid biosynthesis pathway (FASII; Figure 1), is a target for INH (11–13). InhA is inhibited by an adduct formed between INH and NAD(H), following activation of INH by KatG, the mycobacterial catalase-peroxidase enzyme (Scheme 1) (11, 14–19). Since the predominant mechanism of INH resistance arises from mutations in KatG (3, 20–22), compounds that inhibit the ultimate molecular target(s) of INH, but that do not require activation by KatG, have tremendous promise as novel drugs for combating MDRTB (23, 24). Importantly, Jacobs and co-workers have validated InhA as a drug target by demonstrating that inactivation of InhA using a temperature-sensitive mutation causes the same phenotypic response as INH administration (25).

ABSTRACT Novel chemotherapeutics for treating multidrug-resistant (MDR) strains of *Mycobacterium tuberculosis* (MTB) are required to combat the spread of tuberculosis, a disease that kills more than 2 million people annually. Using structure-based drug design, we have developed a series of alkyl diphenyl ethers that are uncompetitive inhibitors of InhA, the enoyl reductase enzyme in the MTB fatty acid biosynthesis pathway. The most potent compound has a K_i value of 1 nM for InhA and MIC_{99} values of 2–3 $\mu\text{g mL}^{-1}$ (6–10 μM) for both drug-sensitive and drug-resistant strains of MTB. Overexpression of InhA in MTB results in a 9–12-fold increase in MIC_{99} , consistent with the belief that these compounds target InhA within the cell. In addition, transcriptional response studies reveal that the alkyl diphenyl ethers fail to upregulate a putative efflux pump and aromatic dioxygenase, detoxification mechanisms that are triggered by the lead compound triclosan. These diphenyl ether-based InhA inhibitors do not require activation by the mycobacterial KatG enzyme, thereby circumventing the normal mechanism of resistance to the front line drug isoniazid (INH) and thus accounting for their activity against INH-resistant strains of MTB.

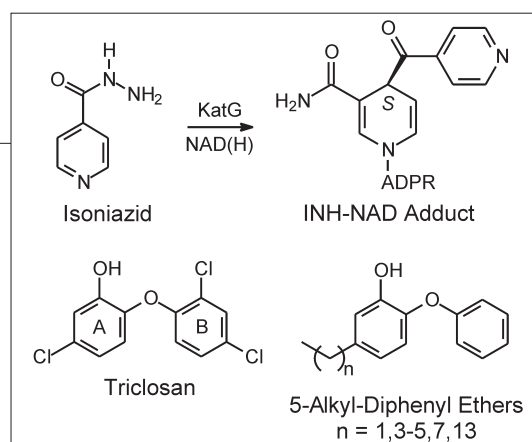
*To whom correspondence should be addressed.
E-mail: richard.slayden@colostate.edu;
caroline.kisker@virchow.uni-wuerzburg.de;
peter.tonge@sunysb.edu.

Received for review October 20, 2005
and accepted January 23, 2006.

Published online February 10, 2006

10.1021/cb0500042 CCC: \$33.50

© 2006 American Chemical Society



Scheme 1. Enoyl Reductase Inhibitors

Previous inhibitor design efforts have focused on the FASII enoyl reductase from both MTB (InhA) as well as from other organisms such as *Escherichia coli*, *Staphylococcus aureus*, and *Plasmodium falciparum* (23, 26–31). Using high throughput screening, Kuo *et al.* (23) identified two classes of InhA inhibitors based upon substituted piperazines and pyrazoles, the most potent of which had IC_{50} values for InhA of 0.16 μM ((4-(9H-fluoren-9-yl)piperazin-1-yl)(indolin-5-yl)methanone) and 2.4 μM (4-(trifluoromethyl)-2-(4,5-dihydro-4-(2,4-dinitrophenyl)pyrazol-1-yl)pyrimidine), respectively. While less potent *in vitro* than the piperazine-based compound, the pyrazole inhibited the growth of drug-resistant MTB strains with MIC_{99} values of 1–30 μM . Since several of the strains have mutations in the *katG* gene, these data are consistent with the premise that InhA inhibitors not requiring activation by KatG should be effective against INH-resistant MTB.

Our own efforts to develop a novel class of InhA inhibitors have focused on the biocide triclosan (Scheme 1), which was originally thought to have a

nonspecific mode of action but has now been shown to target the FabI enzyme in a number of organisms (see ref 26 and references therein) including InhA (8, 32). Triclosan is a picomolar inhibitor of the *E. coli* FabI enzyme (ecFabI) (33, 34), but only a sub-micromolar inhibitor of InhA (35). Using the SAR data on ligands binding to FabI and InhA, we have now designed a series of InhA inhibitors with nanomolar inhibition constants (Scheme 1; Table 1). The increase in inhibition of InhA correlates with a concomitant reduction in the MIC_{99} values for inhibiting the growth of both drug-sensitive (H37_{Rv}) and drug-resistant strains of MTB. 5-Octyl-2-phenoxyphenol (8PP), the most potent compound so far identified, has a K_i value of 1 nM for InhA and MIC_{99} values of 2–3 $\mu\text{g mL}^{-1}$ (6–10 μM) for H37_{Rv} as well as for five clinical strains of MTB with differing drug resistance profiles. Some of the clinical strains contain polymorphisms in KatG, and the antibacterial activity of the alkyl diphenyl ethers thus supports the hypothesis that compounds that inhibit InhA via a mechanism independent of KatG activation should be active against INH-resistant strains of MTB. Evidence that the compounds target InhA within the cell is provided by the observation that the MIC_{99} values for growth inhibition are increased 9–12-fold upon overexpression of InhA in MTB. Finally, transcriptional profiling suggests that the improved antibacterial activity of the alkyl diphenyl ethers compared to triclosan may be that the former compounds evade detoxification mechanisms triggered by triclosan.

Figure 1. The FASII pathway in *M. tuberculosis*. The type II fatty acid biosynthesis pathway in MTB (FASII) elongates fatty acids provided by the MTB type I pathway (FASI). The growing fatty acid is carried by the MTB-specific acyl carrier protein, AcpM, and the elongation reactions utilize malonyl-AcpM formed from holo AcpM and malonyl-CoA by FabD, the malonyl CoA:AcpM acyltransferase. The FASII cycle is initially primed by FabH, the β -ketoacyl-AcpM synthase III, which catalyzes the condensation of malonyl-AcpM with a long chain acyl-CoA provided by the FASI pathway. This reaction occurs with the loss of CO_2 , and each elongation cycle adds two carbons to the growing fatty acid. The β -ketoacyl-AcpM is reduced by the NADPH-dependent 3-ketoacyl-AcpM reductase (MabA or FabG) to 3-hydroxyacyl-AcpM, which is then dehydrated by a dehydrase to give 2-enoyl-AcpM. The last step in the cycle is catalyzed by the NADH-dependent enoyl-AcpM reductase (InhA or FabI) which reduces the enoyl-AcpM to a saturated fatty acid. Subsequent rounds of elongation are primed by a β -ketoacyl-AcpM synthase I which catalyzes the condensation of the acyl-AcpM formed by InhA with malonyl-AcpM. The resulting β -ketoacyl-AcpM is reduced, dehydrated, and reduced by the actions of MabA, the dehydrase, and InhA. There are two β -ketoacyl-AcpM synthase I enzymes in MTB that are designated, KasA and KasB. KasA is thought to function primarily in the initial rounds of elongation, while KasB is thought to be specific for very long chain fatty acids.

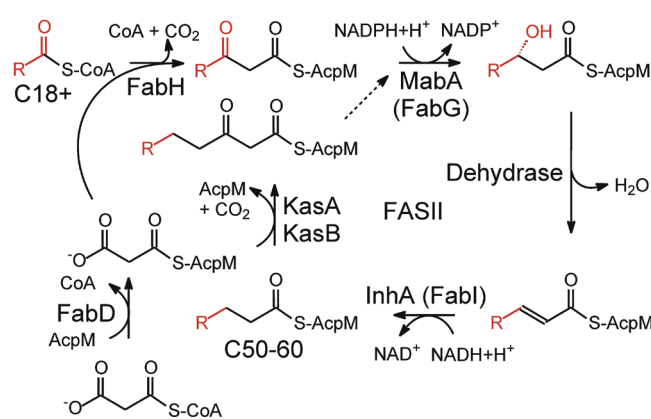


TABLE 1. Enzyme inhibition and MIC₉₉ data for triclosan, isoniazid, and the alkyl diphenyl ethers

Compound ^a	IC ₅₀ (nM) ^b	K _i ^c (nM) ^c	MIC ₉₉ , μg mL ⁻¹ (μM) ^d						
			H37 _{Rv}	H37 _{Rv} ^e pMH29:inhA	W210	TN587	NHN20	HN335	NHN382
Triclosan	1000 ± 100	220 ± 20	12.5 ± 0 (43.1 ± 0)	33.3 ± 12.9 (115 ± 45)	14.7 ± 3.8 (50.8 ± 13)	12.5 ± 0 (43.1 ± 0)	12.5 ± 0 (43.1 ± 0)	18.8 ± 6.3 (64.9 ± 21.7)	12.5 ± 0 (43.1 ± 0)
2PP	2000 ± 700		3.8 ± 0 (17.5 ± 0)						
4PP	80 ± 15								
5PP	17 ± 5	11.8 ± 4.5	1.0 ± 0 (3.5 ± 0)						
6PP	11 ± 1	9.4 ± 0.5	2.1 ± 0.9 (7.8 ± 3.3)	18.8 ± 6.8 (69 ± 25)	2.9 ± 0.4 (10.7 ± 1.5)	2.0 ± 1.0 (7.4 ± 3.7)	3.1 ± 0 (11.5 ± 0)	3.7 ± 0.9 (13.7 ± 3.3)	3.1 ± 0 (11.5 ± 0)
8PP	5.0 ± 0.3	1.1 ± 0.2	1.9 ± 0.5 (6.4 ± 1.7)	22.9 ± 5.1 (77 ± 17)	2.6 ± 0.4 (8.7 ± 1.3)	2.0 ± 1.0 (6.7 ± 3.4)	2.4 ± 0.76 (8.0 ± 2.6)	3.1 ± 0 (10.4 ± 0)	2.6 ± 0.9 (8.7 ± 3.0)
14PP	150 ± 24	30.3 ± 4.7	175 (460)						
INH		0.75 ± 0.08 ^f	0.05 ± 0 (0.37 ± 0)		0.03 ± 0.02 (0.22 ± 0.15)	2.4 ± 1.3 (17.5 ± 9.5)	0.03 ± 0 (0.22 ± 0)	0.03 ± 0 (0.22 ± 0)	1.6 ± 0 (11.7 ± 0)

^a2PP, 4PP, 5PP, 6PP, 8PP, and 14PP are the diphenyl ethers with ethyl, butyl, pentyl, hexyl, octyl, and tetradecyl substituents at the 5 position. ^bIC₅₀ determined by varying inhibitor concentration at a fixed substrate concentration. ^cK_i^c is the inhibition constant for uncompetitive inhibition of the enzyme. ^dMIC₉₉ is the concentration of inhibitor that resulted in complete inhibition in growth of MTB. Numbers reported were determined in triplicate and standard deviations are given. H37_{Rv} is a drug-sensitive strain of MTB, while W210, TN587, NHN20, HN335, and NHN382 are five clinical strains of MTB with various sensitivities to INH. ^eH37_{Rv} pMH29:inhA is the strain of MTB transformed with an InhA overexpression vector. ^fK_i for the inhibition of InhA by the INH-NAD adduct from Rawat *et al.* (24).

RESULTS AND DISCUSSION

Since INH resistance results predominantly from mutations in KatG, compounds that inhibit the molecular target(s) of INH but that do not require KatG activation should be effective against the majority of INH-resistant strains of MTB. Although the mechanism of INH action is undoubtedly complex (8), there is convincing evidence that InhA, the FASII enoyl reductase, is a target for activated INH. Using structure-based drug design, we have developed a series of alkyl diphenyl ethers that are potent *in vitro* inhibitors of InhA and that prevent the growth of both sensitive and INH-resistant MTB strains with MIC₉₉ values of 2–3 μg mL⁻¹ (6–10 μM). Overexpression of InhA in MTB results in a 9–12-fold increase in MIC₉₉, substantiating InhA as the intracellular target for these compounds. Finally, gene expression profiling suggests that the decrease in MIC₉₉ for the alkyl diphenyl ethers compared to the

parent compound triclosan may partly result from their ability to evade detoxification mechanisms triggered by triclosan.

Rational Design of Diphenyl Ether-Based InhA

Inhibitors. Levy and co-workers were the first to provide evidence that triclosan targets the FabI enoyl reductase enzyme in the bacterial FASII pathway (36). Subsequently, triclosan has been shown to inhibit the FabI enzyme from a number of organisms (see ref 26 and references therein), inhibiting the *E. coli* FabI (ecFabI) with a K₁ value of 7 pM (33, 37). SAR studies have explored the interaction of triclosan derivatives with the enzyme from several different sources (30, 34, 38–40), while Freundlich and co-workers (31) have studied in detail the importance of the triclosan 4' substituent (B ring para-substituent) on the inhibition of the *P. falciparum* FabI (pfFabI). The latter study identified several compounds that inhibited pfFabI

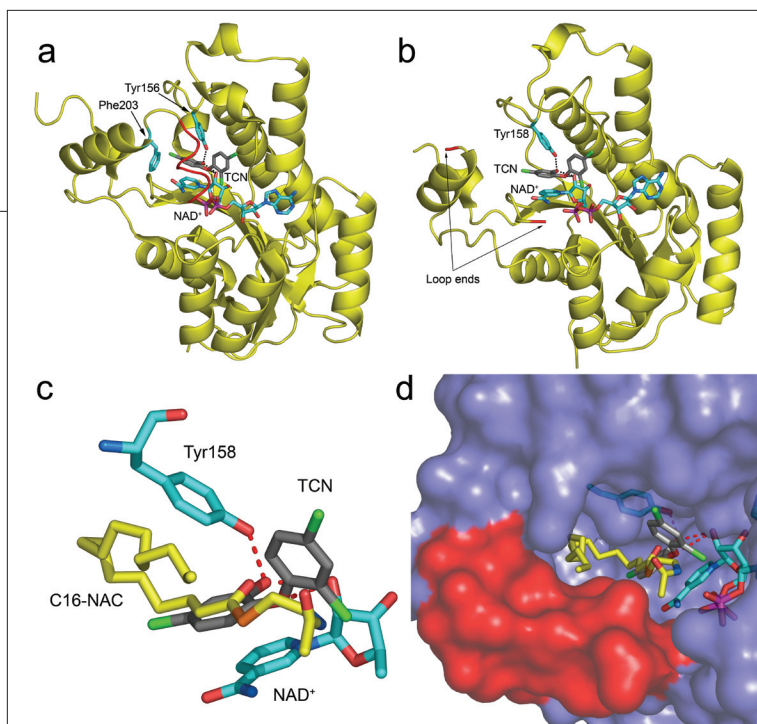


Figure 2. Structures of triclosan bound to the *E. coli* FabI and InhA. Structures of triclosan (TCN; gray molecule) in complex with NAD⁺ (blue molecule) and a) the *E. coli* FabI (1QSG.pdb) and b) InhA. In the FabI-triclosan complex, the active site loop is ordered (colored red) and residue Phe203 (blue) is shown adjacent to the loop, while in the InhA-triclosan complex, the loop is disordered (loop ends colored red). In panels c and d, the InhA-triclosan structure has been overlaid with the structure of the C16-*N*-acetylcytseamine substrate (C16-NAC; yellow) bound to InhA and NAD⁺ (1BRV.pdb; (41)). In the C16-NAC complex, the active site loop (red surface) is ordered. The figures were made with PyMol (<http://www.pymol.org>).

with IC₅₀ values similar to that of the parent compound (IC₅₀ = 73 nM) and that were active against drug-resistant strains of the parasite. Although triclosan is a picomolar inhibitor of ecFabI, this compound inhibits InhA with a K_i value of only 0.2 μM (33, 35, 37). InhA and ecFabI share a high degree of structural homology (1.6 Å rms deviation in α-carbon position) and are 36% identical in sequence (41). Consequently, our goal was to understand the 30 000-fold difference in the affinity of triclosan for ecFabI and InhA and to use this information to design high-affinity InhA inhibitors. To supplement our 1.75 Å structure of triclosan bound to ecFabI (Figure 2, panel a) (42), we determined the structure of triclosan bound to InhA. Consistent with the uncompetitive inhibition of InhA by triclosan (35), the inhibitor binds to InhA in the presence of NAD⁺ (Figure 2, panel b). The 2.3 Å resolution structure presented here reveals that the hydroxyl-substituted ring of triclosan (the “A” ring; Scheme 1) stacks with the nicotinamide ring of NAD⁺ and forms a hydrogen bond to Tyr158 and the 2′-hydroxy group of NAD⁺ (Figure 2, panel b); a highly conserved hydrogen-bonding pattern observed in complexes of triclosan with other enoyl reductase enzymes (Figure 2, panel b, black dotted lines) (23, 42, 43). The remaining inter-

actions between triclosan and InhA are predominantly hydrophobic, and the dichlorophenyl ring (the “B” ring; Scheme 1) is positioned orthogonally to the A ring with one of the chlorine atoms pointing toward the NAD⁺, while the other is solvent-exposed. The asymmetric unit contains one and a half InhA tetramers, and the mode of triclosan binding is the same in all six InhA monomers. Kuo *et al.* have also reported the crystal structure of InhA bound to triclosan (23). In the latter case, the crystals were obtained in a different space-group and only contained two InhA molecules in the symmetric unit. While triclosan was bound to the active site of each molecule in a conformation similar to that observed here, one of the InhA molecules contained a second triclosan in the active site in an inverted orientation relative to the first triclosan. This second binding mode is not present in our crystals, and we observed only one triclosan molecule per active site.

While the structures of triclosan bound to ecFabI and InhA are generally very similar, an important difference concerns the ordering of a loop of amino acids that covers the active site of each enzyme (42). This “substrate-binding” loop (residues 195–200; Figure 2, panel a, red loop) becomes ordered when triclosan binds to ecFabI (42) or when FabI is complexed with a diazaborine inhibitor (44). As triclosan is a slow, tight-binding inhibitor of ecFabI (33, 37), a plausible explanation for the slow step in triclosan binding is the ordering of this loop. In contrast, residues 197–211 in the substrate-binding loop are disordered in the InhA–triclosan complex (Figure 2, panel b, loop ends colored red), and slow onset kinetics are not observed in the inhibition of InhA by triclosan (35). Taken together, these data suggest that ordering of the FabI (InhA) active site loop is coupled to high-affinity enzyme inhibition. In support of this hypothesis, we note that the InhA loop is ordered in the structure of the INH–NAD adduct bound to InhA (19) and that the INH–NAD adduct is a slow, tight-binding inhibitor of InhA with a K_i value of 0.75 nM (24).

In previous studies on the interaction of triclosan with ecFabI, we noted that interactions between the chlorine atom on the triclosan A ring and Phe203, a residue adjacent to the substrate-binding loop in FabI (Figure 2, panel a, blue residue), were critical for high-affinity inhibition of FabI. Phe203 is mutated in *E. coli* strains with increased resistance to triclosan (36); the F203L FabI enzyme has been shown to bind triclosan

40-fold less tightly than the wild-type enzyme (37). In addition, subtle alterations in the A ring chlorine have a major impact on binding to FabI (34). In contrast to the data with ecFabI, replacement of the A ring chlorine with a fluorine or methyl group had only a minor impact on binding to InhA (unpublished data). Since InhA has a larger substrate-binding loop than ecFabI (19, 41), we speculated that larger substituents on the diphenyl ether A ring might result in additional contacts between triclosan and residues comprising the InhA substrate-binding loop. This hypothesis was substantiated following superposition of the structures of InhA in complex with triclosan (Figure 2, panel c, gray molecule) and hexadecenoyl-*N*-acetylcysteamine (C16-NAC; Figure 2, panel c, yellow molecule; PDB code 1BVR) (41). Significantly, the substrate-binding loop is ordered in the InhA–C16-NAC complex (Figure 2, panel d, red surface), and superposition of the two structures revealed that the thioester carbonyl group and the first three carbons of the acyl chain of the C16-substrate superimpose accurately with half of the triclosan A ring and the chlorine atom of ring A (Figure 2, panel c, yellow and gray molecules, respectively). Consequently, this lipophilic chlorine atom and those of ring B were removed, and the former was replaced with an alkyl chain of varying length resulting in the alkyl diphenyl ethers shown in Scheme 1.

Interaction of the Alkyl Diphenyl Ethers with InhA.

IC_{50} values for the inhibition of InhA by the alkyl diphenyl ethers are presented in Table 1. Under the conditions used, the IC_{50} value for triclosan was 1 μ M. Removal of the two chlorines on the diphenyl ether B ring and replacement of the A ring chlorine with an ethyl group resulted in a 2-fold increase in IC_{50} compared to triclosan. Thereafter, as the alkyl substituent at the meta-position on the A ring was enlarged from C2 to C8, there was a corresponding decrease in IC_{50} value for InhA inhibition from 2 μ M to 5 nM. K_i' values were determined for several inhibitors, and these values followed the trend observed in the IC_{50} values. In each case, the compounds were rapid, reversible, uncompetitive inhibitors of InhA suggesting that they bind to the InhA–NAD⁺ product complex as observed for triclosan (35). The K_i' value for triclosan was 0.22 μ M and 1.1 nM for 8PP, the most potent diphenyl ether reported in the present studies. In addition to compounds with alkyl chains of up to eight carbons, we also tested a diphenyl ether with a C14 alkyl substituent. This

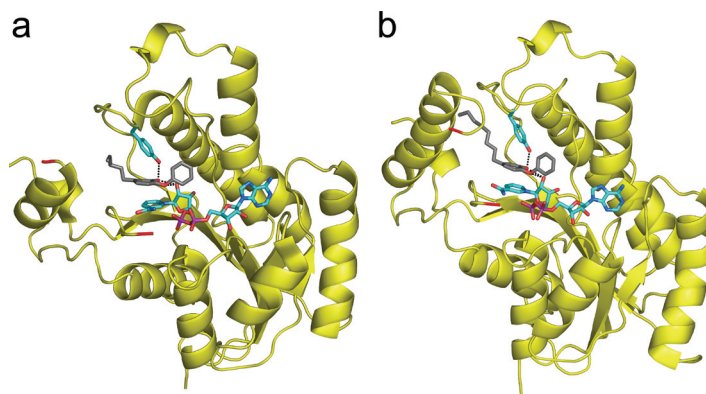


Figure 3. Structures of InhA in complex with the alkyl diphenyl ethers 5PP and 8PP. InhA in complex with NAD⁺ (blue molecule) and a) 5-pentyl-2-phenoxyphenol (5PP, gray) and b) 5-octyl-2-phenoxyphenol (8PP, gray). The color scheme is the same as that used in Figure 1, panel b. In both subunits, the active site loops are disordered and the loop ends are colored red. Also shown are hydrogen bonds (black dots) between the inhibitor (gray) and Tyr158 (blue molecule) as well as the 2'-hydroxyl group of NAD⁺ (blue molecule). The figures were made with PyMol (<http://www.pymol.org>).

compound, while still an uncompetitive inhibitor of InhA, had a K_i' value 30-fold larger than that for 8PP.

To probe the interaction of the alkyl diphenyl ethers with InhA, we solved the structures of the enzyme complexes with the pentyl (5PP) and octyl (8PP) derivatives to 2.8 and 2.6 Å resolution, respectively. As predicted from the superposition, the rings of these inhibitors bind in a similar orientation as the triclosan rings (Figure 3, gray molecules). In addition, a similar hydrogen-bonding pattern is observed between inhibitor hydroxyl substituents and Tyr158 as well as the 2'-hydroxyl group of NAD⁺ (Figure 3, black dotted lines). The 5PP alkyl chain is in a similar position in the active site as the alkyl chain of the C16 substrate. The pentyl chain forms predominantly hydrophobic interactions with residues Phe149, Met155, Tyr158, and Leu218. Superposition of the 8PP inhibitor structure with the C16 structure shows that the first four carbons of the octyl chain superimpose well with the C16 fatty-acyl chain. However, in contrast to the C16 substrate, which bends in the active site and forms a U-shape within the peptide gallery, the octyl chain adopts a linear conformation and burrows into the protein forming hydrophobic contacts with Phe149, Met155, Ala157, Pro156, Ile215, and Leu218. To accommodate the linear conformation, Leu218 is shifted by an average of 2 Å.

Similar to the InhA–triclosan complex, the majority of the substrate-binding loop is disordered in both the

pentyl and octyl ligand structures (Figure 3, loop ends colored red). Thus, in agreement with our hypothesis, compounds that are rapid, reversible inhibitors of InhA do not cause loop ordering. We note that the pentyl and octyl chains interact only with the very last residues of the loop, and further elaboration of the inhibitors is presumably required to cause loop ordering. We predict that compounds with this property will be slow, tight-binding InhA inhibitors with sub-nanomolar affinities for the enzyme.

Antibacterial Activity against Drug-Sensitive and Drug-Resistant Strains of MTB. Table 1 contains MIC₉₉ values for triclosan and several of the substituted diphenyl ethers against H37_{Rv}, a drug-sensitive laboratory strain of MTB, and five clinical strains of MTB, W210, TN587, NHN20, HN335, and NHN382 with multiple drug resistance profiles and, specifically, with various levels of susceptibility to the InhA inhibitor INH. Both TN587 and NHN382 are INH-resistant due to KatG alterations, the most common mechanism of resistance to INH. TN587 carries the KatG S315T mutation, while in NHN382, the *katG* gene has been deleted. We also included clinical strains NHN20 and HN335 which, while having a wild-type copy of *katG*, have been reported to have low levels of INH resistance.

Triclosan has an MIC₉₉ value of 12.5 μg mL⁻¹ (43 μM) for H37_{Rv}, which decreases to 3.8 and 1 μg mL⁻¹ (17.5 and 3.5 μM) for the diphenyl ethers with ethyl (2PP) and pentyl (5PP) substituents, respectively. 6PP and 8PP have MIC₉₉ values of 2 μg mL⁻¹ (7.8 and 6.4 μM, respectively) against H37_{Rv}. However, when the alkyl chain was extended to 14 carbons, a much larger MIC₉₉ value of 175 μg mL⁻¹ (460 μM) was obtained. Thus, there is a correlation between the antibacterial activity of the diphenyl ethers and the affinity of these compounds for InhA. This is significant, especially when we take into account that modulation of the triclosan skeleton has also improved the therapeutic index for these InhA inhibitors. Thus, while triclosan is toxic to Vero cells at 5 × MIC, 6PP and 8PP display no toxicity up to 10 × MIC (unpublished data).

Triclosan, 6PP, and 8PP were also evaluated against the five clinical MTB strains and were each shown to have MIC₉₉ values close to those exhibited against H37_{Rv}. In contrast, while the MIC₉₉ values for INH against three of the clinical strains (W210, NHN20, and HN335) were similar to the value obtained against H37_{Rv} (0.05 μg mL⁻¹, 0.37 μM), strains TN587 and

NHN382 had MIC₉₉ values 30–50-fold higher than for H37_{Rv}. Thus, not only are 6PP and 8PP more potent than triclosan at inhibiting the growth of MTB, but they are also active against clinical strains resistant to the front line TB drug INH. Thus, in support of our hypothesis, InhA inhibitors that do not require KatG activation are active against INH-resistant MTB. The correlation between the *K_i* and MIC₉₉ values suggests that the antibacterial activity of the compounds results from inhibition of InhA within the mycobacterium, a proposal supported by the observation that the C14-substituted diphenyl ether has a reduced affinity for InhA and a substantially poorer MIC₉₉ value for H37_{Rv} compared to 6PP and 8PP (Table 1). To provide further mechanistic insight into the antibacterial activity of the diphenyl ethers, we evaluated the effect of InhA overexpression on the MIC₉₉ values and used QRT-PCR to monitor the effect of compound treatment on the transcriptional response of several mycobacterial genes that discriminate between the antimycobacterial compounds INH, thiolactomycin (TLM), and triclosan (45, 46).

Mode of Action Studies. Previously, it was shown that overexpression of InhA in MTB using the mycobacterial vector pMH29:*inhA* resulted in a 7-fold increase in MIC₉₉ for triclosan (8). These data are in agreement with genetic selection experiments (32) and support the hypothesis that InhA is an important intracellular target for triclosan. We have now extended these studies to include the alkyl diphenyl ethers and have shown that transformation of MTB with pMH29:*inhA* causes a 9–12-fold increase in MIC₉₉ values for 6PP and 8PP (Table 1). The increase in MIC₉₉ values are of the same magnitude as the expected increase in InhA levels resulting from use of the pMH29:*inhA* vector, and thus, these data provide strong support for the belief that the antibacterial activity of the alkyl diphenyl ethers results from an inhibition of InhA within the cell.

While the gene dosage experiments suggest that InhA is an intracellular target for both triclosan and the alkyl diphenyl ethers, the latter compounds are considerably more potent than the parent pharmacophore. To provide further insight into the mode of action of the alkyl diphenyl ethers, we evaluated the effect of these compounds on the expression levels of several genes that had previously been identified by Betts *et al.* who studied the transcriptional response of MTB to treatment with the antimycobacterial drug INH

and the *in vitro* bacterial inhibitors TLM and triclosan (46). The latter experiments were undertaken to provide more information on the molecular basis for INH action which, while known to inhibit InhA, has also been proposed to target KasA, the FASII β -ketoacyl-AcpM synthase I (7). Both INH and TLM were shown to strongly induce genes in the *kas* operon, which encodes components of the FASII system including AcpM, KasA, and KasB (8, 45, 46). Since TLM is a Kas inhibitor, these data support the hypothesis that KasA is one of the targets for INH (7, 8). In contrast, triclosan has no effect on the expression of the *kas* operon, consistent with the proposal that triclosan inhibits InhA within the mycobacterium (8, 32). Instead, triclosan strongly upregulates *rv1685c-rv1686c-1687c*, an operon which encodes a transcriptional regulator and components of an ABC-type multidrug transport system, as well as *rv3160c-rv3161c*, which encodes a putative aromatic dioxygenase that converts aromatic compounds to non-aromatic *cis*-diols.

Thus, the strongest response of MTB to triclosan treatment is to upregulate detoxification mechanisms that metabolize triclosan and pump it out of the cell (46). Interestingly, neither triclosan nor INH affected the expression of the *mab* operon that encodes the FASII enzymes MabA and InhA.

Using the gene expression data, we chose discriminant genes encoding putative drug detoxification proteins (*rv3160c-rv3161c* and *rv1685c-rv1687c*), fatty acid biosynthetic proteins related to mode of action (*fabD*, *kasA*, and *inhA*), and proteins indicative of general metabolic activity (*dnaA*) and of unknown function (*rv1072*, *rv2253*, and *rv3486*), in addition to a normalization gene (*sigA*), to evaluate the mode of action of 6PP and 8PP in comparison to the parent pharmacophore triclosan. The transcriptional response of the majority of the genes analyzed revealed that 6PP and 8PP have a similar mode of action to triclosan (Table 2). Thus, like triclosan, 6PP and 8PP failed to differentially alter the expression of the *kas*

operon, in contrast to INH and TLM (8, 46). However, 6PP and 8PP also failed to induce expression of two operons, *rv3160c-rv3161c* and *rv1685c-rv1687c*, that were highly induced in response to triclosan treatment (Table 2). Thus, one difference in MIC₉₉ values for 6PP and 8PP compared to triclosan may be that the former compounds, unlike triclosan, do not upregulate proteins involved in triclosan detoxification.

While the gene dosage experiments provide strong evidence that the alkyl diphenyl ethers target InhA within the cell, transcriptional profiling suggests a mode of action for these compounds different from that of the front line TB drug INH. This is of interest since INH is a known *in vitro* InhA inhibitor, and InhA is thought to be a principal *in vivo* target for the activated form of the drug. While we firmly believe that InhA is a target for INH, the variation in transcriptional profiles between INH and the alkyl diphenyl ethers suggests that one or both of these compounds have additional targets

TABLE 2. Transcriptional changes in discriminant genes^a

ORF	Gene	Gene product ^b	Ratio (treated/untreated)		
			Triclosan	6PP	8PP
Rv0001	<i>dnaA</i>	Chromosomal replication initiator	0.5 ± 0.2	0.9 ± 0.6	0.6 ± 0.1
Rv0077c		Probable oxidoreductase	1.6 ± 0.6	1.0 ± 0.3	1.5 ± 1.3
Rv0711	<i>atsA</i>	Possible arylsulfatase	1.2 ± 0.6	1.2 ± 0.4	1.2 ± 0.4
Rv1072		Probable conserved membrane protein	1.1 ± 0.2	1.3 ± 0.8	0.4 ± 0.1
Rv1484	<i>inhA</i>	Enoyl reductase	0.8 ± 0.0	1.1 ± 0.5	1.0 ± 0.1
Rv1557	<i>mmpL6</i>	Probable conserved transmembrane transport protein	1.3 ± 0.6	0.6 ± 0.3	1.2 ± 0.4
Rv1685c		Unknown	10.6 ± 4.6	1.1 ± 0.3	1.9 ± 0.1
Rv1686c		Probable integral membrane protein ABC transporter	15.0 ± 1.0	1.0 ± 0.3	0.8 ± 0.1
Rv1687c		Probable ATP binding protein ABC transporter	146 ± 3	0.7 ± 0.4	1.7 ± 0.2
Rv2243	<i>fabD</i>	Malonyl CoA-AcpM acyltransferase	0.7 ± 0.4	1.4 ± 0.5	1.2 ± 0.5
Rv2245	<i>kasA</i>	β -Ketoacyl synthase	1.0 ± 0.2	1.4 ± 0.1	1.0 ± 0.1
Rv2253		Unknown	0.3 ± 0.2	1.3 ± 0.9	0.7 ± 0.4
Rv3160c		Possible transcriptional regulator	4.7 ± 2.7	1.2 ± 0.5	1.1 ± 0.6
Rv3161c		Possible dioxygenase	7.4 ± 6.0	1.1 ± 0.3	1.4 ± 0.8
Rv3486		Unknown	0.6 ± 0.3	0.6 ± 0.2	0.5 ± 0.3

^aTranscriptional response of MTB to 2 h treatment with triclosan, 5-hexyl-2-phenoxyphenol (6PP), or 5-octyl-2-phenoxyphenol (8PP) normalized to *sigA* at 2 × MIC₉₉. Ratios are calculated from independent biological replicates of treated compared to untreated. Standard deviations were determined from three biologically independent experiments. ^bFunction taken from the tuberculist web site (<http://genolist.pasteur.fr/Tuberculist/>).

within the cell. Given that INH has also been proposed to target the mycobacterial FASII KasA enzyme (7) and that the activated form of INH is a highly reactive species, we believe that the data underline the complexity in the mode of action of INH (8). Thus, it seems increasingly likely that INH has multiple targets within the mycobacterium. Further insight into this question will be provided by a genomewide analysis of the effect of the alkyl diphenyl ethers on gene transcription.

Finally, it is clear that modification of the triclosan structure has resulted in compounds that are not detected by the bacterial machinery that senses and deals with exposure to chemicals that abrogate mycobacterial survival. Consequently, future studies will be focused on understanding the molecular mechanisms triggered by drug administration, since this is of paramount importance in directing the discovery of novel TB chemotherapeutics.

METHODS

Bacterial Strains. H37_{Rv} is a drug-sensitive laboratory strain of MTB. W210, TN587, NHN20, HN335, and NHN382 are clinical MTB isolates with differing drug resistance profiles. TN587 and HN335 have been described before (7, 47), while W210, NHN20, and NHN382 were obtained from Drs. Musser and Kreiswirth.

Synthesis of the Alkyl Diphenyl Ethers. The general reactions for the synthesis of the alkyl diphenyl ethers are shown in Figure 4. The bridging diphenyl ether bond was formed using Buchwald's method to couple iodobenzene with 4-chloro-2-methoxyphenol (reaction A) (48). Subsequently, a Negishi

coupling reaction was used to replace the chloro group with an alkyl chain (reaction B) (49), and then the final product was generated using BBr₃ at low temperatures (reaction C) (50). Details of the synthesis may be found in Supporting Information.

Inhibition Kinetics.

Kinetic assays using *trans*-2-dodecenoyl-Coenzyme A (DD-CoA) and wild-type InhA were performed as described previously (35). Reactions were initiated by addition of substrate to solutions containing InhA, inhibitor, and NADH in 30 mM PIPES and 150 mM NaCl,

pH 6.8, buffer. IC₅₀ values were determined by varying the concentration of inhibitor in reactions containing 250 μM NADH, 1 nM InhA, 85 μM DD-CoA, 8% (v/v) glycerol, and 0.1 mg mL⁻¹ BSA. The experimental data were analyzed using eq 1, where *I* is the inhibitor concentration and *y* is percent activity, and using a slope factor (*s*) of 1.0.

$$y = 100\% / [1 + (I/IC_{50})^s] \quad (1)$$

Inhibition constants (*K*_i) were calculated by determining the *K*_{cat} and *K*_m (DDCoA) values at several fixed inhibitor concentrations using the same assay conditions to those described above. The inhibition data were analyzed using the standard equation for uncompetitive inhibition. For compounds with *K*_i values in the low nanomolar range, initial velocities were

determined at a fixed substrate concentration and the data were fit to eq 2 using a value of *K*_m (DDCoA) = 29 μM

$$v_i/v_0 = (K_m + S)/(K_m + S[1 + I/K_i]) \quad (2)$$

where *v*_i and *v*₀ are the initial velocities in the presence and absence of inhibitor. The substrate concentration was fixed at 15 μM, and the inhibitor concentration was varied from 3.5 to 1200 nM. Data fitting was performed using Grafit 4.0 (Erithacus Software Ltd.).

X-ray Crystallographic Analysis of InhA-Inhibitor Complexes.

Native crystals of InhA were grown by vapor diffusion at 22 °C, equilibrating equal volumes of protein solution (10 mg mL⁻¹) and precipitant solution containing 12.5% (w/v) PEG 4000, 6 mM DMSO, 100 mM ADA, pH 6.8, 2 mM NAD⁺, and 100 mM ammonium acetate against a reservoir solution identical to the precipitant solution minus NAD⁺. To prepare the complex with 5-octyl-2-phenoxyphenol (8PP), the InhA crystals were transferred into soaking solution (12.5% (w/v) PEG 4000, 6 mM DMSO, 100 mM ADA, pH 6.8, 25% (v/v) glycerol, 2 mM NAD⁺, and 150 mM NaCl) containing 18 mM inhibitor and incubated for 5 days before being cryocooled in liquid nitrogen. Diffraction data were collected at beam line X26C at the National Synchrotron Light Source at Brookhaven National Laboratory and indexed, integrated, and scaled using the HKL software (51). Crystals of InhA bound to 5-pentyl-2-phenoxyphenol (5PP) were obtained by vapor diffusion, co-crystallizing the protein in the presence of inhibitor. InhA (4.2 mg mL⁻¹) was equilibrated for 2.5 h in inhibitor solution (0.4 mM NAD⁺, 150 mM NaCl, 30 mM PIPES, pH 6.8, and 2.8 mM 5PP). Equal volumes of protein-inhibitor solution and precipitant solution (14% (w/v) PEG 4000, 6 mM DMSO, 100 mM ADA, pH 6.8, and 100 mM ammonium acetate) were mixed and equilibrated against a reservoir solution identical to the precipitant solution at 22 °C. The crystals were transferred to precipitant solutions containing 3 mM 5PP, 2 mM NAD⁺, 150 mM NaCl, and increasing amounts of glycerol in steps of 5% (v/v) until a final concentration of 25% glycerol was obtained. The crystals were subsequently cryocooled in liquid nitrogen, and data were collected, indexed, integrated, and scaled as described above. Crystals of InhA bound to triclosan were also obtained by co-crystallization. Triclosan (Irgasan DP 300) was a gift from Ciba Specialty Chemicals Corp. Equal volumes of inhibitor solution (8.5 mg mL⁻¹ InhA, 30 mM PIPES, pH 6.8, 150 mM NaCl, 1 mM EDTA, pH 8.0, 1.5 mM triclosan, and 1.5 mM NAD⁺) and precipitant solution (20% (w/v) PEG 4000, 6 mM DMSO, 100 mM ammonium acetate, and 100 mM ADA, pH 6.8) were mixed and equilibrated against the precipitant solution at 22 °C. These crystals were transferred to precipitant solutions containing 1.5 mM triclosan, 1.5 mM NAD⁺, 150 mM NaCl, and increasing amounts of DMSO in 5% (v/v) increments to a final DMSO concentration of 25%.

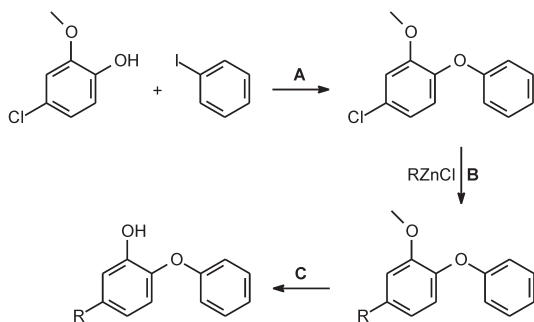


Figure 4. Synthesis of the alkyl diphenyl ethers. Conditions for each step are as follows. Reaction A: (CuOTf)₂·PhH, Cs₂CO₃, EtOAc, [ArCO₂H], toluene, 110 °C. Reaction B: 2% Pd(P(*t*-Bu)₃)₂, THF/NMP, 130 °C. Reaction C: BBr₃, dry CH₂Cl₂, -78 °C.

The crystals were cryocooled, and data were collected, indexed, integrated and scaled as described above. All three crystal forms contain six molecules in the asymmetric unit. Four of the molecules in the asymmetric unit form a homotetramer, while the remaining two subunits form half of a second homotetramer, which is completed by a crystallographic two-fold symmetry axis.

The 2.3 Å structure of InhA bound to triclosan was solved by molecular replacement with AMoRe, using the protein portion of the previously characterized structure of InhA bound to a C16-substrate (PDB code 1BVR) as a search model (41). From the MR model, phases and the associated figure of merit of the structure factors were calculated using the CCP4 programs SFALL and SIGMAA. This information was used to run the prime-and-switch procedure of RESOLVE to remove model bias. The NAD⁺ cofactor and triclosan molecules were easily identified in the resulting maps and built using the program O. Iterative model building and refinement were carried out using O and CNS, respectively.

The 2.8 Å structure of InhA bound to 5PP was solved by molecular replacement with AMoRe, using the protein portion of the triclosan structure as a search model. The MR model was used as the input for CNS rigid body refinement and simulated annealing. The NAD⁺ cofactor and 5PP could be identified in the resulting $F_o - F_c$ difference density map and were built into the electron density maps using O. Iterative model building and refinement were carried out using O and CNS, respectively.

The 2.6 Å structure of InhA bound to 8PP was solved by molecular replacement with AMoRe, using a single monomer of a previously characterized InhA (PDB code 1ENY) as a search model. The MR model was used as input for REFMAC rigid body refinement. The NAD⁺ cofactor and the 8PP were built into the resulting $F_o - F_c$ difference density map using O. Iterative model building and refinement were carried out using O and REFMAC, respectively. For all three structures, six-fold noncrystallographic symmetry (NCS) restraints were maintained throughout refinement. The tightness of restraints was chosen to minimize the free *R*-value. For the InhA-triclosan complex, all six subunits contain a single triclosan molecule. Crystallographic statistics are given in Supplementary Table 1.

MIC₉₉ Determinations. MIC₉₉ values were determined using the microplate dilution method, as previously described (52). *M. tuberculosis* strains (H37_{Rv} and clinical isolates W210, TN587, NHN20, HN335, and NHN382) were cultivated in Middlebrook 7H9 liquid medium containing 10% OADC enrichment and 0.05% Tween-80 to an optical density of ~0.4 at 600 nM. The bacterial cultures were then prepared for testing by diluting 1:100 in 7H9 Middlebrook liquid medium containing 10% OADC enrichment and 0.05% Tween-80. A total of 50 µL of each culture was added to each well of a 96-well optical plate. INH was prepared at 90 µM in sterile H₂O, while triclosan and the alkyl diphenyl ethers were prepared at 60 µM in 5% DMSO. Compound stock solutions were diluted 1:2 in Middlebrook 7H9 liquid medium containing 10% OADC enrichment and 0.05% Tween-80 and then distributed in the plate as 2-fold serial dilutions to achieve a concentration range of 30–0.47 µM (45–0.02 µM for INH) in a total final volume of 100 µL. The plates were incubated at 37 °C for 5–7 days and evaluated for the presence of bacterial growth or non-growth using an inverted plate reading method. The MIC₉₉ was determined to be the lowest concentration of compound that inhibited visible growth of the bacterial culture. Reported MIC₉₉ values represent measurements performed in triplicate. The effect of gene dosage on the MIC₉₉ values was evaluated by overexpressing InhA in MTB H37_{Rv} using the pMH29:*inhA* vector as described previously (8). The pMH29:*inhA* vector causes an 8–10-fold increase in the concentration of InhA in the mycobacterial cells

Quantitative Real Time PCR (QRT-PCR). Total bacterial RNA

from *M. tuberculosis* strain H37_{Rv} was prepared for quantitative real time PCR analysis from cells grown in 1000 mL of 7H9 Middlebrook liquid medium supplemented with 10% OADC enrichment and 0.05% Tween-80C at 37 °C. The culture was shaken at a constant speed of 200 rpm until an OD₆₀₀ of ~0.4 was reached, and then the early mid-logarithmic phase culture was divided into eight 100 mL aliquots in 250 mL Erlenmeyer flasks; four of these aliquots were subjected to one of the following drugs: 6PP, 8PP, or triclosan at 2, 2, and 35 µM, respectively. DMSO was added to the remaining 100 mL aliquots to account for the vehicle used to solubilize the compounds, and these were considered untreated controls. The 100 mL cultures were incubated at 37 °C with shaking at a constant speed of 200 rpm for 2 h. Upon completion of the incubation, the cells in the cultures were pelleted by centrifugation and resuspended in 10 mL of TRIzol reagent to stabilize the RNA. Total bacterial RNA was isolated from these cultured bacteria by sonic disruption of the TRIzol-bacterial mixture for a total of 3 min on ice. RNA was partitioned into the aqueous phase from the TRIzol bacterial homogenate mixture by the addition of 2 mL of chloroform and centrifuged at 27 000g for 30 min. RNA was further purified by precipitation in isopropyl alcohol overnight at -80 °C and centrifugation at 2800g for 1 h. Final preparation of total RNA was performed utilizing an RNeasy miniprep kit (Qiagen).

cDNA was generated from total RNA using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). Primer sequences for *sigA*, *rv2243*, *rv2245*, *rv1685c*, *rv1687c*, *rv3160c*, and *rv3161c* were identical to those described by Betts *et al.* (46) (Supplementary Table 2). Primer sequences for *rv0001*, *rv0077c*, *rv0711*, *rv1072*, *rv1484*, *rv1557*, *rv1686c*, *rv2253*, and *rv3486* were designed using Primer3 (<http://frodo.wi.mit.edu/cgi-bin/primer3/primer3-www.cgi>) (Supplementary Table 2). QPCR was performed with an iCycler iQ real-time PCR detection system (Bio-Rad Laboratories) using Platinum SYBR Green qPCR SuperMix UDG (Invitrogen). PCR was performed at 55 °C for 5 min, then 95 °C for 2 min, followed by 45 cycles of 95 °C for 15 s, 60 °C for 30 s, and 72 °C for 45 s. The final steps involved an increase of 1 °C every 10 s beginning at 30 °C for 65 cycles. The relative number of transcripts for each gene was determined from standard curves established for each primer set by linear regression of data obtained using 100, 10, and 1 ng of H37_{Rv} genomic DNA. Amplification results were visualized and analyzed using the iCycler software, the quantification of each gene was normalized to the number of *sigA* molecules, and the differential gene expression for each gene was determined in each treated sample relative to paired untreated samples.

Accession Codes. Atomic coordinates for each structure have been deposited in the Protein Data Bank with the ID codes 2B35 (InhA inhibited by triclosan), 2B36 (InhA inhibited by 5-pentyl-2-phenoxyphenol), and 2B37 (InhA inhibited by 5-octyl-2-phenoxyphenol).

Acknowledgment: This work was supported by NIH Grant AI44639 to P.J.T. and AI055298 to R.A.S. P.J.T. is an Alfred P. Sloan Fellow. We thank A. Amin and D. Chatterjee for performing the Vero cell toxicity studies.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. World Health Organization (2005) Fact sheet on tuberculosis, <http://www.who.int/mediacentre/factsheets/fs104/en/print.html>.
2. Bloom, B. R., and Murray, C. J. (1992) Tuberculosis: commentary on a reemerging killer, *Science* 257, 1055–1064.

3. Heym, B., Honore, N., Truffot-Pernot, C., Banerjee, A., Schurra, C., Jacobs, W. R., Jr., van Embden, J. D., Grosset, J. H., and Cole, S. T. (1994) Implications of multidrug resistance for the future of short-course chemotherapy of tuberculosis: a molecular study, *Lancet* **344**, 293–298.
4. Perlman, D. C., ElSadr, W. M., Heifets, L. B., Nelson, E. T., Matts, J. P., Chirgwin, K., Salomon, N., Telzak, E. E., Klein, O., Kreiswirth, B. N., Musser, J. M., and Hafner, R. (1997) Susceptibility to levofloxacin of *Mycobacterium tuberculosis* isolates from patients with HIV-related tuberculosis and characterization of a strain with levofloxacin monoresistance, *AIDS* **11**, 1473–1478.
5. Rattan, A., Kalia, A., and Ahmad, N. (1998) Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives, *Emerging Infect. Dis.* **4**, 195–209.
6. Brennan, P. J., Rooney, S. A., and Winder, F. G. (1970) The lipids of *Mycobacterium tuberculosis* BCG: fractionation, composition, turnover and the effects of isoniazid, *Ir. J. Med. Sci.* **3**, 371–390.
7. Mdluli, K., Slayden, R. A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., Crane, D. D., Musser, J. M., and Barry, C. E., III (1998) Inhibition of a *Mycobacterium tuberculosis* beta-ketoacyl ACP synthase by isoniazid, *Science* **280**, 1607–1610.
8. Slayden, R. A., Lee, R. E., and Barry, C. E., III (2000) Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*, *Mol. Microbiol.* **38**, 514–525.
9. Larsen, M. H., Vilcheze, C., Kremer, L., Besra, G. S., Parsons, L., Salfinger, M., Heifets, L., Hazbon, M. H., Alland, D., Sacchettini, J. C., and Jacobs, W. R. (2002) Overexpression of inhA, but not kasA, confers resistance to isoniazid and ethionamide in *Mycobacterium smegmatis*, *M. bovis* BCG and *M. tuberculosis*, *Mol. Microbiol.* **46**, 453–466.
10. Kremer, L., Dover, L. G., Morbidoni, H. R., Vilcheze, C., Maughan, W. N., Baulard, A., Tu, S. C., Honore, N., Deretic, V., Sacchettini, J. C., Locht, C., Jacobs, W. R., and Besra, G. S. (2003) Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria, *J. Biol. Chem.* **278**, 20547–20554.
11. Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., Collins, D., de Lisle, G., and Jacobs, W. R., Jr. (1994) inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*, *Science* **263**, 227–230.
12. Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., Jr., and Blanchard, J. S. (1995) Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*, *Biochemistry* **34**, 8235–8241.
13. Dessen, A., Quemard, A., Blanchard, J. S., Jacobs, W. R., Jr., and Sacchettini, J. C. (1995) Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*, *Science* **267**, 1638–1641.
14. Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*, *Nature* **358**, 591–593.
15. Johnsson, K., King, D. S., and Schultz, P. G. (1995) Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis, *J. Am. Chem. Soc.* **117**, 5009–5010.
16. Marcinkeviciene, J. A., Magliozzo, R. S., and Blanchard, J. S. (1995) Purification and characterization of the *Mycobacterium smegmatis* catalase-peroxidase involved in isoniazid activation, *J. Biol. Chem.* **270**, 22290–22295.
17. Basso, L. A., Zheng, R. J., and Blanchard, J. S. (1996) Kinetics of inactivation of WT and C243S mutant of *Mycobacterium tuberculosis* enoyl reductase by activated isoniazid, *J. Am. Chem. Soc.* **118**, 11301–11302.
18. Quemard, A., Dessen, A., Sugantino, M., Jacobs, W. R., Jr., Sacchettini, J. C., and Blanchard, J. S. (1996) Binding of catalase-peroxidase-activated isoniazid to wild-type and mutant *Mycobacterium tuberculosis* enoyl-ACP reductases, *J. Am. Chem. Soc.* **118**, 1561–1562.
19. Rozwarski, D. A., Grant, G. A., Barton, D. H. R., Jacobs, W. R., Jr., and Sacchettini, J. C. (1998) Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*, *Science* **279**, 98–102.
20. Stoeckle, M. Y., Guan, L., Riegler, N., Weitzman, I., Kreiswirth, B., Kornblum, J., Laraque, F., and Riley, L. W. (1993) Catalase-peroxidase gene sequences in isoniazid-sensitive and -resistant strains of *Mycobacterium tuberculosis* from New York City, *J. Infect. Dis.* **168**, 1063–1065.
21. Musser, J. M., Kapur, V., Williams, D. L., Kreiswirth, B. N., van Soolingen, D., and van Embden, J. D. (1996) Characterization of the catalase-peroxidase gene (katG) and inhA locus in isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance, *J. Infect. Dis.* **173**, 196–202.
22. Ramaswamy, S. V., Reich, R., Dou, S. J., Jasperse, L., Pan, X., Wanger, A., Quitugua, T., and Graviss, E. A. (2003) Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*, *Antimicrob. Agents Chemother.* **47**, 1241–1250.
23. Kuo, M. R., Morbidoni, H. R., Alland, D., Sneddon, S. F., Gourlie, B. B., Staveski, M. M., Leonard, M., Gregory, J. S., Janjigian, A. D., Yee, C., Kreiswirth, B., Iwamoto, H., Perozzo, R., Jacobs, W. R., Jr., Sacchettini, J. C., and Fidock, D. A. (2003) Targeting tuberculosis and malaria through inhibition of enoyl reductase: compound activity and structural data, *J. Biol. Chem.* **278**, 20851–20859.
24. Rawat, R., Whitty, A., and Tonge, P. J. (2003) The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase; adduct affinity and drug resistance, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13881–13886.
25. Vilcheze, C., Morbidoni, H. R., Weisbrod, T. R., Iwamoto, H., Kuo, M., Sacchettini, J. C., and Jacobs, W. R., Jr. (2000) Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASII end products and cell lysis of *Mycobacterium smegmatis*, *J. Bacteriol.* **182**, 4059–4067.
26. Heath, R. J., White, S. W., and Rock, C. O. (2001) Lipid biosynthesis as a target for antibacterial agents, *Prog. Lipid Res.* **40**, 467–497.
27. Suroliya, N., and Suroliya, A. (2001) Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*, *Nat. Med.* **7**, 167–173.
28. Heerding, D. A., Chan, G., DeWolf, W. E., Fosberry, A. P., Janson, C. A., Jaworski, D. D., McManus, E., Miller, W. H., Moore, T. D., Payne, D. J., Qiu, X., Rittenhouse, S. F., Slater-Radosti, C., Smith, W., Takata, D. T., Vaidya, K. S., Yuan, C. C., and Huffman, W. F. (2001) 1,4-Disubstituted imidazoles are potential antibacterial agents functioning as inhibitors of enoyl acyl carrier protein reductase (FabI), *Bioorg. Med. Chem. Lett.* **11**, 2061–2065.
29. Payne, D. J., Miller, W. H., Bery, V., Brosky, J., Burgess, W. J., Chen, E., DeWolf, W. E., Jr., Fosberry, A. P., Greenwood, R., Head, M. S., Heerding, D. A., Janson, C. A., Jaworski, D. D., Keller, P. M., Manley, P. J., Moore, T. D., Newlander, K. A., Pearson, S., Polizzi, B. J., Qiu, X., Rittenhouse, S. F., Slater-Radosti, C., Salyers, K. L., Seefeld, M. A., Smyth, M. G., Takata, D. T., Uzinskas, I. N., Vaidya, K., Wallis, N. G., Winram, S. B., Yuan, C. C., and Huffman, W. F. (2002) Discovery of a novel and potent class of FabI-directed antibacterial agents, *Antimicrob. Agents Chemother.* **46**, 3118–3124.
30. Perozzo, R., Kuo, M., Sidhu, A. S., Valiyaveetil, J. T., Bittman, R., Jacobs, W. R., Jr., Fidock, D. A., and Sacchettini, J. C. (2002) Structural elucidation of the specificity of the antibacterial agent triclosan for malarial enoyl acyl carrier protein reductase, *J. Biol. Chem.* **277**, 13106–13114.

31. Freundlich, J. S., Anderson, J. W., Sarantakis, D., Shieh, H. M., Yu, M., Valderramos, J. C., Lucumi, E., Kuo, M., Jacobs, W. R., Jr., Fidock, D. A., Schiehsler, G. A., Jacobus, D. P., and Sacchettini, J. C. (2005) Synthesis, biological activity, and X-ray crystal structural analysis of diaryl ether inhibitors of malarial enoyl acyl carrier protein reductase. Part 1: 4'-substituted triclosan derivatives, *Bioorg. Med. Chem. Lett.* **15**, 5247–5252.
32. McMurry, L. M., McDermott, P. F., and Levy, S. B. (1999) Genetic evidence that InhA of *Mycobacterium smegmatis* is a target for triclosan, *Antimicrob. Agents Chemother.* **43**, 711–713.
33. Ward, W. H. J., Holdgate, G. A., Rowsell, S., McLean, E. G., Paupit, R. A., Clayton, E., Nichols, W. W., Colls, J. G., Minshull, C. A., Jude, D. A., Mistry, A., Timms, D., Camble, R., Hales, N. J., Britton, C. J., and Taylor, I. W. F. (1999) Kinetic and structural characteristics of the inhibition of enoyl (acyl carrier protein) reductase by triclosan, *Biochemistry* **38**, 12514–12525.
34. Sivaraman, S., Sullivan, T. J., Johnson, F., Novichenok, P., Cui, G., Simmerling, C., and Tonge, P. J. (2004) Inhibition of the bacterial enoyl reductase FabI by triclosan: a structure-reactivity analysis of FabI inhibition by triclosan analogues, *J. Med. Chem.* **47**, 509–518.
35. Parikh, S. L., Xiao, G., and Tonge, P. J. (2000) Inhibition of InhA, the enoyl-reductase from *Mycobacterium tuberculosis*, by triclosan and isoniazid, *Biochemistry* **39**, 7645–7650.
36. McMurry, L. M., Oethinger, M., and Levy, S. B. (1998) Triclosan targets lipid synthesis, *Nature* **394**, 531–532.
37. Sivaraman, S., Zwahlen, J., Bell, A. F., Hedstrom, L., and Tonge, P. J. (2003) Structure-activity studies of the inhibition of FabI, the enoyl reductase from *Escherichia coli*, by triclosan: kinetic analysis of mutant FabIs, *Biochemistry* **42**, 4406–4413.
38. Heath, R. J., Yu, Y. T., Shapiro, M. A., Olson, E., and Rock, C. O. (1998) Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis, *J. Biol. Chem.* **273**, 30316–30320.
39. Heath, R. J., Li, J., Roland, G. E., and Rock, C. O. (2000) Inhibition of the *Staphylococcus aureus* NADPH-dependent enoyl-acyl carrier protein reductase by triclosan and hexachlorophene, *J. Biol. Chem.* **275**, 4654–4659.
40. Muralidharan, J., Suguna, K., Surolia, A., and Surolia, N. (2003) Exploring the interaction energies for the binding of hydroxy-diphenyl ethers to enoyl-acyl carrier protein reductases, *J. Biomol. Struct. Dyn.* **20**, 589–594.
41. Rozwarski, D. A., Vilcheze, C., Sugantino, M., Bittman, R., and Sacchettini, J. C. (1999) Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate, *J. Biol. Chem.* **274**, 15582–15589.
42. Stewart, M. J., Parikh, S., Xiao, G., Tonge, P. J., and Kisker, C. (1999) Structural basis and mechanism of enoyl reductase inhibition by triclosan, *J. Mol. Biol.* **290**, 859–865.
43. Heath, R. J., Rubin, J. R., Holland, D. R., Zhang, E., Snow, M. E., and Rock, C. O. (1999) Mechanism of triclosan inhibition of bacterial fatty acid synthesis, *J. Biol. Chem.* **274**, 11110–11114.
44. Baldock, C., Rafferty, J. B., Sedelnikova, S. E., Baker, P. J., Stuitje, A. R., Slabas, A. R., Hawkes, T. R., and Rice, D. W. (1996) A mechanism of drug action revealed by structural studies of enoyl reductase, *Science* **274**, 2107–2110.
45. Wilson, M., DeRisi, J., Kristensen, H. H., Imboden, P., Rane, S., Brown, P. O., and Schoolnik, G. K. (1999) Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12833–12838.
46. Betts, J. C., McLaren, A., Lennon, M. G., Kelly, F. M., Lukey, P. T., Blakemore, S. J., and Duncan, K. (2003) Signature gene expression profiles discriminate between isoniazid-, thioamycin-, and triclosan-treated *Mycobacterium tuberculosis*, *Antimicrob. Agents Chemother.* **47**, 2903–2913.
47. Gutacker, M. M., Smoot, J. C., Migliaccio, C. A., Ricklefs, S. M., Hua, S., Cousins, D. V., Graviss, E. A., Shashkina, E., Kreiswirth, B. N., and Musser, J. M. (2002) Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains, *Genetics* **162**, 1533–1543.
48. Marcoux, J. F., Doye, S., and Buchwald, S. L. (1997) A general copper-catalyzed synthesis of diaryl ethers, *J. Am. Chem. Soc.* **119**, 10539–10540.
49. Dai, C. Y., and Fu, G. C. (2001) The first general method for palladium-catalyzed Negishi cross-coupling of aryl and vinyl chlorides: use of commercially available Pd(P(t-Bu)₃)₂ as a catalyst, *J. Am. Chem. Soc.* **123**, 2719–2724.
50. Neumeyer, J. L., Baidur, N., Yuan, J., Booth, G., Seeman, P., and Niznik, H. B. (1990) Development of a high-affinity and stereoselective photoaffinity label for the D-1 dopamine receptor: synthesis and resolution of 7-(¹²⁵I)iodo-8-hydroxy-3-methyl-1-(4'-azidophenyl)-2,3,4,5-tetrahydro-1H-3-benzazepine, *J. Med. Chem.* **33**, 521–526.
51. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* **276**, 307–326.
52. Slayden, R. A., and Barry, C. E., III (2002) The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *Mycobacterium tuberculosis*, *Tuberculosis* **81**, 1–12.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California-Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Timothy P. Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroiyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stülz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Enhancing the Web Experience

Few would argue that the Internet and the World Wide Web have not revolutionized the way we communicate. We e-mail our colleagues for information; we post our thoughts in our blogs; the latest news is fed to our handheld devices. And, we turn first to electronic media for our scientific information, searching databases for particular genes, structures, and papers. We download and read PDFs. *ACS Chemical Biology* is working to enhance our readers' Web experience through a number of online-only features, including Ask the Expert, WIKI pages, as well as extensive links between these features and our print content.

We include links within our content to Web-only features. For example, you can read about the lives, research interests, and backgrounds of the scientists who are featured in the Ask the Expert Web forum in the Profiles (1) section of the journal. On our website, you can easily navigate from the HTML of the Profile to the Ask the Expert section and back again. In Ask the Expert, you may ask questions of the featured scientists about the techniques they use in their research or topics on which they might have additional insight. For instance, we are currently highlighting neuroscientists Ehud Isacoff and Richard Kramer and organic chemist Dirk Trauner. One user asked these experts to speculate on why nature designed ion channels as homo-multimeric proteins instead of single polypeptide chains. Ask the Expert is designed for scientists at all career stages, from graduate students wanting to learn about a particular technique to established faculty members planning on shifting their research focus.

In the current issue of the journal, you will also note a glossary of keywords within the review by Schneider and Shilatfard (2). These keywords are linked from the HTML page to our Chemical Biology WIKI page where *you* can contribute your insight to these definitions. In addition to the growing glossary list, Chemical Biology WIKI asks for input on topics, such as, what is chemical biology and where is chemical biology being done? It is a real-time forum for users to shape discussions pertinent to chemical biology.

Spend some time with our website; read research papers; and try out our new community building features. We firmly believe that these online tools are a great way to enhance learning, reach out to new colleagues, and challenge us all to think differently about chemical biology.

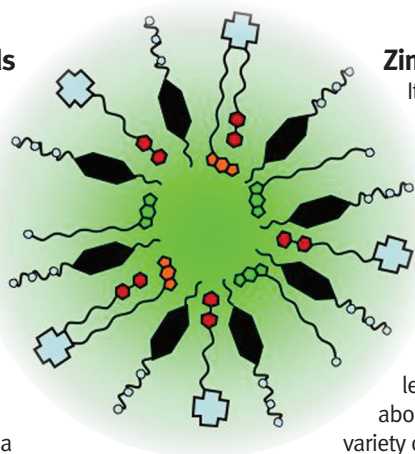


Sarah B. Tegen
Managing Editor

1. Brownlee, C. (2006) Gateways to Collaboration, *ACS Chem. Biol.* 1, 10–13.
2. Schneider, J., and Shilatfard, A. (2006) Histone Demethylation by Hydroxylation: Chemistry in Action, *ACS Chem. Biol.* 1, 75–81.

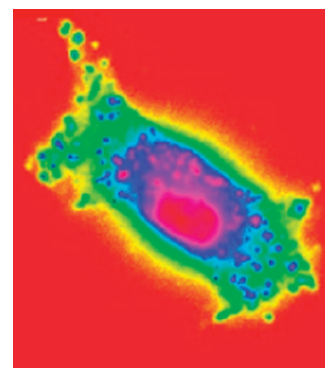
Discriminating Heads

Phospholipase A (PLA) enzymes are widely distributed in the human body and are implicated in many diseases such as cancer and autoimmune, cardiovascular, and neurological disorders. All PLA enzymes bind phospholipids and cleave them to generate a free fatty acid (such as arachidonic acid) and a lysophospholipid. Both of these products mediate a host of cellular signaling cascades. The subcellular localization and regulation of PLA isoforms in various cell types and under different physiological conditions is an area of intense study. How and when individual PLA enzymes selectively interact with their target phospholipid is not clear. To understand the substrate specificity of PLAs, Tyler and Prestwich (p 83; see the accompanying Point of View by Cho) synthesize fluorogenic analogues of the phospholipids phosphatidic acid (PA), phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylglycerol (PG) as model PLA substrates. They use these new molecules to determine, in real time, the influence of head group modifications on PLA activity *in vitro* and *in vivo*. The data indicate that these new probes will be useful in analyzing the selectivity of PLA enzymes and the potential overlap in the roles of individual phospholipases in pathological processes.



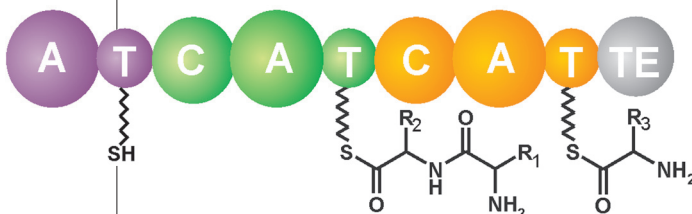
Zinc Supplements

It is estimated that 2–3 g of zinc is distributed among 1000 proteins, making zinc the second most abundant trace element in the body. This element is vital for normal function of numerous enzymes and transcription factors involved in the immune response, reproductive system, and neuronal circuits. For example, zinc is essential for proper brain function; however, excess zinc in the brain is toxic to cells. Therefore, precise control of zinc levels within cells is crucial for health. Little is known about the processes that distribute and control the levels of this metal in the body or about the incorporation of zinc into a variety of metalloproteins. Now, Bozym *et al.* (p 103; see the accompanying Point of View by Barrios) develop a sensitive fluorescence assay to measure the concentration of the free (or rapidly exchangeable) zinc in live cells. They find that in resting PC-12 and CHO cells the concentration of zinc is 5–10 pM in the cytoplasm and in the nucleus. These new zinc sensors open the door to understanding the biology of this metal ion.



Blooming Toxins

Microcystins are responsible for the poisoning of animals and humans who come in contact with cyanobacteria-contaminated water. These toxins are particularly stable as a result of their cyclic structure. Very few cyanobacterial toxins have been characterized, and methods to detect and analyze their biosynthetic pathways are under development. Previous biochemical and genetic analysis has shown that the cyclic microcystins are assembled on a large multienzyme complex via a nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS) pathway. Comparison of microcystin synthetase gene clusters with those from other NRPS/PKS systems suggested a biosynthetic mechanism for the synthesis of these toxins. However, *in vivo* data do not support all aspects of this synthetic scheme. Microcystin G polyketide synthase (McyG-PKS) contains an adenylation-peptidyl carrier protein (A-PCP) loading didomain that is believed to activate and load the starter unit phenylacetate for subsequent steps of the synthesis. Using a combination of assays and mass spectrometry techniques to observe enzyme-bound intermediates, Hicks *et al.* (p 93) report the *in vivo* and *in vitro* processing of the loading module McyG A-PCP. They find that McyG A-PCP preferentially activates phenylpropanoids rather than phenylacetates. Other results suggest that the microcystin biosynthetic pathway contains some novel biochemical reactions. Further analysis should elucidate the biosynthetic pathway of this hepatotoxin.



Published online March 17, 2006.

10.1021/cb6000922 CCC: \$33.50

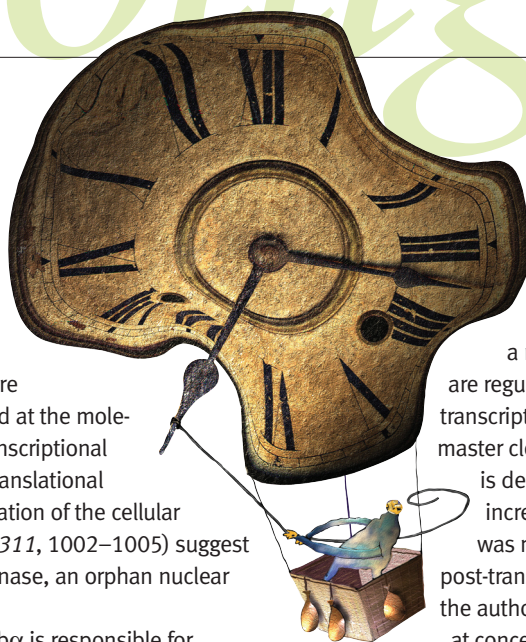
© 2006 by American Chemical Society

Spotlight

Clockwork Lithium

Lithium, a common treatment for bipolar disorder, carries a number of side effects. Sleep disturbance is among the most common. Sleep patterns are controlled by circadian rhythm, the 24-hour internal clock in organisms. These rhythms, which are also displayed by cells, are maintained at the molecular level through interconnected transcriptional feedback loops of clock genes. Post-translational modification also plays a role in regulation of the cellular clocks. Now, Yin *et al.* (*Science* 2006, 311, 1002–1005) suggest a molecular link between lithium, a kinase, an orphan nuclear receptor, and the clock genes.

The orphan nuclear receptor Rev-erb α is responsible for modulating the transcription and rhythmic expression of Bmal1, one of many clock proteins. Previous studies showed that Rev-erb α turns down its own transcription and that of Bmal1 during circadian night. Rev-erb α contains several potential phosphorylation sites for glycogen synthase kinase 3 β (GSK3 β). Interestingly, this kinase is selectively inhibited by lithium ions and a mutation found in the fly homologue of GSK3 β affects circadian timing. To test if this



Steve Campbell/Getty Images

kinase modulates the activity of Rev-erb α , the authors depleted GSK3 β from tissue culture cells using RNA interference. To their surprise, depletion of the kinase resulted in the disappearance of Rev-erb α as well. Since a number of timing events in the circadian circuit are regulated by positive and negative feedback on transcription, the mRNA levels of both Rev-erb α and master clock regulator Bmal1 were tested. When GSK3 β is depleted, the mRNA levels of both regulators increased. These data suggested that GSK3 β activity was modulating the amount of Rev-erb α through post-translational modification. To test this hypothesis, the authors inhibited GSK3 β kinase activity with lithium at concentrations that mimic patient serum levels.

They found very little Rev-erb α protein suggesting that it was being degraded. The authors show that ubiquitin-mediated proteasomal degradation of Rev-erb α is modulated by two key serine residues which are GSK3 β phosphorylation sites. These and other data indicate that the kinase activity of GSK3 β is critical for stabilizing Rev-erb α and explain how lithium inhibition of a kinase alters the circadian clock. **JU**

Trigger Points

Many age-related degenerative diseases, including Alzheimer's, and diseases connected with medical therapy, such as dialysis-related amyloidosis (DRA), are associated with the oligomerization of protein into amyloid fibrils. The constituents of these amyloids vary widely, but the fibrils share common structural properties. *In vitro* studies indicate that fibril formation begins with a nucleation event that is followed by a rapid and cooperative

association. As the amyloid product is a heterogeneous aggregate, the molecular basis for these changes are, to date, poorly understood. Now, Eakin *et al.* (*Nat. Struct. Mol. Biol.* 2006; 13, 202–208) determine the chemical basis and resultant structural changes required for oligomerization of β -2-microglobulin (β 2m) in DRA.

β 2m is a subunit necessary for the cell surface expression of class I-like complexes such as the major histocompatibility

complex (MHC). As part of MHC turnover, β 2m is released and subsequently broken down by the kidney. The concentration of β 2m is higher in patients with kidney disease, but this increased level of the protein by itself does not result in aggregation. Under physiological conditions, the transient presence of stoichiometric Cu^{2+} is required for β 2m fibrillogenesis.

The authors show that Cu^{2+} catalyzes the formation of

(continued on page 59)

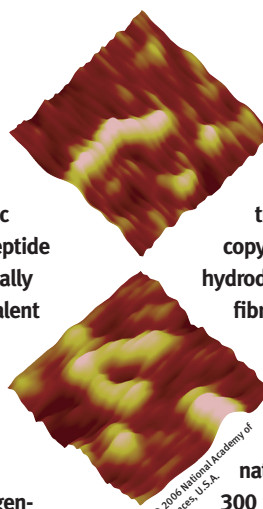
Designer Collagen

Collagen, which is Greek for “glue producer”, is the most abundant protein in humans. True to its name, collagen literally holds the human body together as the major component of connective tissue. It also contributes to skin strength and elasticity and has a greater tensile strength than steel. Collagen’s many unique properties have stimulated attempts to develop materials for medicinal and nanotechnological applications. However, natural collagen is difficult to modify and can cause pathological side effects when transplanted into humans. Furthermore, generation of synthetic collagen has proved difficult to control. Now Kotch and Raines (*Proc. Natl. Acad. Sci. U.S.A.* 2006, 103, 3028–3033) have used synthetic chemistry to create unique fragments that self-assemble into collagen.

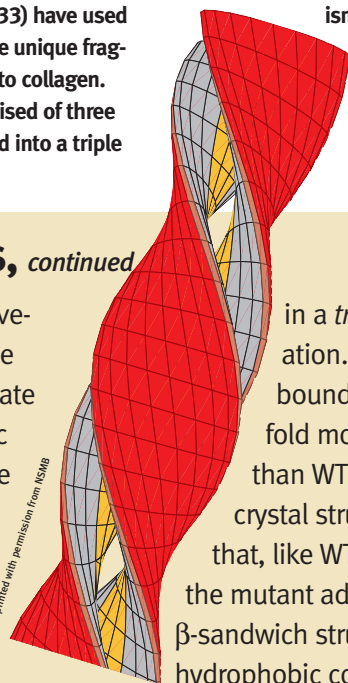
Natural collagen is comprised of three polypeptide strands that fold into a triple

helix, with each strand containing a repeating tripeptide sequence. The researchers generated synthetic peptides containing these tripeptide sequences along with strategically placed cysteines to enable covalent linkage of the strands through disulfide bonds. The peptides were designed so that the locations of the cysteines resulted in “sticky ends” that enabled self-assembly of collagen-like triple helices. The authors used a variety of techniques to characterize the synthetic collagen triple helices. Circular dichroism spectra confirmed a triple helical structure, and thermal stability experiments demonstrated cooperative denaturation,

characteristic of triple helices. Dynamic light scattering, atomic force microscopy, and transmission electron microscopy provided information about the hydrodynamic radius and length of the fibrils, which had a width near 1 nm and a length approaching 1 μm , depending on amino acid composition, temperature, and solvent. The length of natural collagen is approximately 300 nm; thus this “sticky-ends” method for producing synthetic collagen is capable of producing strands longer than those of natural collagen. This unique approach presents intriguing opportunities for creating collagen-based materials for diverse applications, ranging from tissue engineering to providing nanowires for electronic devices. **EG**



© 2006 National Academy of Sciences, U.S.A.



Reprinted with permission from KSMG

Trigger Points, *continued*

M^* , an alternative native-like conformation of the protein. This intermediate assembles into dimeric building blocks that are stabilized by Cu^{2+} . The authors hypothesized that divalent cation-mediated backbone isomerization of one proline residue was responsible for M^* formation. To test this hypothesis, Eakin *et al.* mutated the proline to alanine trapping this residue

in a *trans* conformation. This mutant bound Cu^{2+} 10 000 fold more tightly than WT $\beta 2\text{m}$. The crystal structure showed that, like WT $\beta 2\text{m}$, the mutant adopts a β -sandwich structure but its hydrophobic core is repacked. These and other structural rearrangements result in the formation of an amphipathic dimer. The dimeric structure reveals a basis for its role as

an intermediate building block for the formation of higher oligomeric states.

Metal ions are implicated in other amyloid diseases. The proteins in these systems, like $\beta 2\text{m}$, show cation-triggered conformational transitions possibly involving backbone rearrangements. Understanding the relationships between these metal-triggered transitions and fibril formation will be valuable in understanding amyloid diseases. **EJ**

Assisted Suicide

Some anticancer agents work by inducing cancer cells to undergo apoptosis or commit suicide. Cancer cells unfortunately are notorious for acquiring resistance to apoptosis-inducing drugs, including the DNA topoisomerase II-based DNA-damaging agent doxorubicin. Smukste *et al.* (*Cancer Cell* 2006, 9, 133–146) have developed an assay to identify small molecules that boost the toxicity of doxorubicin in resistant cancer cells.

Human papillomavirus (HPV), the major causative factor in cervical cancer, produces an oncogenic protein called E6. This protein increases the resistance of cancerous cells to anticancer drugs through various mechanisms. E6 forms a complex with E6 Associated Protein, an E3 ligase, and induces p53 ubiquitination and degradation. This allows tumor cells to become resistant to apoptosis. E6 also causes the degradation of a proapoptotic protein BAK and activates telomerase, which extends the life of cancer cells. To identify small molecules that overcome

E6-mediated drug resistance, the researchers used an E6-expressing colon carcinoma cell line called RKO-E6 that is two to four times more resistant to doxorubicin-mediated apoptosis than RKO cells alone. Approximately 27 000 small molecules were screened for their ability to increase the toxicity of doxorubicin, and 88 compounds were identified that selectively increase doxorubicin's lethality by at least 20%. The compounds were grouped into five structural classes, including quaternary ammonium compounds, protein synthesis inhibitors, 11-deoxyprostaglandin E1 analogues, 1,3-bis(4-morpholinylmethyl)-2-imidazolidinethione analogues, and a collection that the authors named indoxins.

Examination of the mechanisms by which the compounds exerted their effects revealed that doxorubicin resistance is tied to regulation of its

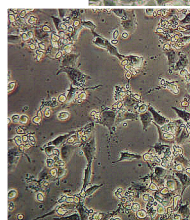
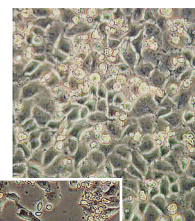


Image courtesy of B. Stowell

target, topoisomerase II α . Many of the compounds were found either to increase cellular topoisomerase II α levels or to induce S phase arrest. The

latter could affect topoisomerase II α transcription. The indoxins were especially interesting because they upregulated topoisomerase II α and caused S phase arrest, increasing the sensitivity of resistant cells to these compounds. Photoaffinity pull-down experiments revealed that the molecular target of the indoxins is a nuclear actin-related protein complex involving myosin 1c and ARP2, proteins that are known to be involved in transcriptional regulation. The mechanisms of action of the compounds identified in this screen shed light on pathways of doxorubicin resistance in cancer, and this knowledge could lead to improved treatment strategies for drug-resistant cancers. **EG**

Reforming Bubbles

How do cells know where to end transcription of a gene, and what do they do when the transcription process stalls?

Transcription termination is the answer, and *E. coli* have evolved a number of different mechanisms to accomplish this task. Termination can be accomplished by an intrinsic mechanism relying on the formation of a hairpin structure within the RNA molecule. Alternatively termination can rely on an enzymatic process involving either the Rho terminator factor or the Mfd protein, which is important for removing stalled polymerases. Parks and Roberts (*Proc. Natl. Acad. Sci U.S.A.*; doi:10.1073/pnas.0600145103) have explored the similarities between these termination mechanisms and have found

there are common elements in each of them, all involving the transcription bubble.

The transcription bubble is the region of unwound DNA in which RNA polymerase elongates the new RNA molecule. Termination and release of the RNA involve the rewinding of the DNA duplex within this region and the unwinding of the DNA/RNA hybrid. Intrinsic RNA termination requires the formation of a hairpin structure that disrupts the RNA/DNA hybrid. This disruption allows the DNA transcription bubble to be partially reformed, favoring dissociation of the RNA/DNA hybrid. Parks and Roberts examined whether Rho-mediated or Mfd-mediated

termination also involved the rewinding of the transcription bubble. They show that transcription bubbles prevented from reannealing by incorporation of mismatches within the bubble region are less efficient at releasing transcripts in both Rho-dependent and Mfd-dependent termination. They further show that both Rho and Mfd translocate the transcription bubble downstream, effectively moving the transcription bubble and further favoring dissociation of the RNA from the DNA template. These results indicate that all known mechanisms of termination involves the remodeling of the transcription bubble, and yield a clearer picture of the mechanism of prokaryotic transcription termination. **ST**

Capturing the Capsid

Many viruses, including hepatitis B virus (HBV), West Nile virus, and HIV, have macromolecular protein cores, or capsids, that contain the viral genome. Assembly of the core is an integral part of the viral life cycle, has no counterpart in the infected cell, and may be particularly sensitive to small changes in local conformation, making it an attractive target for antiviral therapy. However, capsid assembly has not been extensively pursued as a drug target, in part due to difficulty in adapting relevant assays for high throughput screening. Stray *et al.* (*Nat. Biotechnol.* 2006, 24, 358–362) has developed an *in vitro* assay for HBV capsid assembly, facilitating the search for small molecule inhibitors. This assay can be readily adapted to other virus assembly systems (such as avian influenza).

In HBV, the icosahedral core consists of 240 copies of capsid protein that enclose the viral DNA and reverse transcriptase. *In vitro*, assembly of the capsid protein into a core-like particle is dependent on many factors, such as protein concentration, pH, and ionic conditions, and is nucleated by a trimer of protein dimers. This assembly process results in close proximity of the C termini of the capsid proteins, suggesting that fluorescence quenching could be employed to monitor assembly progression. The researchers gener-

ated a fluorescent capsid protein, termed C150BO. This modified protein was highly fluorescent prior to assembly, but fluorescence dropped markedly upon capsid assembly due to fluorescence quenching. Two small molecules known to disrupt capsid assembly, urea and a heteroaryldihydropyrimidine (HAP-1), were used to determine whether perturbation of capsid assembly could be detected using C150BO. When HAP-1, which is known to accelerate capsid assembly and nucleate the formation of irregular particles, was added to the C150BO assembly reactions, a decrease in fluorescence was observed. This result suggests that HAP-1 increased the rate and extent of C150BO assembly as it does with wild-type capsid protein. Unlike HAP-1, urea inhibits capsid assembly. When it was added to C150BO assembly reactions, an increase in fluorescence was observed, confirming the decrease in assembly. These results indicate the C150BO assay can identify small molecule inhibitors of virus assembly. Furthermore, this assay is amenable to a high throughput format, providing a novel method for discovering small molecule antiviral agents that misdirect capsid formation to yield aberrant and noninfectious virus particles. EG

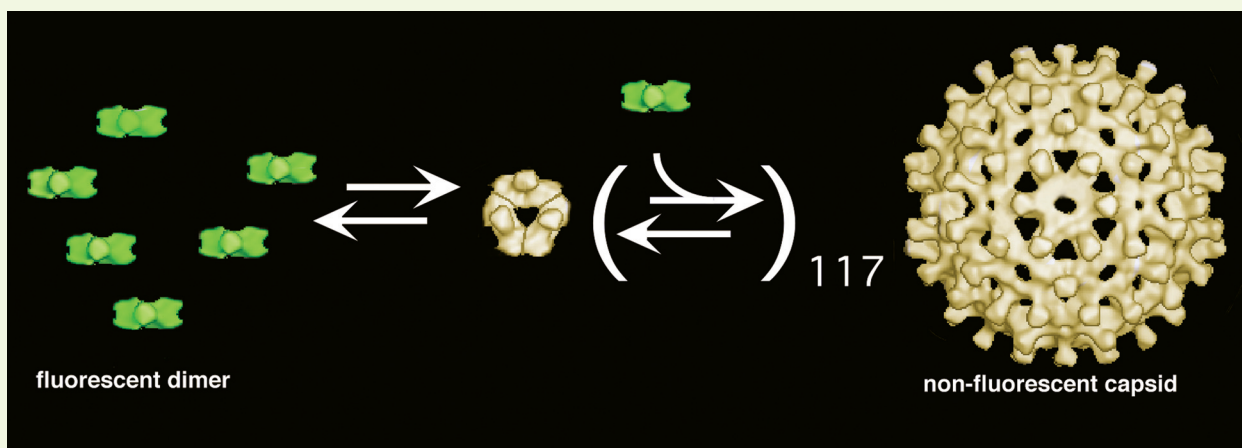


Image courtesy of A. Zlotnick

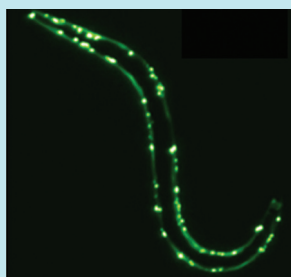
Misfortunes of Misfolding

Several neurodegenerative disorders, such as Huntington's, Alzheimer's, Parkinson's, and amyotrophic lateral sclerosis (Lou Gehrig's disease), are associated with the misfolding of proteins. In some of these disorders, protein misfolding and aggregation are caused by expansion of polyglutamine (polyQ) residues in specific proteins, but the link between aggregation and disease pathogenesis is not clearly defined. Gidalevitz *et al.* (*Science*, 2006, 311, 1471–1474.) have uncovered a connection between polyQ expansion and cellular malfunction that helps explain the role of aberrant protein folding in neurodegenerative conditions.

The researchers employed genetic methods using *Caenorhabditis elegans* (*C. elegans*) temperature-sensitive (ts) mutations

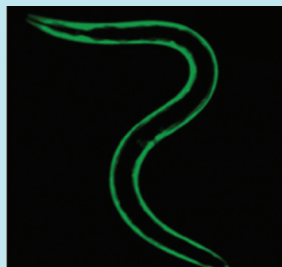
to explore the effects of the presence of an aggregation-prone protein on cell homeostasis.

Temperature-sensitive mutant proteins are dependent on the folding environment, so the extent of aggregation versus folded protein in the ts mutant worms serves as an indicator of the state of the protein-folding quality-control system. *C. elegans* strains containing either nonaggregating or aggregation-prone polyQ proteins were crossed



with several structurally and functionally unrelated ts mutants. The authors observed that the presence of an aggregation-prone polyQ protein, but not the nonaggregating protein,

was sufficient for the worms to exhibit the mutant phenotype at the permissive condition. These observations indicate



that expression of the aggregation-prone polyQ protein generally interferes with the protein-folding capacity of the cell.

The results are surprising considering that normally the presence of misfolded protein activates a stress response that increases protein refolding. The researchers hypothesize that the chronic presence of aggregation-prone proteins, while insufficient to activate the stress response, initiates a positive feedback mechanism that enhances disruption of protein folding. Over time, the gradual accumulation of misfolded and damaged proteins compromises cellular function and results in a disease state. This work puts forward a general mechanism for the catastrophic effects of protein misfolding, and further discernment of this process could lead to new therapeutic strategies to treat neurodegenerative disorders. **EG**

Integrating Your RNA

Is data coming at you faster than you can catalogue it? You're not alone but don't despair. A group of RNA scientists, the RNA Ontology Consortium (ROC) is developing an RNA Ontology (RO) (RNA 2006, published online February 16, 2006, 10.1261/rna.2343206). According to Wikipedia, ontology has one basic question: "What actually exists?" For the RNA scientists what exist are the kinds, structures, and sequences of RNA, and the events and processes in which they are involved. RNA Ontology is an initiative to create a taxonomy of all this information along with

the methodologies used to examine and interpret it. RO also aims to create software to bring computational tools to the bench scientists and to make possible precise searches for all RNA information. With this aggregate of taxonomies, tools, and common vocabulary, scientists with different backgrounds, experiences, and interests can discuss their favorite RNA using a meaningful, mutually intelligible dialect. The ROC invites those interested in RNA to contribute to the on-going discussion of RNA ontology through their Community Discussion Board (<http://roc.bgsu.edu>) **GM**

Distinguishing Cathepsins

Cathepsins are cysteine proteases whose expression levels and activity are increased in human and mouse cancer cells. In some tumors, cathepsins are also mislocalized to the cell surface and can be secreted into the extracellular space. There are 11 cathepsin family members but their individual functions in cancer and normal cells are not clear. In a recent study, Gocheva *et al.* (*Genes Dev.* 2006, 20, 543–556) used a genetic approach to define the roles of four cathepsin genes in cancer cells.

The author generated knockouts of four different cathepsin genes in cancer bearing mice and examined the phenotypes of the mutants. Mutations in cathepsins B and S impaired the formation of tumors and angiogenesis. Removing cathepsin B, L or S resulted in reduced tumor growth. Absence of any one of these three cathepsins impaired tumor invasion. In contrast, removing cathepsin C had no significant effect on the tumor parameters examined. Further analysis showed that cathepsin B, L, and S cleaved E-cadherin, a protein involved in cell adhesion, *in vitro*, but cathepsin C did not. In mice lacking cathepsin B, L, and S, the levels of E-cadherin levels are maintained. The authors propose that the cathepsin B, L, and S-mediated tumor invasion is likely due to loss of cell adhesion resulting from the cleavage of E-cadherin. The molecular mechanism by which each cathepsin mediates other aspects of tumorigenesis requires additional studies. However, the data from Gocheva *et al.* may be useful in designing inhibitors that target specific cathepsins and thereby specific stages of cancer progression. **EJ**

A Life or Death Decision

Cellular response to the proinflammatory cytokine tumor necrosis factor (TNF) can result in cell survival or cell death, depending on the circumstances. TNF activates the c-Jun NH²-terminal kinase (JNK) signaling pathway, but understanding how JNK determines the cell's destiny requires deciphering its apparent dual role in mediating both life and death. TNF-induced activation of

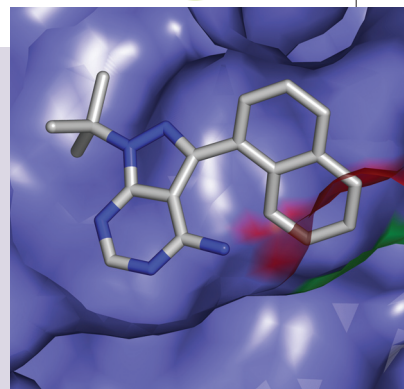


Image courtesy of R. Davis

JNK is characterized by an early, robust but transient phase followed by a late, low-level but sustained period. Ventura *et al.* (*Mol. Cell* 2006, 21, 701–710) have used chemical genetics to investigate how the time course of JNK activity affects the fate of the cell.

The redundancy of kinases in the cell has hampered the ability to find specific small molecule inhibitors to probe kinase function, and JNK is no exception. To circumvent this problem, the researchers created mutant JNK (m-JNK) containing an enlarged ATP binding pocket and an inhibitor, 1-naphthylmethyl-4-amino-1-*tert*-butyl-3-(*p*-methylphenyl)pyrazolo[3,4-*d*]pyrimidine (1NM-PP1), that binds specifically to this pocket but not to wild-type JNK or other kinases. The researchers also created murine embryonic fibroblasts (MEFs) that express wild-type JNK, no JNK, or m-JNK to facilitate interpretation of the specific function of JNK in the cell.

The temporal role of JNK in apoptosis and survival pathways was investigated using these unique biological tools. The authors found that JNK activity during the late phase of JNK signaling mediates proapoptotic signaling. In contrast, it was observed that the early phase of TNF-induced JNK activation is critical for signaling survival. These data explain the apparent paradoxical roles of JNK in the cell and indicate that a cell's fate is critically dependent on the temporal regulation of this kinase. The details of how JNK controls each signaling process will require additional mechanistic studies. **EG**

UPCOMING CONFERENCES

Channels, Receptors and Synapses

Cold Spring Harbor Meeting
April 18–21, 2006
Cold Spring Harbor Laboratory, NY

Molecular Chaperones and the Heat Shock Response

May 3–7, 2006
Cold Spring Harbor Laboratory, NY

First Annual Meeting

Japanese Society for Chemical Biology
May 8–9, 2006
Tokyo, Japan

Yale Chemical Biology Symposium

May 12, 2006
New Haven, CT

71st symposium :

Regulatory RNAs
May 31–June 5, 2006
Cold Spring Harbor Laboratory, NY

RNA 2006

Annual Meeting of the RNA Society
June 20–25, 2006
Seattle, WA

Spotlights written by Eva Gordon, Evelyn Jabri, Grace Miller, Sarah Tegen, and Jason Underwood.

Seeing Is Believing: Real-Time Cellular Activity Assay for Phospholipase A₂

Wonhwa Cho*

Department of Chemistry (M/C 111), 845 West Taylor Street, University of Illinois at Chicago, Chicago, Illinois 60607-7061

A majority of cellular enzymes work in concert with other interacting proteins or function as part of a multiprotein complex (1). Also, cellular activities of enzymes are tightly and dynamically regulated in a spatiotemporally specific manner by different mechanisms, including reversible post-translation modification and binding to small molecules. Therefore, real-time activity assays of cellular enzymes within the context of their natural physiological environment, that is, cells, tissues, or whole organisms, are an essential tool for elucidating their physiological function and regulation and also for developing and validating specific enzyme inhibitors as drug candidates.

Real-time cell assays pose a much greater technical challenge than *in vitro* assays because the activity of a particular enzyme must be measured in a highly heterogeneous and dynamic environment, which entails a sensitive, rapid, specific, and high-resolution detection and quantification of signals. In general, fluorescence measurements meet these requirements, and thus, microscope-based assays using fluorogenic substrates have been the most popular means of real-time cell assays. Fluorescence-based, real-time cellular activity assays have been developed for several classes of enzymes (2), including proteases, protein kinases, phosphatases, and phospholipases.

Phospholipase A₂ (PLA₂) is one of several phospholipases whose cellular actions have been investigated by means of fluorescence-based, real-time cellular assays (3–7). PLA₂ is a superfamily of

intracellular (calcium-dependent and -independent) and secreted enzymes that catalyze the hydrolysis of the *sn*-2 ester of phospholipids (Figure 1). Since PLA₂'s have been implicated in various physiological processes (*e.g.*, inflammation) and diseases (*e.g.*, rheumatoid arthritis, asthma, etc.) through the production of bioactive fatty acids (*e.g.*, arachidonic acid) and lysophospholipids, these enzymes have been extensively studied. However, determination of physiological functions and regulation of PLA₂ isoforms and the sites of their cellular action has been elusive. Therefore, development of a set of fluorogenic substrates that can be specifically recognized by individual PLA₂ isoforms in the cell would be a boon to PLA₂ research.

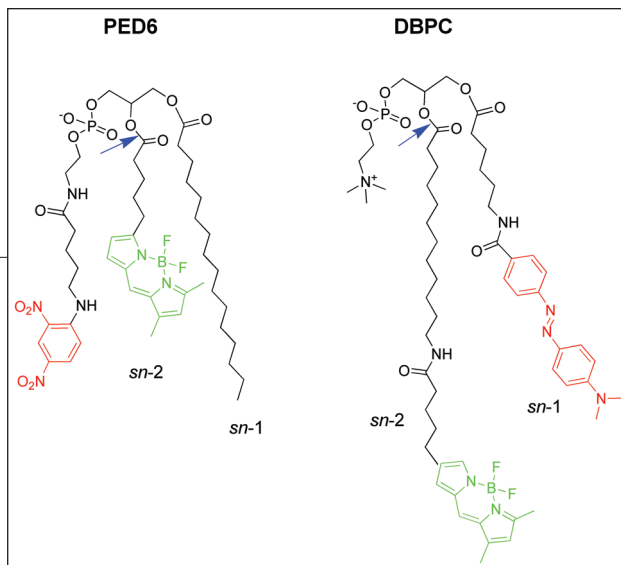
PLA₂'s can have two types of substrate selectivity, headgroup selectivity and *sn*-2 acyl group selectivity. In general, PLA₂'s have diverse headgroup selectivity but exhibit a lesser degree of *sn*-2 acyl group selectivity; only group IVA cytosolic PLA₂ (cPLA₂α) shows pronounced selectivity for the *sn*-2 arachidonoyl group (8, 9). The first generation fluorogenic substrate for the cell PLA₂ assay, *N*-((6-(2,4-dinitrophenyl)amino)hexanoyl)-1-hexadecanoyl-2-(BODIPY-pentanoyl)-*sn*-glycero-3-phosphoethanolamine (PED6) (3), contains a BODIPY fluorophore in the *sn*-2 acyl chain, the fluorescence of which is quenched by a 2,4-dinitrophenyl moiety appended to the lipid headgroup. The release of the BODIPY fluorophore by PLA₂ catalysis would relieve this quenching, hence, the increase in fluorescence intensity (Figure 1). Although recent cell studies using PED6 and its

ABSTRACT Quantitative real-time *in situ* activity assays are necessary for determining the physiological function and regulation of enzymes. A paper in this issue reports the synthesis of a series of new fluorogenic phospholipids that allow fast real-time measurements of cellular activity and head group selectivity of an important family of enzyme, phospholipases.

*To whom correspondence should be addressed.
E-mail: wcho@uic.edu.

Published online March 17, 2006
10.1021/cb600079t CCC: \$33.50
© 2006 by American Chemical Society

Figure 1. Structures of fluorogenic substrates used for the real-time cell activity assay of PLA₂. PED6 has a BODIDY fluorophore (green) in the *sn*-2 acyl group whose fluorescence is quenched by a 2,4-dinitrophenyl moiety (red) in the *sn*-3 headgroup. DBPC also contains a BODIPY group (green) appended to the end of *sn*-2 acyl chain, but in this molecule, the BODIPY fluorescence is quenched by a Dabcyl group (red) in the *sn*-1 acyl chain. This design allows variation in the headgroup structure and thereby a more versatile PLA₂ activity assay. For both PED6 and DBPC, PLA₂-catalyzed hydrolysis of *sn*-2 ester (blue arrows), which liberates fatty acid and lysophospholipid as products, relieves the fluorescence quenching and thus enhances the emission intensity of BODIPY. For DBPC, phospholipase A₁ which catalyzes the hydrolysis of the *sn*-1 acyl group can also give the same fluorescence signal.



derivative have provided valuable information about the function and regulation of different PLA₂'s (3–7), these substrates allow determination of neither headgroup nor acyl group selectivity of PLA₂'s. Thus, it is difficult to identify the PLA₂ responsible for the hydrolysis of these substrates unless specific enzyme inhibition or overexpression is carried out independently. Also, some PLA₂'s, such as cPLA₂α, show extremely low activity toward these phospholipids due to a bulky, unnatural headgroup (6, 7).

On page 83 of this issue of *ACS Chemical Biology*, Rose and Prestwich report a series of new fluorogenic phospholipids that allow not only robust and fast *in vitro* determination of headgroup selectivity of PLA₂ isoforms but also a real-time cellular activity measurement for a diverse group of PLA₂'s, including cPLA₂α (10). A new twist is that in these substrates the fluorescence of the BODIPY fluorophore appended to the *sn*-2 acyl chain is quenched by a Dabcyl group attached to the *sn*-1 acyl chain (Figure 1), allowing variation in the headgroup structure from phosphatidylcholine (1-(6-*p*-methyl red)aminohexanoyl)-2-*O*-(12-(5-BODIDY-pentanoyl)aminododecanoyl)-*sn*-3-glycerophosphocholine; DBPC), -ethanolamine, and -glycerol to phosphatic acid (DBPA). Therefore, these substrates

should be extremely valuable for measuring the physiological activities of those PLA₂ isoforms with strong headgroup selectivity. For example, the formation of potent bioactive lipids, lysophosphatidic acid and lysophosphatidylcholine by PA- and PC-specific PLA₂ isoforms, respectively, can be directly monitored under physiological conditions using DBPA and DBPC, respectively, as substrates.

Undoubtedly, the work by Rose and Prestwich (10) represents important progress toward developing an *in vitro* or cell-based, high-throughput screening system for PLA₂ inhibitors. However, some key issues still need to be addressed to establish a robust and versatile *quantitative* real-time cell assay of PLA₂'s. First, current substrate designs make it difficult to incorporate the *sn*-2 arachidonoyl group in fluorogenic substrates. As a result, cPLA₂α, which plays a key role in cellular arachidonic acid formation (11, 12) due to its unique *sn*-2 arachidonoyl selectivity (8, 9), cannot be specifically assayed. Another important factor that applies to all cell-based assays is the cellular localization of fluorogenic substrates. Because PLA₂'s, cPLA₂α in particular, act in a spatially specific manner (13), uniform distribution of fluorogenic lipids over cell membranes is essential for unbiased monitoring of PLA₂ activities and

for determining the sites of PLA₂ actions. Unfortunately, the cellular distribution of fluorogenic phospholipids cannot be readily controlled or monitored. Finally, current fluorogenic phospholipids allow only qualitative to semiquantitative activity measurements because neither ratiometric measurement nor substrate concentration determination through calibration is feasible. Improvement in fluorogenic substrate design as well as in fluorescence microscopy detection would lead to the development of a true quantitative real-time *in situ* assay for PLA₂ (and other phospholipases) that can be used for functional and mechanistic studies of PLA₂'s and PLA₂-specific drug development in various cells, tissues, and whole organisms.

REFERENCES

1. Cho, W. (2006) *Sci. STKE* 2006, pe7.
2. Baruch, A., Jeffery, D. A., and Bogoy, M. (2004) *Trends Cell Biol.* 14, 29–35.
3. Hendrickson, H. S., Hendrickson, E. K., Johnson, I. D., and Farber, S. A. (1999) *Anal. Biochem.* 276, 27–35.
4. Farber, S. A., Olson, E. S., Clark, J. D., and Halpern, M. E. (1999) *J. Biol. Chem.* 274, 19338–19346.
5. Farber, S. A., Paack, M., Ho, S. Y., Johnson, I. D., Wagner, D. S., Dosch, R., Mullins, M. C., Hendrickson, H. S., Hendrickson, E. K., and Halpern, M. E. (2001) *Science* 292, 1385–1388.
6. Kim, Y. J., Kim, K. P., Rhee, H. J., Das, S., Rafter, J. D., Oh, Y. S., and Cho, W. (2002) *J. Biol. Chem.* 277, 9358–13174.
7. Wijewickrama, G. T., Kim, J. H., Kim, Y. J., Abraham, A., Oh, Y., Ananthanarayanan, B., Kwatia, M., Ackerman, S. J., and Cho, W. (2006) *J. Biol. Chem.*, published online Feb 13, <http://dx.doi.org/10.1074/jbc.M512657200>.
8. Clark, J. D., Lin, L. L., Kriz, R. W., Ramesha, C. S., Sultzman, L. A., Lin, A. Y., Milona, N., and Knopf, J. L. (1991) *Cell* 65, 1043–1051.
9. Sharp, J. D., White, D. L., Chiou, X. G., Gooden, T., Gamboa, G. C., McClure, D., Burgett, S., Hoskin, J., Skatrud, P. L., Sportsman, J. R., Becker, G. W., Kang, L. H., Roberts, E. F., and Kramer, R. M. (1991) *J. Biol. Chem.* 266, 14850–14853.
10. Rose, T. M., and Prestwich, G. D. (2006) *ACS Chem. Biol.* 1, 83–92.
11. Bonventre, J. V., Huang, Z., Taheri, M. R., O'Leary, E., Li, E., Moskowitz, M. A., and Sapirstein, A. (1997) *Nature* 390, 622–625.
12. Uozumi, N., Kume, K., Nagase, T., Nakatani, N., Ishii, S., Tashiro, F., Komagata, Y., Maki, K., Ikuta, K., Ouchi, Y., Miyazaki, J., and Shimizu, T. (1997) *Nature* 390, 618–622.
13. Evans, J. H., Spencer, D. M., Zweifach, A., and Leslie, C. C. (2001) *J. Biol. Chem.* 276, 30150–30160.

Intracellular Metal Detectors

Amy M. Barrios*

Department of Chemistry, University of Southern California, Los Angeles, California 90089

Zinc is the second most abundant “trace” element in the body with an estimated 2–3 g distributed among over 1000 different proteins (1, 2). Zinc plays important roles in numerous biological processes including gene expression, apoptosis, enzyme regulation, immune system modulation and metabolism, to name a few (3). Although the total cellular concentration of zinc is approximately 200 μM (2), most is tightly bound to biomolecules and is not readily accessible to chelation by small molecules (3). However, it has been established that in specialized cells in the brain, pancreas, and prostate, significant pools of readily exchangeable zinc are accumulated (3). Because of the well-established roles of protein-bound and exchangeable zinc in both healthy and diseased tissue, the question of how much zinc is available in a cell has emerged at the forefront of chemical biology. On page 103 of this issue of *ACS Chemical Biology*, Bozym *et al.* report the first use of a fluorescence resonance energy transfer (FRET) based ratiometric zinc sensor to directly image and quantify the concentration of zinc in resting eukaryotic cells (4).

Interest in the development of reagents to sense intracellular and extracellular zinc levels has skyrocketed over the past decade, resulting in a multitude of probes based on molecules, peptides, or proteins that exhibit zinc-dependent fluorescence (5). These probes exhibit various dynamic ranges, excitation and emission profiles, and cellular permeabilities and localize differently in cells. Biological metal ion sensing continues to be the subject of

intense research in part because it is clear that no one probe will be ideal for all applications. However, there are certain features that are desirable in zinc probes for a biological toolkit. First, the probes must exhibit excellent selectivity for zinc over metal ions such as calcium and magnesium, which are present in cells at a much higher concentration. Second, probes should be useful over a broad concentration range, for use in imaging the very low resting concentrations of zinc present in most cells as well as the high resting zinc concentrations in certain specialized cell types. Third, probes may also need fast zinc association and dissociation rates in order to facilitate dynamic imaging of changes in zinc concentrations over time and in response to various stimuli. Fourth, the probes should ideally localize in the organelle of interest, distribute evenly throughout the cell, or remain in the extracellular environment, depending on the experiment. Finally, the probes should not be cytotoxic or perturb the resting state of the cells.

Despite recent advances in intracellular zinc sensing, it remains difficult to quantify the concentration of exchangeable zinc in resting eukaryotic cells that do not accumulate large amounts of zinc. The zinc biosensor reported by Bozym *et al.* (4) is based on FRET from a zinc-bound aryl sulfonamide to a fluorophore covalently attached to carbonic anhydrase (CA), an enzyme with exquisite selectivity and sensitivity to zinc (Figure 1) (4). By labeling CA with a TAT cell-penetrating peptide, the authors are able to introduce the zinc sensor into both PC-12 (a rat pheochromo-

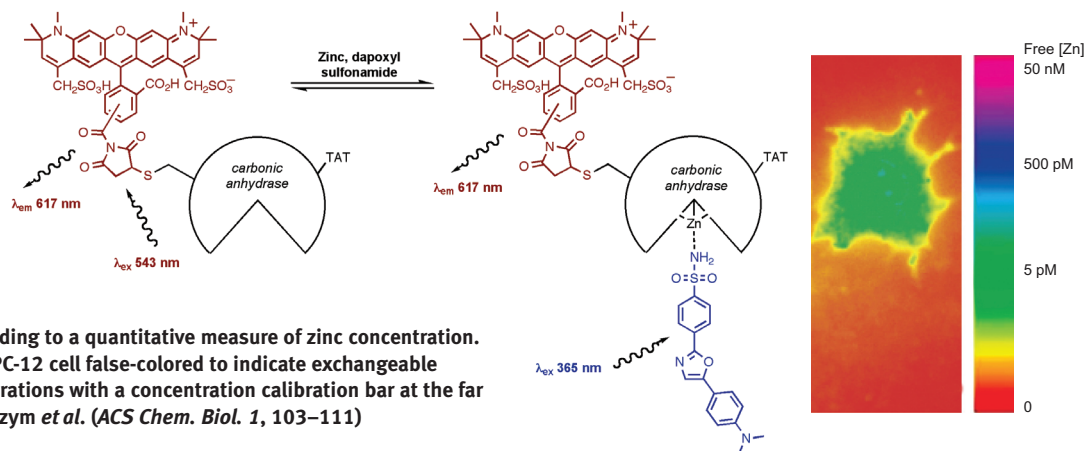
ABSTRACT Metal ions play numerous crucial roles in biology, and there is great interest in obtaining an accurate measurement of the intracellular concentrations of both tightly bound and exchangeable metal ions. Measuring the concentration of readily exchangeable transition metal ions in a cell has been particularly difficult because of the extremely small concentrations involved, interference from other metal ions and biomolecules, and the challenge of introducing probes into the cell with minimal perturbations. Recent work has made quantification of the intracellular exchangeable zinc pool possible for the first time using a cell-permeable, ratiometric, fluorescence resonance energy transfer based zinc biosensor.

*To whom correspondence should be addressed.
E-mail: amy.barrios@usc.edu.

Published online March 17, 2006
10.1021/cb600080a CCC: \$33.50
© 2006 by American Chemical Society

With the right toolkit, the tasks of quantifying metal ion concentrations and imaging changes in metal concentrations in response to external stimuli, intracellular signaling events, physiological and pathological processes will become routine.

Figure 1. Ratiometric zinc determination using Alexa Fluor 594-modified carbonic anhydrase and dapoxyl sulfonamide. The emission intensity at 617 nm following excitation at 365 nm is normalized to the emission intensity at 617 nm following excitation at 543 nm, leading to a quantitative measure of zinc concentration. On the right is shown a PC-12 cell false-colored to indicate exchangeable cytoplasmic zinc concentrations with a concentration calibration bar at the far right. Cell image from Bozym *et al.* (*ACS Chem. Biol.* 1, 103–111)



cytoma cell line) and CHO (Chinese hamster ovary) cells with no apparent ill effects, circumventing the need for microinjection, electroporation, or other, less “gentle” methods of introducing sensors into cells, one of the major difficulties in this field. The zinc concentration inside the cells is quantitatively measured from the ratio of fluorescence intensities at two different excitation wavelengths (6). The advantages of ratiometric imaging methods include minimal background from variations in excitation intensity, specimen thickness, and fluorophore concentration (7).

With a CA-based zinc sensor, the intracellular concentration of available zinc is measured at approximately 5 pM throughout both types of cells (Figure 1, inset). This concentration, while quite low, is significantly higher than the femtomolar levels proposed for prokaryotic cells (8). In these experiments, the zinc seems to be evenly distributed throughout the cytoplasm. One unexpected result of this research is the rapid equilibration between apoprotein and metal-bound protein in the cells. *In vitro*, the zinc-binding kinetics of apo-carbonic anhydrase are quite slow, and equilibration of the apoprotein with picomolar levels of intracellular exchangeable zinc would be expected to require several hours. In the imaging experiments, however, the equilibrium seemed to occur in minutes. This

suggests that catalysis of zinc insertion into the CA may be occurring.

This work represents a critical step forward in sensing intracellular metal ion concentrations in that it provides the first example of a quantitative measurement of zinc levels in eukaryotic cells in a resting state. As is often the case, the results raise even more questions and highlight the prospects for future research in this field. For example, is 5 pM a standard level of exchangeable zinc in many types of differentiated eukaryotic cells? Is this zinc really distributed evenly throughout the cytoplasm, and how do exchangeable zinc levels differ in various subcellular compartments? What effects, if any, do the probes have on the cells?

We can expect many more exciting breakthroughs in sensing biological metal ion concentrations over the next several years, building upon recent advances in the field. Genetically encoded biosensors, “reagentless” sensing systems, and cell-permeable small molecule sensors are all promising directions to pursue. Chemists are charged with designing novel probes with high selectivity, optimal binding affinities, and on/off rates that can be selectively delivered to a desired subcellular compartment. Techniques that allow improved spatial resolution as well as reliable quantification of analyte concentrations will

be of paramount importance. Biologists are presented with the challenge of developing “gentle” methods for delivering sensors into cells, elucidating the biological effects of the added sensors, and interpreting the results of metal ion concentration measurements. With the right toolkit, the tasks of quantifying metal ion concentrations and imaging changes in metal concentrations in response to external stimuli, intracellular signaling events, physiological and pathological processes will become routine.

REFERENCES

- Berg, J. M., and Shi, Y. (1996) The galvanization of biology: a growing appreciation for the roles of zinc, *Science* 271, 1081–1085.
- Maret, W. (2001) Zinc biochemistry, physiology, and homeostasis — recent insights and current trends, *BioMetals* 14, 187–190.
- Vallee, B. L., and Falchuk, K. H. (1993) The biochemical basis of zinc physiology, *Physiol. Rev.* 73, 79–118.
- Bozym, R., Thompson, R., Stoddard, A., and Fierke, C. (2006) Measuring picomolar intracellular exchangeable zinc in PC-12 cells using a ratiometric fluorescence biosensor, *ACS Chem. Biol.* 1, 103–111.
- Kikuchi, K., Komatsu, K., and Nagano, T. (2004) Zinc sensing for cellular application, *Curr. Opin. Chem. Biol.* 8, 182–191.
- Fierke, C. A., and Thompson, R. B. (2001) Fluorescence-based biosensing of zinc using carbonic anhydrase, *BioMetals* 14, 205–222.
- Tsien, R. Y. (1989) Fluorescent Probes of Cell Signaling in *Annual review of neuroscience*, pp 227–253, Annual Review, Inc., Palo Alto, CA.
- Finney, L. A., and O'Halloran, T. V. (2003) Transition metal speciation in the cell: insights from the chemistry of metal ion receptors, *Science* 300, 931–936.

Linking Genome Structure and Function through Specific Histone Acetylation

Jeffrey C. Hansen*

Department of Biochemistry and Molecular Biology, 1870 Campus Delivery, Colorado State University, Fort Collins, Colorado 80523

Efforts to understand histone acetylation lie at the crossroads of chemical and biological research. It is well-established that addition of a two-carbon acetyl group to a specific lysine side chain on one of the core histone N-terminal “tail” domains (NTDs) can stimulate an entire biological process such as gene transcription (1, 2). However, we know almost nothing about the molecular mechanism(s) that transduce the change in the local chemistry of the histone tail domain into potent biological regulation. A very recent publication by Shagren-Knack *et al.* (3) addresses the molecular basis of acetylation function by determining how specific acetylation of K16 of the H4 NTD influences the salt-dependent condensation of model nucleosomal arrays and native chromatin fibers *in vitro*. The exciting results obtained by Shagren-Knaak *et al.* (3) show that K16 acetylation completely inhibits the ability of the H4 NTD to mediate chromatin fiber condensation and that H4 K16 acetylation is a potent modulator of the nucleosome–nucleosome interactions that drive chromatin fiber condensation. These findings strongly implicate changes in genome architecture as a primary mechanism through which K16acet regulates processes such as transcription. At the molecular level, these results indicate that conversion of a single, strategically placed lysine side chain from a positively charged amine to a neutral hydrophobic moiety is sufficient to completely abolish the ability of the H4 NTD to engage in nucleosome–nucleosome interactions.

Core Histone NTDs, Chromatin Fiber Condensation, and Genome Architecture.

A schematic depiction of the hierarchical organization of interphase chromosomes is shown in Figure 1. Chromosomal DNA is assembled with repetitively spaced core histone octamers into nucleosomal arrays. A short stretch of nucleosomal arrays in many ways can be thought of as the “subunit” of an interphase chromosome. Nucleosomal arrays that are complexed with non-histone proteins are called chromatin fibers (hence, there are many different specific types of chromatin fibers embedded within chromosomes). Local nucleosome–nucleosome interactions result in formation of a highly condensed irregular helical structure traditionally termed “the 30 nm fiber” (4). Longer range organizational levels beyond 30 nm fibers have been well-documented in chromosomes (5), although little is known at the biochemical level about how they are assembled and maintained.

In vitro studies of chromatin fiber dynamics have been performed for over 30 years. Early studies were limited by the use of heterogeneous endogenous chromatin fragments and the at times staggering complexity of the system; even a short 12-mer array of nucleosomes is composed of nearly 100 histone proteins and 2500 bp of DNA, has a mass of 3 million Da, and exists in many different conformational states in solution. Progress during the last 2 decades has been driven by the availability of model systems that can be assembled *in vitro* from defined sequence

ABSTRACT A recent publication shows that a simple chemical event, acetylation of lysine 16 on the histone H4 N-terminal tail domain (NTD), completely abolishes the ability of the H4 NTD to mediate the nucleosome–nucleosome interactions involved in chromatin condensation. This result provides novel insight into the molecular mechanism of histone acetylation and also implicates H4 K16acet-dependent changes in chromatin fiber architecture as a central mechanism for generating transcriptionally active genomic domains.

*To whom correspondence should be addressed.
E-mail: jeffrey.c.hansen@colostate.edu.

Published online March 17, 2006
10.1021/cb6000894 CCC: \$33.50
© 2006 by American Chemical Society

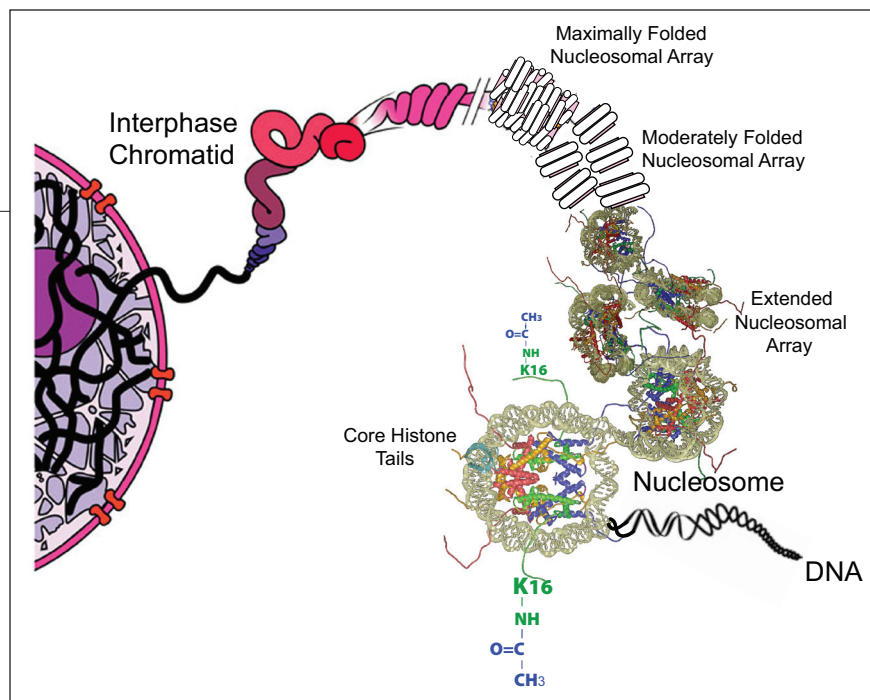


Figure 1. Schematic illustration of the interphase chromosome organization. Extended, moderately folded, and maximally folded nucleosomal arrays are discussed in detail in the text. The first nucleosome (lower right-hand corner) is shown as having lysine 16 acetylation on each H4 NTD. The acetyl groups are shown in blue.

DNA and pure histones, in effect yielding a homogenous preparation of length and compositionally defined chromatin fibers. Moreover, the recent use of native and mutant recombinant core histones (6–9) has opened the doors to a much better understanding of the histone contributions to chromatin fiber architecture.

Model system studies have provided the following essential background for the Shagren-Knaak *et al.* article (3): (a) under physiological ionic conditions, nucleosomal arrays are in equilibrium between extended (“beads-on-a-string”), moderately folded, extensively folded, and oligomeric conformational states (4); (b) the extensively folded “30 nm” chromatin fiber appears to be a two-start helix (6), although the detailed structure remains a mystery; (c) assembly of both folded and oligomeric nucleosomal arrays requires cation-dependent DNA charge neutralization *and* additional functions mediated by the core histone NTDs (4); (d) the H4 NTD mediates local nucleosome–nucleosome interactions by binding to a specific acidic domain formed by H2A/H2B on the surface of adjacent nucleosomes (7–9); and (e) a threshold level of nonspecific core histone acetylation inhibits the nucleosome–

nucleosome interactions involved in array condensation (4). Other equally important studies have shown that H4 K16acet is associated with transcriptionally active euchromatic domains *in vivo* (10–12). With this as background, we can now better understand the importance of the central questions addressed by Shagren-Knaak *et al.* (3). Does H4 K16acet inhibit folding and/or oligomerization of model nucleosomal arrays and endogenous, transcriptionally active chromatin fragments? Is specific acetylation a biochemical switch that destabilizes the repressive nucleosome–nucleosome interactions involved in chromatin condensation? Is modulation of genome architecture one of the mechanisms through which specific acetylation accomplishes its biological functions?

The Key to Success: Overcoming Technical Challenges. As is so often the case, surmounting technical barriers paved the way for success. Specifically, it was necessary to assemble preparations of defined nucleosomal arrays that uniformly contained H4 NTDs acetylated only on lysine 16. Shagren-Knaak *et al.* (3) conquered this problem by using recombinant core histones together with a “native chemical ligation” strategy. In

this protocol, a peptide was synthesized that consisted of residues 1–22 of H4 and was acetylated on K16. This peptide was chemically ligated to a recombinantly expressed H4 fragment (residues 23–104) to yield full-length H4 that was acetylated exclusively on K16. This acetylated H4 was mixed with recombinant full-length H2A, H2B, and H3 using standard protocols to yield H4 K16acet core histone octamers. Defined 12-mer nucleosomal arrays were assembled from core histone octamers and nucleosome positioning DNA using classical salt dialysis reconstitution. Three different types of histone octamers were used for the reconstitutions: wild-type, H4 K16acet, and H4 NTD[−] (octamers lacking H4 residues 1–22). The latter provides a critical control that defined the condensation behavior of arrays that lack a functional H4 NTD. Whereas assembly of wild-type and H4 NTD[−] arrays has been accomplished previously and is straightforward (6–9), assembly of the model K16acet arrays represented an elegant solution to a very difficult problem, one that allowed direct determination of the structural effects of a histone modification that increased the mass of a 3 MDa nucleosomal array by roughly 1/100 of 1%! This technical rigor laid the foundation for the success of the subsequent biochemical and biophysical experiments discussed below.

H4 K16acet Is a Potent Chemical Switch That Regulates Genome Architecture.

To assay for salt-dependent folding, native and H4 K16acet and H4 NTD[−] nucleosomal arrays in the absence and presence of 1 mM Mg²⁺ were analyzed by sedimentation velocity in the analytical ultracentrifuge. Nucleosomal arrays (12-mer) undergo a change in sedimentation coefficient from ~30 to ~55 S as they progress from the fully extended to maximally folded conformations, making sedimentation velocity an ideal assay for defining the extent of compaction under any given set of solution conditions (4). The exciting result obtained

by Shagren-Knaak *et al.* (3) was that H4 K16acet arrays sedimented identically as the H4 NTD⁻ arrays and were incapable of forming the maximally folded 30 nm structures formed by the native nucleosomal arrays under the same ionic conditions. Subsaturated arrays (*i.e.*, <12 octamers per DNA) cannot condense beyond the moderately folded conformation due to the nucleosome-free gaps in the array (4). Subsaturated H4 K16acet and H4 NTD⁻ arrays also sedimented identically in Mg²⁺ and much slower than native arrays, indicating that K16acet also completely abolished the ability of the H4 NTD to interact with adjacent nucleosomes. These intriguing results indicate that specific H4 K16acet is as effective at inhibiting the ability of H4 to mediate local nucleosome–nucleosome interactions as is the removal of the 22 N-terminal most H4 residues.

Oligomerization of model nucleosomal arrays is reversible, highly cooperative, and produces very large aggregates. This structural transition can be rapidly and effectively characterized by a simple differential centrifugation pelleting assay (4). A great deal of indirect evidence suggests oligomerization is related to organizational levels beyond 30 nm fibers in chromosomes (4). Shagren-Knaak *et al.* (3) found that H4 K16acet and H4 NTD⁻ arrays yielded identical oligomerization profiles and required more Mg²⁺ than wild-type arrays to achieve comparable levels of oligomerization. Thus, as with folding, H4 K16acet was as disruptive to oligomerization as was complete removal of the H4 NTD. A complimentary approach was taken to address the physiological relevance of the oligomerization results. Nuclease-digested HeLa cell chromatin fragments were mixed with different MgCl₂ concentrations, the samples pelleted by centrifugation, and the H4 K16acet levels determined in the supernatant (Mg²⁺-“soluble”) and pellet fractions. Results showed that the endogenous

chromatin that remained in the supernatant in 0.5–2.0 mM MgCl₂ was enriched in H4 K16acet. The Mg²⁺-solubility experiment is essentially identical to performing an oligomerization assay at a single Mg²⁺ concentration. Hence, H4 K16acet also appears to inhibit the Mg²⁺-dependent oligomerization of native chromatin fragments isolated from intact nuclei.

Taken together, the biochemical and biophysical studies of Shagren-Knaak *et al.* (3) document that H4 K16acet serves as a molecular switch that regulates genome architecture by controlling the ability of the H4 NTD to mediate local and global nucleosome–nucleosome interactions.

Chemical and Biological Ramifications.

The experiments of Shagren-Knaak *et al.* (3) have many important implications relating to eukaryotic genome structure and function. For chemists, they raise the fundamental question of how a simple two-carbon addition to a lysine side chain can completely abolish the ability of the H4 NTD to engage in nucleosome–nucleosome interactions. Because of their high positive charge density, the core histone NTDs historically were believed to mediate chromatin condensation by binding to DNA and neutralizing backbone charge (4). The finding that a reduction in a single positive charge is sufficient to completely abolish H4 NTD function is perhaps the strongest direct evidence against nonspecific DNA binding. Instead, an increasingly large body of evidence points to a mechanism involving interaction of residues 14–19 of the H4 NTD with a specific “acidic patch” domain formed by H2A and H2B on the surface of neighboring nucleosomes (7–9). If one assumes that the H4 NTD interacts with the acidic charge patch similarly to that of the viral LANA peptide (13), it makes sense that K16 acetylation could abolish nucleosome–nucleosome interactions. Acetylation not only neutralizes a positive charge that may be needed to interact with the acidic domain, but also caps the lysine

side chain with a hydrophobic methyl group that will further destabilize this interaction. In this manner, a simple chemical event, H4 K16acet, can act as a switch that helps regulate the stability of condensed chromatin in any given region of the genome.

The popular histone code hypothesis states that specific patterns of core histone post-translational modifications are “marks” that are recognized by functionally important chromatin-associated proteins (1, 2). Obviously, the results of Shagren-Knaak *et al.* (3) cannot be explained by the histone code hypothesis. Instead, H4 K16acet appears to function by fundamentally altering the local protein chemistry of the H4 NTD, perhaps in conjunction with the intrinsically disordered nature of the domain (14). The extent to which other histone post-translational modifications function through alteration of the local chemical environment of the NTD remains to be determined and will be a fruitful area of future research.

Extensively folded nucleosomal arrays and chromatin fibers are generally considered to be repressive to transcription initiation and elongation (4, 15). Although the H4 NTD has a central role, each of the four *unmodified* core histone NTDs contribute significantly to the extensive network of nucleosome–nucleosome interactions that help stabilize condensed chromatin fiber structures (16). Any given nucleosome in a chromatin fiber has the potential to engage in 16 distinct NTD-dependent interactions with neighboring nucleosomes, eight involving its own NTDs interacting with other nucleosomes and eight from other nucleosomes interacting with that nucleosome. Assuming that there are specific post-translational modifications that can switch off the chromatin-condensing functions of the H2A, H2B, and H3 NTDs in the same manner as K16acet does for the H4 NTD, the local network of nucleosome–nucleosome interactions will allow for fine-tuned biological regulation.

The local network of nucleosome–nucleosome interactions will allow for fine-tuned biological regulation.

One extreme is the extensively condensed, biologically inactive fiber structures formed when all the NTDs are unmodified. At the other extreme is the extended, more functionally active conformation. In between is an entire spectrum of condensed chromatin structures whose stability and biological activity will be determined by which particular NTDs are unmodified in any given region of the genome. In this manner, chromatin fiber architecture can exert an exquisitely sensitive, rheostat-like control over transcriptional activity and other biological processes that take place in a chromatin milieu.

REFERENCES

1. Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* **15**, 172–183.
2. Peterson, C. L., and Laniel, M. A. (2004) Histones and histone modifications. *Curr. Biol.* **14**, R546–R551.
3. Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847.
4. Hansen, J. C. (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 361–392.
5. Kireeva, N., Lakonishok, M., Kireev, I., Hirano, T., and Belmont, A. S. (2004) Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J. Cell Biol.* **166**, 775–785.
6. Schalch, T., Duda, S., and Richmond, T. J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571–1573.
7. Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J. Mol. Biol.* **327**, 85–96.
8. Schalch, T., Duda, S., and Richmond, T. J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138–141.
9. Fan, J. Y., Rangasamy, D., Luger, K., and Tremethick, D. J. (2004) H2A.Z alters the nucleosome surface to promote hp1alpha-mediated chromatin fiber folding. *Mol. Cell* **16**, 655–661.
10. Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (1997) *mof*, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* **16**, 2054–2060.
11. Akhtar, A., and Becker, P. B. (2000) Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* **5**, 367–375.
12. Suka, N., Luo, K., and Grunstein, M. (2002) Sir2p and Sas2p oppositely regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat. Genet.* **32**, 378–383.
13. Barbera, A. J., Chodaparambil, J. V., Kelley-Clarke, B., Joukov, V., Walter, J. C., Luger, K., and Kaye, K. M. (2006) The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* **311**, 856–861.
14. Hansen, J. C., Lu, X., Ross, E. D., and Woody, R. W. (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* **281**, 1853–1856.
15. Horn, P. J., and Peterson, C. P. (2002) Molecular biology. Chromatin higher order folding: wrapping up transcription. *Science* **297**, 1824–1827.
16. Gordon, F., Luger, K., and Hansen, J. C. (2005) The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. *J. Biol. Chem.* **280**, 33701–33706.

Raising Enzymes from the Dead and the Secrets They Can Tell

Michael A. Marletta*

Departments of Chemistry and Molecular and Cell Biology, University of California, Berkeley, California 94720-1460

We often struggle with how to extend beautifully done *in vitro* experiments with clear molecular explanations into the much more complicated environs of cells and complex organisms. Would it not be great if we could disable a protein of interest, an enzyme for example, and then turn it on at will? Would it not be even better if the control over activity were reversible? This is exactly what Cole and colleagues (1) have accomplished with an inactive mutant of the proto-oncogene, tyrosine kinase *src* that they chemically rescued in cells using the small molecule imidazole. This could easily just fall into the neat trick category; however, their results include two surprises. First, it appears that *Src* is active under basal conditions, an observation that is certain to have impact upon the signaling field. Second, they turn up new substrates for the kinase including CrkL (chicken tumor virus no. 10 [CT10] regulator of kinase). As the authors point out “CrkL is a particularly intriguing *Src* target because of its well-established role in cytoskeletal signaling and its known *Src* connections”.

Src is a tyrosine kinase and the first of that family to be described. The story begins with Peyton Rous and what he endured to convince the community that what was later identified and named Rous sarcoma virus was involved in tumor formation. In time it was shown that *v-src* is an oncogene and essential to the transforming properties of the virus. It is an understatement to say that finding the cellular homologue *c-src* was a major step forward along with the later identifica-

tion of it as a proto-oncogene. Many key discoveries followed and continue up to the present. The Nobel Prizes to Rous, Varmus, and Bishop certainly add to the luster of the great discoveries. Steve Martin has written two very nice accounts of how we have gotten to this point (2, 3). Despite all that has been uncovered, *Src* proteins are tyrosine kinases, how mutations in *src* can lead to transformed cells, and much more, functional questions remain unanswered. The findings by Cole and colleagues provide new insight into the long-standing questions.

Site-directed mutagenesis has been a powerful tool, particularly in the probing enzyme mechanism (4). Mutagenesis to change enzyme specificity represents an important early contribution showing that enzyme engineering was indeed possible. Though many examples exist, the proteases remain among the most instructive (5). The replacement of key protein side-chains that lead to little or no activity and the regaining of activity by building that missing, chemically important residue into the substrate, termed chemical complementation, was an important forerunner to chemical rescue (6). This led the way to chemical rescue with what has also been termed substrate-assisted catalysis, creation of an inactive site directed mutant of a residue involved in catalysis and the rescue of activity by providing a substrate analog the contains chemical functionality that is missing from the mutant residue (7). In this work, Carter and Wells removed the catalytic histidine from *B. subtilis* subtilisin and replaced it with alanine. The enzyme

ABSTRACT The complexity of partners and reaction sequence has made the deciphering eukaryotic signaling pathways particularly difficult. Various approaches have yielded important results. A recent paper reports on an advance that uses chemical rescue of inactive mutant of *Src*. The rescue was done in cells, and the results obtained showing new aspects of *Src* function point the way toward a general use of this chemical tool in the sorting out of complex biological processes.

*To whom correspondence should be addressed.
E-mail: marletta@berkeley.edu.

Published online March 17, 2006

10.1021/cb600110g CCC: \$33.50

© 2006 by American Chemical Society

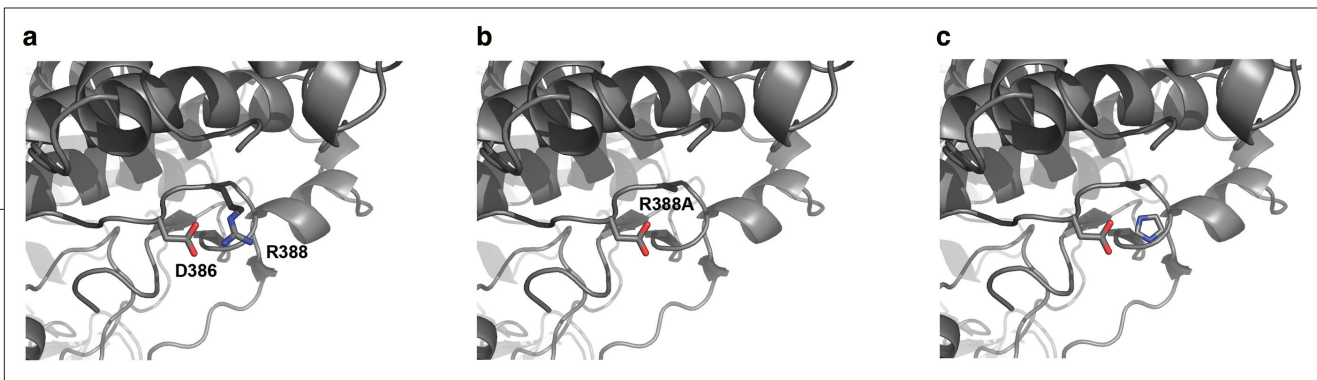


Figure 1. Chemical rescue of Src. a) wild type c-Src showing the active site D386 and R388 (PDB id entry 2PTK); b) model of R388A made by removal of R388 except for the β -carbon. This mutation renders the enzyme inactive; and c) model shown in b) with imidazole placed in the site occupied by R388 in the wild type protein. The addition of imidazole resurrects catalytic activity. The figures were made by Douglas Mitchell at University of California-Berkeley using PyMol (<http://www.pymol.org>).

was severely wounded (k_{cat}/K_m was down a million-fold); however a substrate with a histidine in the P2 position complemented the missing catalytic histidine and significant activity was restored.

The first true chemical rescue was carried out by Toney and Kirsch on aspartate aminotransferase, a pyridoxal phosphate dependent enzyme (8). In this remarkable paper, they removed a critical lysine residue that acts a general base in the essential proton transfer steps in the reaction and replaced it with alanine. The resulting mutant was dead but could be rescued by the addition of exogenous amines. Toney and Kirsch ultimately used 11 different amines and were able to do the first true Brønsted analysis of proton transfer reactions in enzymes (8).

Now, Cole and colleagues (1) used imidazole to rescue an inactive Src mutant, R388A. Arginine 388 is conserved in the Src family and participates in hydrogen bonds with the substrate tyrosine hydroxyl group and an aspartate active site residue, again conserved in the Src family. The inactive mutant is then rescued by imidazole playing the part of the mutated arginine (1)

Imidazole rescue has received attention before, and these past studies have shown the utility of using it to decipher complex reactions and in heme cofactor binding. For example, Ortiz de Montellano and colleagues (9) restored catalytic activity to an inactive mutant (H25A) of heme oxygenase with imidazole. In this case, the added imidazole replaces the proximal His ligand to the heme when bound as a substrate. Goodin and co-workers carried out similar studies with cytochrome c peroxidase

(10). Others have rescued heme cofactor binding with imidazole where the histidine mutation has taken away the proximal ligand to the heme (11, 12). Expression without imidazole leads to apoprotein, and so these latter two studies also strongly suggest that His-iron coordination is an early step in cofactor incorporation, perhaps early along the folding path.

The choice of imidazole to replace arginine by Cole and colleagues (1) circumvents the transport and toxicity issues of guanidines. Imidazole itself is not without toxicity, but they were able to carry out the studies without cellular toxicity as a complicating problem. It is somewhat surprising that 30–50% of the wild-type activity is regained with imidazole when attempting to replace arginine, but it works. With the cellular experiments, Cole and colleagues (1) have taken rescue to a new level and they chose a target where there are plenty of important questions to answer. Despite years of intense study, the cellular functions of Src remain elusive. The ability to control the enzyme activity in cells without laborious conditional knockouts should open new doors, and it has. Addition of imidazole in the absence of growth factor stimulation still led to phosphorylation of tyrosine, so c-Src is basally active. Future experiments that combine stimulation with control of activity will be very useful in sorting out complex signaling pathways. Perhaps most importantly, the authors identify new substrates (Crkl, lamin A/C, and procollagen). Findings like these will prove invaluable to a complete understanding of the complex signaling pathways regulated by proteins like Src.

Chemical rescue is clearly complementary to the use of inhibitors and has some advantages. The paper from Cole and colleagues (1) illustrates the power of the method and points the way toward future discoveries.

REFERENCES

1. Qiao, Y., Molina, H., Pandey, A., Zhang, J., and Cole, P. A. (2006) Chemical rescue of a mutant enzyme in living cells, *Science* 311, 1293–1297.
2. Martin, G. S. (2001) The hunting of the Src, *Nat. Rev. Mol. Cell Biol.* 2, 467–475.
3. Martin, G. S. (2004) The road to Src, *Oncogene* 23, 7910–7917.
4. Peracchi, A. (2001) Enzyme catalysis: removing chemically “essential” residues by site-directed mutagenesis, *Trends Biochem. Sci.* 26, 497–503.
5. Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., and Rutter, W. J. (1985) Redesigning trypsin: alteration of substrate specificity, *Science* 228, 291–297.
6. Hwang, Y. W., and Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein, *J. Biol. Chem.* 262, 13081–13085.
7. Carter, P., and Wells, J. A. (1987) Engineering enzyme specificity by “substrate-assisted catalysis”, *Science* 237, 394–399.
8. Toney, M. D., and Kirsch, J. F. (1989) Direct Brønsted analysis of the restoration of activity to a mutant enzyme by exogenous amines, *Science* 243, 1485–1488.
9. Wilks, A., Sun, J., Loehr, T. M., and Ortiz de Montellano, P. R. (1995) Heme oxygenase His25Ala mutant: replacement of the proximal histidine iron ligand by exogenous bases restores catalytic activity, *J. Am. Chem. Soc.* 117, 2925–2926.
10. McRee, D. E., Jensen, G. M., Fitzgerald, M. M., Siegel, H. A., and Goodin, D. B. (1994) Construction of a bisquo heme enzyme and binding by exogenous ligands, *Proc. Natl. Acad. Sci. U.S.A.* 91, 12847–12851.
11. Barrick, D. (1994) Replacement of the proximal ligand of sperm whale myoglobin with free imidazole in the mutant His-93→Gly, *Biochemistry* 33, 6546–6554.
12. Zhao, Y., Schelvis, J. P., Babcock, G. T., and Marletta, M. A. (1998) Identification of histidine 105 in the beta1 subunit of soluble guanylate cyclase as the heme proximal ligand, *Biochemistry* 37, 4502–4509.

Histone Demethylation by Hydroxylation: Chemistry in Action

Jessica Schneider[†] and Ali Shilatifard^{†,‡,*}

[†]Department of Biochemistry, Saint Louis University School of Medicine, 1402 South Grand Blvd., St. Louis, Missouri 63104, and [‡]Saint Louis University Cancer Center, Saint Louis University School of Medicine, St. Louis, Missouri 63104

Eukaryotic cells wrap their chromosomal DNA around octamers of histones to form nucleosomes. This process is a crucial first step in compacting and packaging DNA into the nucleus. In addition to a role in DNA packaging, nucleosomal assembly and disassembly regulate accessibility of the transcriptional machinery to gene coding and regulatory regions.

Electron microscopy studies have shown chromatin as a series of “beads on a string,” with the “string” as linker DNA and the “beads” as individual nucleosomes consisting of eight core histone proteins (two each of H3, H4, H2A, and H2B) (1, 2) (Figure 1a). The core histones are wrapped by 147 base pairs of DNA (1.65 turns around the histone octamer), forming the intact nucleosome (3). The amino termini of histone tails protrude away from the nucleosome and are therefore available for interactions with DNA and other proteins and many histone tail-modifying enzymes (3). The posttranslational modifications of histone tails so far include: acetylation, phosphorylation, ubiquitination, and methylation. Multiple modifications can decorate each histone tail, and some amino acids within the histone tail can be modified in several different ways (4, 5). In fact, the combinatorial effect of such histone tail modifications can serve to elicit a multitude of different responses. This “epigenetic regulation” denotes an inherited state of gene regulation that is independent of the genetic information encoded within DNA itself.

Lysine or arginine residues within histones can be posttranslationally modified via the enzymatic addition of methyl groups from the donor S-adenosylmethionine (SAM) (6). A class of enzymes containing SET domains [this domain takes its name from the *Drosophila* proteins *Su(var)3-9*, *Enhancer of zeste (E(z))*, and *trithorax (trx)*] can catalyze methylation of the lysine residue of

ABSTRACT Histone methylation plays an essential role in epigenetic regulation and has been thought to be an irreversible and stable modification of histones. However, several enzymes have recently been discovered to demethylate mono- and dimethylated lysine residues of histone H3 as well as monomethylated arginines via either amine oxidation or deimination, respectively. The JmjC domain-containing histone demethylase 1 (JHDM1), which is conserved from yeast to human, has been demonstrated to demethylate mono- and di- but not trimethylated H3 K36 via hydroxylation of the methyl moiety within the methylated lysine residue. This study broadens our understanding of different types of reaction mechanisms and cofactor requirements for a different category of histone demethylating machinery.

*To whom correspondence should be addressed.
E-mail: shilatia@slu.edu.

Received for review January 26, 2006
and accepted February 21, 2006

Published online March 17, 2006

10.1021/cb600030b CCC: \$33.50

© 2006 American Chemical Society

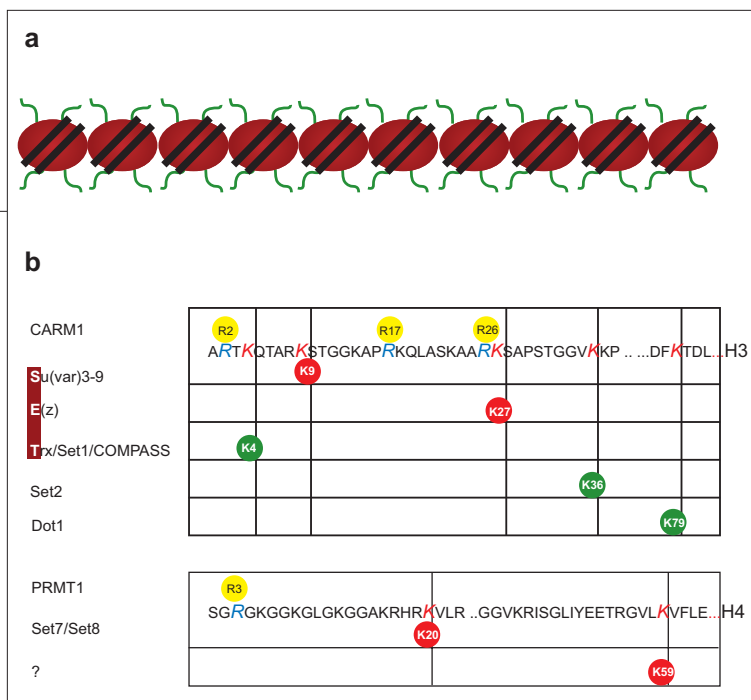


Figure 1. The histone tails. a) Nucleosomes are involved in processes ranging from DNA compaction to transcriptional regulation. They are also considered to be major carriers of epigenetically inherited information as well. As shown, the repeating nucleosomes with intervening “linker” DNA form the 10-nm fiber, known descriptively as “beads on a string”. The histone tails protrude away from each nucleosome and therefore are available for interactions with DNA and other proteins and many histone tail-modifying enzymes. b) The site of modification by methylation and the enzymatic machinery involved in histone tail methylation. The N-terminal amino acid sequences of histones H3 and H4 along with positions of specific methylation sites and the known enzymes required for such modifications are shown above.

histones on the ϵ -nitrogen (Figure 1b). The SET domain-containing enzymes are specific in their substrate selection, unlike histone acetyltransferases. Therefore, methylation of each lysine residue within histones requires a specific SET domain-containing enzyme (4, 6). For example, the Set1 protein in yeast, which is found as a component of the macromolecular complex

COMPASS, is specific for methylation of lysine 4 of histone H3 (Figure 1b) (7–9). Even its human counterpart, the MLL protein, exists in a similar macromolecular complex and has the same histone substrate selectivity as COMPASS (10, 11). Other lysines within histones, such as lysine 36 of histone H3, require the catalytic activity of other SET domain-containing enzymes, such as Set2 (Figure 1b) (12). In addition to SET domain-containing histone methyl-

transferases, there are also non-SET domain-containing enzymes capable of methylating lysine residues within histones. The enzyme Dot1 (disruptor of telomeric silencing 1) is a histone lysine methyltransferase that lacks the characteristic SET domain, and it has been shown to methylate the lysine residue of histone H3 on lysine 79 (Figure 1b) (13–15).

Following the identification of enzymatic machinery capable of methylating the lysine residues within histone proteins, it was demonstrated that the methylation of histones is a stable and irreversible modification. The relatively high stability of lysine methylation compared to other reversible modifications such as phosphorylation and acetylation was in part attributed to the fact that the N-CH₃ bond is thermodynamically highly stable. Furthermore, experimental studies demonstrated that once methylated, the lysine residues exhibit a half-life similar to that of unmodified histones (16, 17). For example, H3 K9 methylation is required for the regulation of epigenetic silencing and the maintenance of heterochromatin; therefore, one would expect this modification to be static rather than dynamic (18, 19). In addition to the importance of particular sites being methylated on histones, the state of methylation also plays an essential role, as lysine residues may be either mono-, di-, or trimethylated (20, 21). Transition from the tri- to dimethyl or the di- to monomethyl form may be required for the proper response to developmental or environmental signals. Therefore, regulation of the transition from fully methylated to partially or fully unmethylated lysine residues within histones is likely to be indispensable for the biological outcomes associated with histone methylation; regulation of this process could be one of the main reasons for the presence of histone demethylating machinery.

Although histone methylation is considered stable, several different histone demethylating enzymes have recently been identified to reverse some, but not all, forms of histone methylation (Figure 2) (22). PADI4 gene product was discovered as a histone deiminase that antagonizes histone arginine methylation (23, 24). It was demonstrated that either free or monomethylated arginine can be cleaved at the guanidine C-N bond (Figure 2a) by the arginine deiminase PADI4, generating the products citrulline and methylammonium. Although PADI4 is capable of deimination of free and monomethyl arginine, dimethylation of arginines prevents deimination by PADI4. Although

KEYWORDS

Nucleosome: Fundamental repeating unit in chromosomes made up of DNA and histone proteins. Nucleosomes are found in eukaryotic nuclei and appear as bead-like structures along the DNA when viewed by electron microscopy. Routinely referred to as the “beads on a string.”

Chromatin: Mass of genetic material located in the cell nucleus containing DNA and proteins that condenses to form chromosomes in a eukaryotic cell.

Epigenetic information: Heritable changes in gene function not encoded by eukaryotic chromosomal DNA.

Histone methylation: Posttranslational modifications occurring on lysine or arginine residues within histone tails that carry epigenetic information.

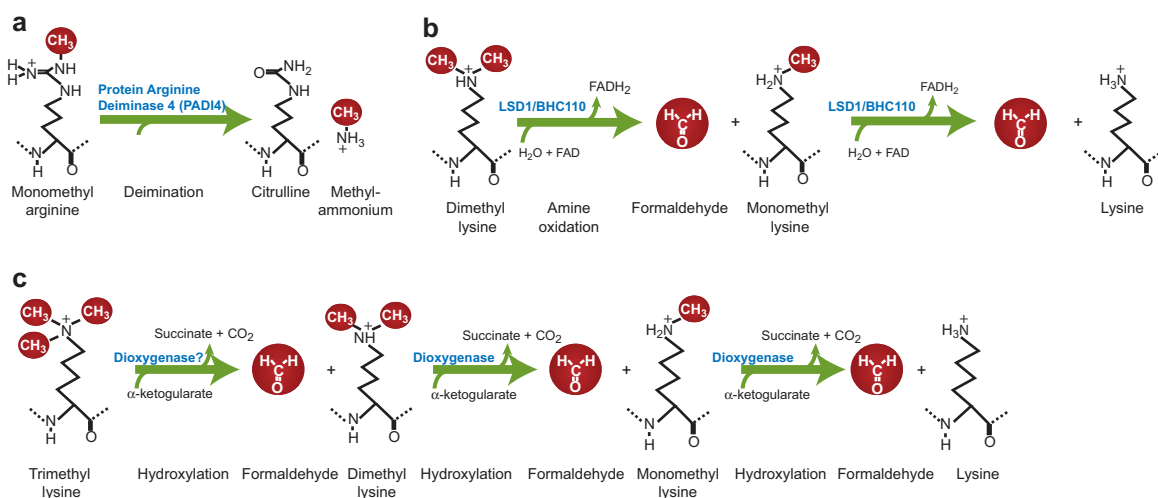


Figure 2. Enzymatic mechanisms required to remove the methyl moiety from modified histones. **a)** It has been demonstrated that the arginine and monomethylated arginine can serve as substrates for protein arginine deiminase 4 (PADI4). PADI4 can reverse arginine methylation by cleaving the guanidino C-N bond via the deamination mechanism. The byproduct of such reaction is citrulline and methyl-ammonium. Lysine methylation can be reversed via either **b)** amine oxidation or **c)** direct hydroxylation of the methyl moiety. The byproducts of both reaction mechanisms in **b** and **c** are unmodified lysine and formaldehyde. Since the formation of an imine intermediate via transfer of two hydrogen atoms to FAD requires a protonated nitrogen, amine oxidation can only demethylate mono- and dimethylated lysine substrates. Since trimethylated lysine has been found in nature, it has been proposed that hydroxylation can represent an alternative demethylating mechanism. In this mechanism, a direct radical attack on the methyl-carbon by Fe(II) in α -ketoglutarate-dependent dioxxygenases such as the JmjC-domain containing proteins can lead to the formation of an unstable carbinolamine, which result in the generation of unmodified lysine and formaldehyde. It has been proposed that this reaction mechanism can be employed to demethylate trimethylated lysine residues.

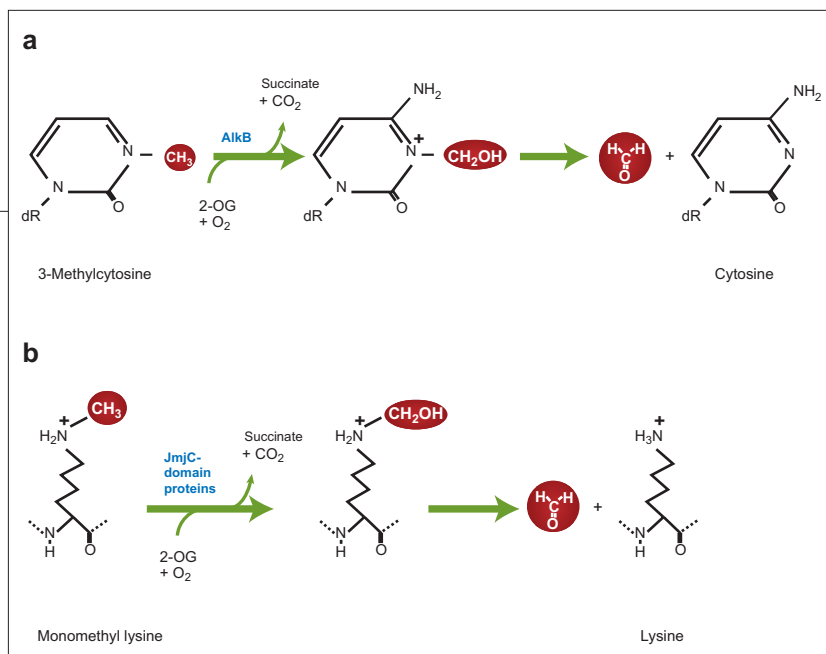
PADI4 can prevent dimethylation of arginine residues, its discovery does not fully address how cells manage dimethylated arginine residues. Other possibilities include enzymatic machinery capable of demethylating dimethylated arginine or the involvement of histone replacement machinery.

Amine oxidation was the first mechanism to be proposed for lysine demethylation (Figure 2b). The BHC110/LSD1 protein, a histone H3 K4 demethylase (25), is a riboflavin-binding protein, which is a member of a FAD-dependent enzyme family. BHC110/LSD1 is highly conserved in organisms ranging from *Schizosaccharomyces pombe* to human. The amine oxidase domain of this enzyme is found in its carboxyl-terminal end. This enzyme also contains a SWIRM domain, which is a protein-protein interaction domain found in several chromatin-associated proteins. The demethylation of mono- and dimethylated lysine 4 residue of histone H3 by BHC110/LSD1 is an oxidation reaction that requires the presence of the cofactor

flavin adenine dinucleotide (FAD). Formaldehyde and an unmodified lysine residue are the byproducts of the enzymatic reaction catalyzed by BHC110/LSD1 (25). Since the formation of an imine intermediate requires a protonated nitrogen, BHC110/LSD1 can only demethylate mono- and dimethylated lysine residues and not the trimethylated form.

Biochemical investigations have demonstrated that BHC110/LSD1 can associate with the androgen receptor and can act as a coactivator for transcription (26). Surprisingly, the interaction of BHC110/LSD1 with the androgen receptor overlaps with the specific loss of H3 K9 methylation from the androgen-receptor DNA element, with little to no effect on H3 K4 methylation within the same region (26). Furthermore, in the presence of the androgen receptor but not its absence, affinity-purified BHC110/LSD1 can catalyze the demethylation of H3 K9. This indicates that the specificity of BHC110/LSD1 towards its histone substrate can be modulated through its interaction with

Figure 3. Demethylation via hydroxylation. a) A mechanism for demethylation of 3-methylcytosine by AlkB has already been demonstrated. b) Allshire and colleagues have proposed that JmjC domain-containing proteins can function via the same mechanism to demethylate methylated lysine residues on proteins (32). A recent study supports a role for JmjC domain-containing proteins in demethylation of mono- and di- but not trimethylated lysine residues on histones (34).



other cofactors. In addition to the interaction with the androgen receptor, BHC110/LSD1 has been shown to exist in a large macromolecular complex containing the CoREST corepressor complex. Association of BHC110/LSD1 within the CoREST complex can increase the demethylation of histone H3 lysine 4 (H3K4) by nearly 5-fold as compared to that of recombinant BHC110/LSD1 alone (27, 28). More importantly, recombinant BHC110/LSD1 is unable to demethylate H3 K4 on nucleosomes. However, within its complex, nucleosomes are readily demethylated by BHC110/LSD1 (27, 28). On the basis of its biochemical interaction with the CoREST complex and other biological studies, it has been proposed that BHC110/LSD1, within the CoREST complex, is recruited to genes containing the REST-responsive element to participate in gene silencing by demethylating K4 mono- and dimethylated histone H3 within the REST-responsive repressor element.

Although BHC110/LSD1 is capable of demethylating both K4 and K9 of histone H3, several observations have suggested the possibility for the presence of additional demethylases that employ other chemical mechanisms to demethylate lysine residues within proteins. Such observations include that (a) it is chemically impossible for BHC110/LSD1 to demethylate trimethylated H3 K4, but trimethylated K4 exists from yeast to human; (b) the number of homologues of BHC110/LSD1 are limited within a given organism relative to the large number of modified histone residues; and (c) BHC110/LSD1 homologues do not exist in *S. cerevisiae*, but H3 K4 can be mono-, di-, and trimethylated by COMPASS *in vivo* (8, 21, 29).

In theory, the mono-, di-, and trimethylated lysine residues can be demethylated by hydroxylation of

the methyl group. This reaction could be catalyzed by dioxygenases that use Fe(II) in their catalytic center, resulting in hydroxylation of the methyl group and subsequent demethylation (Figure 3). Such an observation has already been reported for the DNA repair demethylase AlkB (Figure 3a) (30, 31). AlkB is a 2-oxoglutarate (2OG)-Fe(II)-dependent dioxygenase, which uses Fe(II) at its active site in order to activate a molecule of dioxygen to form a highly reactive oxoferryl species and to hydroxylate the methyl group of certain forms of damage-induced DNA methylation (30, 31). This oxidized product is highly unstable and can be readily released as formaldehyde, resulting in the release of the methyl group from DNA.

Recently, it was suggested by Allshire and colleagues (32) that Epe1 (33), which is genetically required for the integrity of heterochromatin in *Schizosaccharomyces pombe*, could be a histone demethylase functioning via hydroxylation (Figure 3b). Allshire and colleagues have modeled Epe1 onto the known structure of the factor inhibiting hypoxia inducible factor, which belongs to the 2-OG-Fe(II)-dependent dioxygenase family (32). This family of enzymes can catalyze two electron oxidations using iron in their catalytic core and 2-OG as cosubstrate. Furthermore, sequence alignment of Epe1 and its related family members, the JmjC family, show that Epe1 and many of the JmjC-domain proteins share distinctive features of 2-OG-Fe(II)-dependent dioxygenases. Allshire and colleagues have therefore suggested that JmjC-domain proteins may function as protein demethylases capable of demethylating mono-, di- and trimethylated lysine residues within proteins. Indeed, a recent study by Tsukada *et al.* (34) demonstrated that JHDM1 (JmjC domain-containing histone demethylase 1), a protein

conserved from yeast to human, specifically demethylates mono- or dimethylated histone H3 yet is not capable of demethylating trimethylated nucleosomal substrates either *in vitro* or *in vivo*.

In search for enzymatic activity capable of demethylating nucleosomal substrates, Tsukada *et al.* (34) identified an activity in HeLa cell nuclear extracts capable of demethylating radio-labelled nucleosomal substrates methylated on K36 of histone H3. This enzymatic assay was set up to detect the release of radioactive formaldehyde from K36 methylated histone H3, which was enzymatically synthesized using Set2 and radioactive SAM. To increase the sensitivity of the assay, the radioactive formaldehyde was converted to radioactive 3,5-diacetyl-1,4-dihydrolutidine to facilitate its extraction in the organic phase away from the substrate. Using this assay, Tsukada *et al.* (34) have identified the F-box and leucine-rich repeat protein 11 (FBXL11) as the enzyme capable of demethylating K36 methylated histone H3.

FBXL11 was identified in a bioinformatic search for F-box containing proteins (35). In addition to the F-box domain, the FBXL11 contains three leucine-rich repeats at its C-terminal domain, CxxC and PHD domains at its center, and a JmjC domain towards its N-terminal domain. Mutational analyses of FBXL11 indicated that its JmjC domain is required for the demethylase activity associated with this protein and that the deletion of other domains such as the CxxC, PHD, and leucine-rich domains partially impair its activity. In their report, Allshire and colleagues predicted that amino acids required for coordinating Fe(II) within the dioxygenases should be required for demethylase activity (32). Indeed, a single point mutation of histidine 212 within FBXL11, which is highly conserved among FBXL11 orthologues and is predicted to coordinate Fe(II) in the catalytic center of the enzyme, abolishes the demethylase activity associated with this enzyme. On the basis of the demethylase activity associated with FBXL11, Tsukada *et al.* (34) have renamed this protein JHDM1, which will be used to describe this enzyme henceforth.

To determine the *in vitro* specificity of JHDM1, Tsukada *et al.* (34) tested the enzymatic activity of recombinant JHDM1 towards several methylated nucleosomal substrates differentially methylated on K4, K9, K27, K36, and K79 of histone H3, as well as K20 and R3 of histone H4. This study indicated that JHDM1 preferentially demethylates dimethylated

H3 K36 and is not able to demethylate trimethylated H3 K36. Furthermore, analysis of the consequence of JHDM1 overexpression in 293 T-cells indicated that this enzyme is capable of demethylating dimethylated H3 K36 *in vivo* as well.

The study by Tsukada *et al.* (34) supports the proposed model that JmjC domain-containing proteins demethylate methylated lysine residues within histones via hydroxylation. An advantage of this reaction mechanism is that all forms of methylated lysine residues (mono-, di-, and trimethylated forms) can be demethylated. Enzymatic analyses of JHDM1 indicate that this enzyme is not capable of demethylating trimethylated nucleosomal substrates. This observation can be explained in several ways. First, JHDM1 may require the association with other factors to be able to demethylate trimethylated lysine residues. Regulation of the catalytic activity of an enzyme by its interacting factors within a complex is not unprecedented in chromatin biology. For example, several components of COMPASS, Cps60 and Cps40, are required for histone H3 K4 trimethylation by Set1/COMPASS (21). Therefore, it is also possible that interaction of other factors with JHDM1 regulates its enzymatic activity. A second possibility is that the activity of JHDM1 may be regulated via direct posttranscriptional modification of JHDM1. Since trimethylation for several different sites on H3 appears to have different patterns of localization than the corresponding mono- or dimethyl form, it is possible that JHDM1 demethylase activity may be regulated towards these sites. A third possibility

is that demethylation of trimethylated histones may require the presence of other modifications on the same or other histones. An example for such mechanism was shown by the activity of Rad6/Bre1 in histone H2B monoubiquitination and signaling for histone H3 methylation by COMPASS (29, 36, 37). A fourth possibility is the catalytic pocket of JHDM1 may be too small to accommodate trimethylated histones, and other JmjC domain-containing proteins

KEYWORDS

Histone methyltransferases: Factors/proteins that transfer methyl groups from donor S-adenosylmethionine (SAM) to lysine or arginine residues on histones.

SET-domain: (*Drosophila* Su(var)3-9, Enhancer of zeste (E(z)), and trithorax (trx)). A domain found within a class of histone methyltransferase that catalyzes lysine residue methylation of histones on the ϵ -nitrogen.

COMPASS: (Complex Proteins Associated with Set1). A SET domain-containing complex capable of mono- di- and tri-methylating lysine 4 of histone H3. Its human homologue, the MLL complex, is also found in a COMPASS-like complex capable of methylating lysine 4 of histone H3.

Histone demethylases: Factors/proteins that remove the methyl moiety from methylated lysine or arginine residues on histones.

JmjC domain-containing proteins demethylate methylated lysine residues within histones via hydroxylation

could function in this process. Finally, it is also feasible to consider that once lysine residues within histones are trimethylated, such marks are permanent and irreversible.

Recent studies reported by Tsukada *et al.* (34) extend our knowledge regarding the mechanism of posttranslational modification of histones by methylation. This study demonstrated the existence of reaction mechanisms and cofactor requirements fundamentally different from previously reported histone demethylases. Because JmjC domain-containing proteins are conserved from yeast to human and large numbers of JmjC homologues exist within an organism, this class

of enzymes could be involved in modulating a diverse range of existing histone and protein modifications via methylation. Future genetic and biological studies aiming to define the role of JmjC domain-containing proteins in the regulation of gene expression should shed further light on the role of the JmjC superfamily in pathways involving lysine methylation in histones and other proteins.

Acknowledgment: The authors would like to thank Kristy Wendt and Adam Wood for editorial assistance and critical reading of this review. The work in A.S.'s laboratory is supported by grants from the National Institutes of Health (2R01CA089455 and 1R01GM069905) and the American Cancer Society. A.S. is a Scholar of the Leukemia and Lymphoma Society.

REFERENCES

1. Komberg, R. D. (1974) Chromatin structure: a repeating unit of histones and DNA, *Science* 184, 868–871.
2. Komberg, R. D., and Lorch, Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome, *Cell* 98, 285–294.
3. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature* 389, 251–260.
4. Wood, A., and Shilatifard, A. (2004) Posttranslational modifications of histones by methylation, *Adv. Protein Chem.* 6, 201–222.
5. Workman, J. L., and Kingston, R. E. (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation, *Annu. Rev. Biochem.* 67, 545–579.
6. Zhang, Y., and Reinberg, D. (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails, *Genes Dev.* 15, 2343–3460.
7. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein, *Proc. Natl. Acad. Sci. U.S.A.* 98, 12902–12907.
8. Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M., and Shilatifard, A. (2002) COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression, *J. Biol. Chem.* 277, 10753–10755.
9. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. (2001) The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4, *EMBO J.* 20, 7137–7148.
10. Tenney, K., and Shilatifard, A. (2005) A COMPASS in the voyage of defining the role of trithorax/MLL-containing complexes: linking leukemogenesis to covalent modifications of chromatin, *J. Cell. Biochem.* 95, 429–436.
11. Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A., Biondi, C. A., Kay, G. F., Hayward, N. K., Hess, J. L., and Meyerson, M. (2004) Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus, *Mol. Cell* 13, 587–597.
12. Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z. W., Bone, J. R., Caldwell, J. A., Mollah, S., Cook, R. G., Shabanowitz, J., Hunt, D. F., and Allis, C. D. (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression, *Mol. Cell. Biol.* 22, 1298–1306.
13. Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain, *Curr. Biol.* 12, 1052–1058.
14. Lacoste, N., Utley, R. T., Hunter, J. M., Poirier, G. G., and Cote, J. (2002) Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase, *J. Biol. Chem.* 277, 30421–30424.
15. van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core, *Cell* 109, 745–756.
16. Byvoet, P., Shepherd, G. R., Hardin, J. M., and Noland, B. J. (1972) The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells, *Arch. Biochem. Biophys.* 148, 558–567.
17. Duerre, J. A., and Lee, C. T. (1974) In vivo methylation and turnover of rat brain histones, *J. Neurochem.* 23, 541–547.
18. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases, *Nature* 406, 593–599.
19. Lachner, M., and Jenuwein, T. (2002) The many faces of histone lysine methylation, *Curr. Opin. Cell Biol.* 14, 286–298.
20. Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3, *Nature* 419, 407–411.
21. Schneider, J., Wood, A., Lee, J. S., Schuster, R., Dueker, J., Maguire, C., Swanson, S. K., Florens, L., Washburn, M. P., and Shilatifard, A. (2005) Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression, *Mol. Cell* 19, 849–856.
22. Bannister, A. J., Schneider, R., and Kouzarides, T. (2002) Histone methylation: dynamic or static? *Cell* 109, 801–806.
23. Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004) Histone demethylation antagonizes arginine methylation, *Cell* 118, 545–553.
24. Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D., and Coonrod, S. A. (2004) Human PAD4 regulates histone arginine methylation levels via demethyliminium, *Science* 306, 279–283.
25. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1, *Cell* 119, 941–953.

26. Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R., and Schule, R. (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription, *Nature* **437**, 436–439.
27. Lee, M. G., Wynder, C., Cooch, N., and Shiekhhattar, R. (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation, *Nature* **437**, 432–435.
28. Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T., Shi, Y. (2005) Regulation of LSD1 histone demethylase activity by its associated factors, *Mol. Cell* **19**, 857–864.
29. Dover, J., Schneider, J., Tawiah-Boateng, M. A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6, *J. Biol. Chem.* **277**, 28368–28371.
30. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*, *Nature* **419**, 178–182.
31. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage, *Nature* **419**, 174–178.
32. Trewick, S. C., McLaughli, P. J., and Allshire, R. C. (2005) Methylation: lost in hydroxylation? *EMBO Rep.* **6**, 315–320.
33. Clissold, P. M., and Ponting, C. P. (2001) JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2beta, *Trends Biochem Sci.* **26**, 7–9.
34. Tsukada, Y. I., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2005) Histone demethylation by a family of JmjC domain-containing proteins, *Nature* **439**, 811–816.
35. Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., and Nesvizhskii, A. (2004) The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data, *Mol. Cell. Proteomics* **3**, 531–533.
36. Wood, A., Krogan, N. J., Dover, J., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J. F., Johnston, M., and Shilatifard, A. (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter, *Mol. Cell* **11**, 267–274.
37. Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation, *Mol. Cell* **11**, 721–729.

Fluorogenic Phospholipids as Head Group-Selective Reporters of Phospholipase A Activity

Tyler M. Rose and Glenn D. Prestwich*

Department of Medicinal Chemistry and Center for Cell Signaling, University of Utah, 419 Wakara Way, Suite 205, Salt Lake City, Utah 84108

The PLA are an ever-expanding family of enzymes, both in numbers and newly recognized roles in cell signaling. The PLA₂ (phospholipase A₂) superfamily in mammals consists of over 19 enzymes (1), which are broadly divided into three groups: cytosolic (cPLA₂), secretory (sPLA₂), and Ca²⁺-independent (iPLA₂). PLA₂ isozymes catalyze the hydrolysis of *sn*-2 position acyl chains of phospholipids, while PLA₁ catalyze *sn*-1 cleavage.

All PLA reactions generate a free fatty acid and a lysophospholipid, and each product has the potential to mediate cellular responses. Arachidonic acid, for example, can be a precursor for pro-inflammatory eicosanoids (1). Lysophospholipids such as LPA (lysophosphatidic acid), known to be generated by the action of PLA₁ or PLA₂, often in concert with a PLD (phospholipase D) (2) or lysoPLD/ATX (3), are associated with a variety of cellular events (4). The widespread distribution of PLA in the human body and the bioactive nature of their cleavage products have implicated these enzymes in many human diseases, including autoimmune (5) and cardiovascular diseases (6), neurological disorders (7), and cancer (8).

To explore the effects of particular PLA isozymes in cell physiology, it is important to understand their spatiotemporal activation, diacyl group selectivity, and head group selectivity. For example, PLA isozymes have been reported with selectivities for PA (9), PE (10, 11), PC (12–14), PS (15, 16), and PG (17, 18), but information regarding the biological significance of the head group selectivity is limited. Furthermore, the pathophysiology of PLA activity in various diseases has made these enzymes important targets for isoform-specific drug development. In this case, rapid, sensitive, high-throughput, and real-time fluorescence-based activity

ABSTRACT PLA (phospholipases A) are important mediators of cell signaling, generating bioactive fatty acids and LPLs (lysophospholipids). PLA products having different head groups can initiate vastly different types of signaling. Fluorogenic analogues of the PLs (phospholipids) PA (phosphatidic acid), PC (phosphatidylcholine), PE (phosphatidylethanolamine), and PG (phosphatidylglycerol) were synthesized as PLA substrates for rapidly determining in real time the influence of head group modifications on cell signaling both *in vitro* and in cells. Enzyme-assisted remodeling of the *sn*-2 position of the diacylglycerol moiety with cobra venom PLA₂ and transphosphatidylation with a particular PLD (phospholipase D) were central steps in the preparation of these enzymatic probes. The resulting fluorogenic Dabcyl- and BODIPY-containing PL analogues, DBPA, DBPC, DBPE, and DBPG, were used in mixed micelle assays to determine PLA₂ kinetics. Next, the assays were used to determine the X_i(50) value of a common PLA₂ inhibitor. Finally, the head group selectivities of a series of commercially available PLA₂ enzymes were readily established using the DBPLs (Dabcyl-BODIPY PLs) as substrates.

*To whom correspondence should be addressed.
E-mail: Glenn.Prestwich@hsc.utah.edu.

Received for review November 11, 2005
and accepted January 27, 2006.

Published online March 3, 2006.

10.1021/cb5000014 CCC: \$33.50

© 2006 American Chemical Society

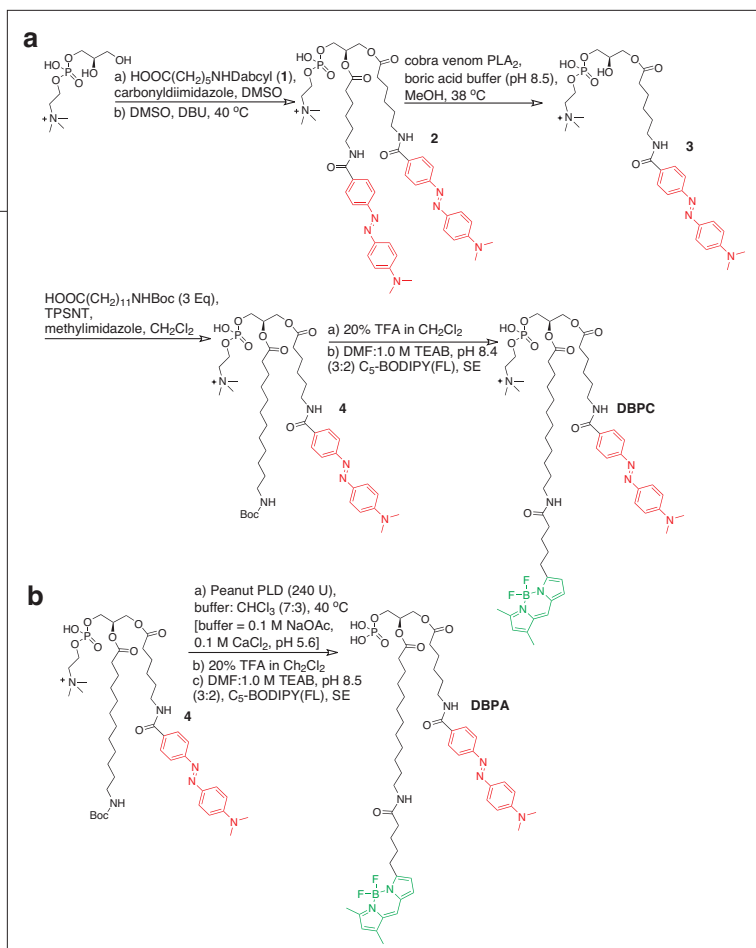


Figure 1. Enzyme-assisted route to DBPC (a) and DBPA (b). See Methods for experimental details.

assays are desirable which would be compatible with isolated enzymes or suitable for cell-based assays.

PLA activity has traditionally been monitored using radiometric, titrametric, or chromatographic endpoint analyses (19), all of which are time-consuming and often do not permit real-time monitoring. Chromogenic assays (20) allow real-time monitoring of PLA *in vitro*, but fail for *in situ* applications. Recently, fluorogenic PLA probes based on a PC skeleton have been developed by our group (21), as well as by others (22, 23). These fluorogenic probes provide rapid, sensitive, real-time monitoring of PLA activity *in vitro* and *in situ*.

We describe here the synthesis and evaluation of fluorogenic PLA probes with four different head groups, PA, PC, PE, and PG, by enzyme-assisted organic synthesis. Each fluorogenic substrate contains the same diacylglycerol moiety, in which the *sn*-1-acyl chain contains an attached fluorescence quencher (Dabcylyl, also known as *p*-methyl red), and the *sn*-2 acyl chain contains an appended BODIPY fluorophore. Intramolecular FRET (fluorescence resonance energy transfer) to the Dabcylyl group quenches BODIPY fluorescence until PLA-mediated substrate cleavage; then, a fluorophore is released when the lysolipid and fatty acid moieties are separated and the intermolecular distance exceeds that

required for efficient energy transfer. We demonstrate herein that these four DBPLs (Dabcylyl-BODIPY PLs), specifically DBPA, DBPC, DBPE, and DBPG, are suitable for *in vitro* monitoring of PLA activity, including applications for inhibitor screening and head group selectivity studies.

RESULTS AND DISCUSSION

The significance of head group selectivity among PLA enzymes is one aspect of their crucial roles in cell signaling that has not been studied in great detail. To address this unmet need, we synthesized fluorogenic phospholipid analogue probes having different head groups using an enzyme-assisted synthetic route. The resulting probes were validated using real-time continuous-monitoring *in vitro* assays. In addition, these substrates are being employed to identify spatio-temporal regulation of PLA activity in living cells.

Enzyme-Assisted Synthesis of DBPC and DBPA.

The original synthesis of DBPC (21) was modified to improve yields and permit access to a variety of phospholipid head groups. The new enzyme-assisted synthetic route is shown in Figure 1a and was based on a route used to prepare photoactive phosphatidic acid derivatives (24). Initial attempts to condense Dabcylyl-linked aminohexanoic acid (1) with PC-glycerol using DCC/DMAP conditions resulted in poor yields of a mixture of monoacyl and diacyl products. Other unsuccessful esterification conditions included use of Sc(OTf)₃ as a catalyst (25), elevation of reaction temperature, and conversion of the fatty acid to an acid chloride. Finally, acyl imidazole chemistry, previously shown to be effective at acylating PC-appended glycerol (26), provided the desired diacyl product 2 in acceptable yield (Figure 1a). After removal of the *sn*-2 fatty acid with cobra venom and re-esterification in high yield, the carbamate 4 was cleaved and the resulting primary amine was conjugated with C₅-BODIPY(FL), SE (Molecular Probes) in 78% yield to give DBPC as the final product. This new synthesis generates a DBPC that has a diacylglycerol moiety with more appropriately matched distances between the glycerol backbone and the appended quencher and fluor.

Treatment of intermediate 4 with peanut PLD yielded its PA analogue, which was deprotected with TFA and then condensed with the active ester C₅-BODIPY(FL), SE to give DBPA. This revised route, using PLA₂ and PLD as synthetic reagents, provided

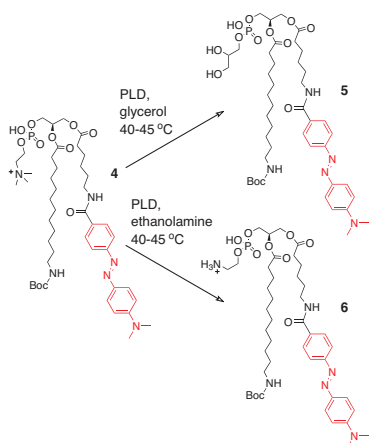


Figure 2. Enzymatic (PLD) reactions leading to DBPE and DBPG precursors. Reagents and conditions are as follows: aqueous nucleophile (glycerol, or ethanolamine) was added incrementally, until a reaction occurred, to Amberlite IRC-50 ion exchange resin and *Streptomyces* sp. PLD(P) (Genzyme) in CHCl_3 and stirred at 40–45 °C. See Methods for experimental details.

phosphatidylation of DBPC was attempted with PLD using the same conditions as above. Unfortunately, these conditions resulted in decomposition of the starting material, with no DBPE product detected. With this knowledge, the routes shown in Figure 3 were used to convert precursors **5** and **6** to final products DBPG and DBPE.

Using the described semienzymatic synthesis, we synthesized four fluorogenic PL analogues with different head groups from a common intermediate (**4**) in two to five steps. Chemical methods for head group introduction from modified diacylglycerol precursors gave unacceptably low yields. Furthermore, chemical introduction of head

groups before incorporation of acyl chains was rejected as inefficient, as it would require separate synthetic

groups before incorporation of acyl chains was rejected as inefficient, as it would require separate synthetic

Synthesis of DBPE and DBPG by

Transphosphatidylation. In the presence of excess alcohol, PLD can preferentially catalyze the alcoholysis of phosphatidylcholine in a process called transphosphatidylation (27–29). It was therefore envisaged that transphosphatidylation of intermediate **4** might yield PS, PE, and PG analogues, which could be processed further to give DBPS, DBPE, and DBPG.

In our hands, the use of anhydrous conditions defined earlier (28) for transphosphatidylation gave little or no product regardless of nucleophile. Addition of nucleophile in buffer, or of buffer alone, was required to drive the PLD reaction to completion. This procedure, a modification of earlier studies (27), was followed using L-serine, ethanolamine, and glycerol as nucleophiles, and DBPC precursor **4** as electrophile. The method gave DBPG and DBPE precursors, **5** and **6**, in good to excellent yields (Figure 2). *Streptomyces* sp. PLD(P) (Genzyme) was the only PLD that consistently catalyzed transphosphatidylation over hydrolysis. Other commercial PLD gave either no reaction or PA analogues. In the case of the L-serine reaction, the product distribution favored DBPA precursor upon buffer addition. Modifying the nucleophile to include variously protected forms of L-serine did not generate the desired DBPS.

To eliminate the need for protection and deprotection of the phosphatidylethanolamine moiety, direct trans-

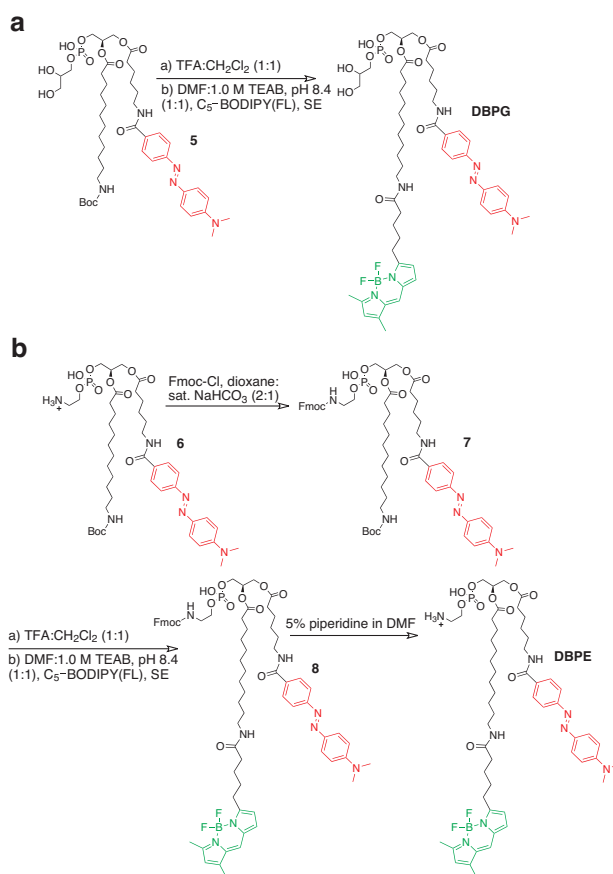


Figure 3. Synthesis of DBPG from **5 (a) and of DBPE from **6** (b). See Methods for experimental details.**

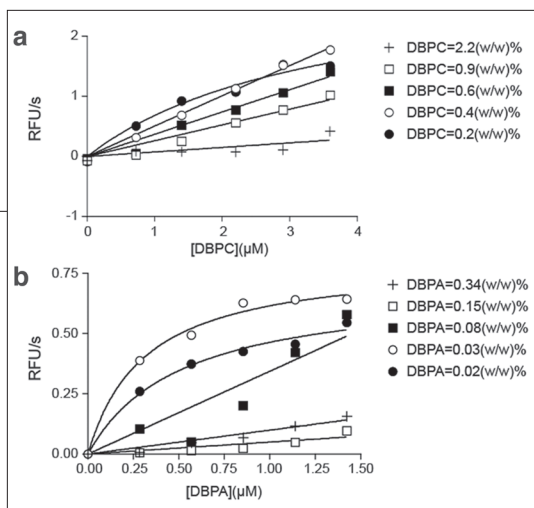


Figure 4. Concentration-dependent hydrolysis of DBPC (a) and of DBPA (b) by PLA₂. An increasingly hyperbolic curve is observed for plots of initial velocity (RFU (relative fluorescence unit)/s) versus probe concentration (μM) as the fraction of probe (either DBPC (a), or DBPA (b)) in Triton X-100 is decreased. LysoMaxS PLA₂ (0.3 U well^{-1}) and bee venom PLA₂ (0.5 U well^{-1}), respectively, were used with DBPC and DBPA in these assays.

fluorescence increases that were dependent on the concentrations of enzyme and probe (data not shown). These assays also revealed that predictable tracking of enzyme activity is predicated on the amount of detergent or phospholipid used as a carrier for the probe. As the fraction of probe in Triton X-100 micelles is decreased, the plots of initial velocity versus probe concentration become increasingly hyperbolic (Figure 4). At $\leq 0.2\%$ (w/w) of DBPC in Triton (0.3 U well^{-1} LysoMaxS PLA₂) and $\leq 0.03\%$ (w/w) of DBPA in Triton (0.5 U well^{-1} bee venom PLA₂), the data fit the Michaelis–Menten equation, and apparent V_{max} and K_{m} values can be determined at a given X_{d} .

Regardless of whether the probes were dispersed in Triton micelles (Figure 4) or phospholipid vesicles (data not shown), a window of robust PLA activity was observed. Above a certain mole fraction of DBPL, the reported enzyme activity stopped following Michaelis–Menten kinetics; at too low a mole fraction of DBPL, the fluorescent signal fell below detection limits. Mole fractions of probe beyond an upper limit presumably yield disrupted aggregation states that are resistant to enzyme catalysis, resulting in lower-than-expected signal generation.

Inhibitor Assay with DBPC.

The Triton mixed micelle assay, despite its intrinsic limitations for obtaining primary kinetic parameters (30), is widely used to obtain relative kinetic information in inhibitor screens and is the basis for a well-developed chromogenic

routes for each probe. Transphosphatidylating by PLD provided a mild, effective way to rapidly introduce head group diversity in our phospholipid probe design.

In vitro Activity Assays with DBPC and DBPA. *In vitro* enzyme assays with DBPC and DBPA produced linear

sPLA₂ assay (31). Triton X-100 has the advantages of being commercially available and relatively inert and is frequently used for isolation of membrane-associated proteins. Triton/DBPC micelles provided a suitable matrix to allow an $X_{\text{d}}(50)$ for thioether amide–PC (32) inhibition of PLA₂ to be quantified (Figure 5).

Increasing concentrations of thioether amide–PC were sonicated into Triton/DBPC mixed micelles and assayed in duplicate with bee venom PLA₂ (Figure 5). The resulting $X_{\text{d}}(50)$ was calculated to be 0.004, corresponding to a thioether amide–PC concentration of $2 \mu\text{M}$. This result correlates with that of a previous measurement of IC_{50} , also $2 \mu\text{M}$, for thioether amide–PC inhibition of cobra venom PLA₂ (32).

Head Group Selectivity Assays. The completed fluorogenic probes DBPA, DBPC, DBPE, and DBPG were used to experimentally determine the head group selectivities of a sampling of commercial PLA₂ in Triton mixed micelles. Each fluorogenic substrate was assayed with LysoMaxS, bee venom, cobra venom, bovine pancreas, *Streptomyces violaceoruber*, and Human Type V PLA₂ in TritonX-100 (reduced) micelles.

Selectivities expressed as a percentage of the slope of the analogue showing the most activity (Figure 6) revealed several interesting trends. First, mammalian enzymes (bovine and human) preferred the PG head group, followed by PC > PE >> PA. Second, the venom and bacterial enzymes preferred the PC head group, followed by PG > PE >> PA. Third, only the venom and pancreatic enzymes significantly catalyzed DBPA hydrolysis. Previous reports on the head group selectivity of these or other closely related PLA₂, as determined using a variety of assay methods, are summarized in Table 1. The degree of agreement between the listed head group selectivity studies is noteworthy, considering the contrasting methods used. Since the PLA head group preferences are conserved from assay platform to assay platform *in vitro*, the same preferences might also legitimately extrapolate to living

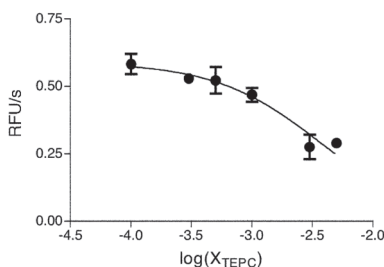


Figure 5. Inhibition of bee PLA₂ by thioether amide–PC. The log of increasing mole fractions of thioether amide–PC sonicated with $0.5 \mu\text{M}$ DBPC at $X_{\text{d}} = 0.001/0.2\%$ (w/w) in Triton X-100 micelles is plotted versus initial velocities ($n = 2$) generated upon addition of 0.01 U well^{-1} of bee venom PLA₂. Calculated $X_{\text{d}}(50) = 0.004$, corresponding to $2 \mu\text{M}$ thioether amide–PC.

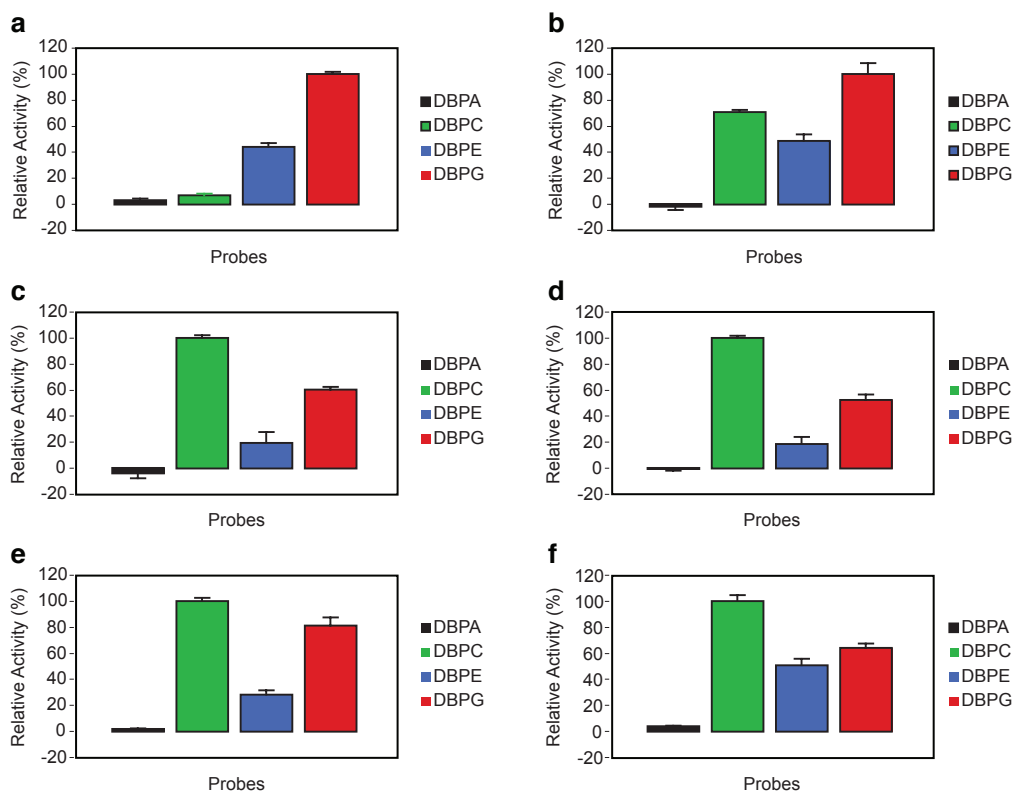


Figure 6. Head group selectivities of commercially available PLA₂. a) 0.1 U well⁻¹ of bovine pancreas, b) 20 ng well⁻¹ of human Type V, c) 0.2 U well⁻¹ of *S. violaceoruber*, d) 0.6 U well⁻¹ of LysoMaxS, e) 0.5 U well⁻¹ of *Naja mossaibica* venom, and f) 0.1 U well⁻¹ of bee venom PLA₂ were assayed against DBPA, DBPC, DBPE, and DBPG in Triton X-100 (reduced) mixed micelles ($X_d = 0.0001/0.02\%$ (w/w) for all probes). Head group selectivities are reported as percentages of the largest slope obtained over a 3 min incubation of enzyme with DBPL mixed micelles.

systems. Thus, these data validate the applicability of the much more easily utilized DBPL fluorogenic assay for further biological studies and inhibitor discovery.

Probe Design Features. Our probe design provides the benefit of unmodified phospholipid head groups and ester linkages that should reflect authentic protein-ligand interactions. Binding models (33, 34) predict shallow insertion by sPLA₂ into the membrane and place the PL acyl chains outside the phospholipid-binding pocket beyond approximately the ninth carbon of the *sn*-2 chain and the fourth carbon of the *sn*-1 chain (35). This leads to the assumption that, for sPLA₂, once a probe is properly inserted into liposomes/micelles, minor modifications to the chain termini compatible with a hydrophobic lipid environment will not be an important factor in enzyme selectivity. The correlation between our head group selectivity data and those of assays using different methods (Table 1) seems to support this assumption. However, the aromatic fluor and quencher groups at the chain termini of the DBPLs may disrupt the micelles sufficiently to result in attenuation of enzyme activity at higher mole fractions (36, 37) and could

also affect subcellular localization of the probes (38). Additionally, the saturated acyl linkages connecting the glyceryl backbone to the Dabcyl and BODIPY groups are expected to resist cleavage by cPLA₂, which insert more deeply into lipid bilayers and thereby maintain a preference for *sn*-2 arachidonic acid (34, 39).

The DBPL probes have also been designed so that fluorescence is completely quenched until cleavage occurs. In this regard, they are like molecular beacons, often used to detect the hybridization of nucleic acids (40). Another paradigm for probe design uses a FRET pair in which the donor and acceptor both emit fluorescence. The acceptor quenches donor fluorescence while simultaneously emitting at another distinct wavelength. The advantage of a design that uses two fluorophores is that the degree of cleavage can be assessed by measuring emission ratios between the donor and acceptor. Such ratiometric probes have been used in single molecule detection studies (41) and as substrates for phosphodiesterases (42) and PLA₂ (23), among other examples. However, bis-fluorophore probes suffer from higher background, lower sensitivity, and a requirement for more complex data analy-

TABLE 1. Comparison of head group selectivity data

PLA ₂	Phospholipid selectivity	Assay system
<i>S. violaceoruber</i>	PC » PE » PA	Triton X-100/chromagenic ^a
<i>S. violaceoruber</i>	PC » PG » PE » PA	Triton X-100/DBPLs
Pig pancreas	PG » PC	Triton X-100 and other detergents/pH stat ^b
Pig pancreas	PG » PE » PC	Continuous fluorescence displacement ^c
Bovine pancreas	PG » PE » PC » PA	Triton X-100/DBPLs
Bee venom	PG » PC » PE	Continuous fluorescence displacement ^c
Bee venom	PE ≥ PC	Triton X-100/pH-stat ^d
Bee venom	PC » PG ≥ PE » PA	Triton X-100/DBPLs
<i>N. naja</i> venom	PC » PG » PE	Continuous fluorescence displacement ^c
<i>N. naja</i> venom	PC » PE	Triton X-100/pH-stat ^d
<i>N. mossambica</i> venom	PC » PG » PE » PA	Triton X-100/DBPLs
Human group V	PC » PG » PA » PE	Mixed POPL vesicles/LC-ESI ^e
Human group V	PG » PC » PE	Polymerized mixed liposomes/ Pyrene-labeled PL ^f
Human group V	PG » PC » PE » PA	Triton X-100/DBPLs

^aAdapted from ref 48. ^bAdapted from ref 49. ^cAdapted from ref 50. ^dAdapted from ref 51.

^eAdapted from ref 52. ^fAdapted from ref 53.

sis. Incomplete quenching and inadequate spectral overlap can result in a high fluorescence background. An extreme example of this occurred when DBPC was compared with BBPC (bis-BODIPY-PC, Invitrogen) and found to generate a vastly superior signal (21).

The nonratiometric data obtained with the DBPL substrates is qualitative in a cellular context: comparative fluorescence intensity is used to monitor subcellular enzyme activity. The straightforward synthetic route to the DBPL analogues would readily allow substitution of the BODIPY–dabcyl pair with a ratiometric fluorophore pair, should applications require such probes. But, notwithstanding the quantitative limitations of the “dequenching” paradigm, a probe design in which fluorescence is completely quenched prior to enzyme cleavage, and only one fluor (rather than two) is released following cleavage, offers an advantage in fluorescence or laser scanning confocal microscopy. When used with these visualization techniques, dequenching probes like the DBPL substrates offer the potential for achieving excellent sensitivity and detailed detection of enzyme activity at subcellular locations without the need for specialized equipment.

Potential Probe Applications. An increasing body of work indicates important cell signaling roles for cPLA₂, sPLA₂, and iPLA₂ that do not necessarily prefer

unsaturated acyl chains and for which the DBPL probes could serve as fluorogenic substrates. For example, serum LPA is biosynthesized primarily by PLD cleavage of LPLs that have been generated either directly by serum LCAT (lecithin–cholesterol acyl transferase) or by cellular sPLA₂ and PS–PLA₁ (2). Direct cleavage of PA by sPLA₂ or mPA–PLA₁ can also give serum LPA in a mechanism that may be important for local signaling events like wound healing and inflammation (2). LPA production is thought to be regulated in cells by initial PLD conversion to PA followed with cleavage by an sPLA₂ or PA–PLA₁ (4).

Further examples include a human Group III sPLA₂ with homology to bee venom sPLA₂ that has an apparent preference for PG. Its unique tissue distribution relative to other sPLA₂ suggests a specialized biological role (18). A lysosomal PLA₂ from macrophages (17) having a demonstrated preference for

PG has been implicated in the biosynthesis of LBPA (lysobisphosphatidic acid), an important component of vesicle structure (43). In addition, *tafazzin*, the gene responsible for Barth syndrome, has been shown to encode a transacylase responsible for remodeling of PG and cardiolipin (44). On the basis of sPLA₂ hydrolysis experiments, LPE or its *N*-acyl derivatives are proposed intermediates in the biosynthesis of *N*-acyl ethanolamines, endogenous ligands for cannabinoid and vanilloid receptors having anti-inflammatory activity (10).

Another example of a signaling sPLA₂ with a demonstrated head group preference, a PC-preferring lysosomal PLA₂, is selectively expressed in alveolar macrophages and may play a role in pulmonary surfactant catabolism (12). LPC generation by iPLA₂ can stimulate the attraction of phagocytes to apoptotic cells (13). And, Group X sPLA₂, which is uniquely located in peripheral and neuronal fibers, demonstrates neuritogenic activity dependent on LPC production (14). The literature contains descriptions of many other sPLA₂ and iPLA₂ with uncertain or unknown head group selectivities.

Furthermore, cPLA₂ enzymes hydrolyze a variety of saturated phospholipids and can be important signal transducers. A cPLA₂ of this kind is likely to play a role in transducing stimulation of muscarinic receptors to

modulation of calcium channel activity (45) in neurons. In a related example, which also demonstrates the use of DBPC in a cell-based enzyme assay, SCG (superior cervical ganglion) neurons from mice expressing cPLA₂ and deficient in sPLA₂ were labeled with DBPC and found to exhibit a distinct increase in gross PLA₂ activity following stimulation with the muscarinic agonist Oxo-M (oxotremorine-M, Figure 7).

In conclusion, DBPA, DBPC, DBPE, and DBPG are useful new substrates for analysis of PLA activity *in vitro* and in living cells and tissues. The fluorogenic assay allows expeditious detection of enzyme activity in real time with a continuous readout. The assay is amenable to high-throughput screening for PLA₂ isozyme inhibitors and for examination of the PL head group selectivity of virtually all PLA₂ isozymes. The utility of DBPC in cell-based enzyme assays has been shown here as well as previously (21). Reports of its use in whole cell monitoring of PLA in studies investigating a putative PLA₂ involved in ovarian cancer progression (46), and for another PLA₂ involved in neuronal L-channel inhibition (Liu, L., Zhao, R., Bai, Y., Stanish, L. F., Evans, J. E.,

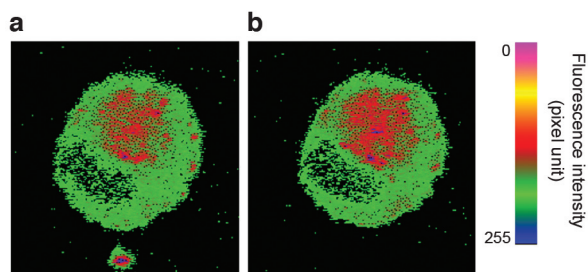


Figure 7. Fluorescence of DBPC in a cell-based assay of stimulation of an adult mouse superior cervical ganglion neuron expressing cPLA₂ but deficient in sPLA₂. PLA₂ activation is shown at 0 min (a, baseline PLA₂ activity), and after 6 min exposure to 10 μM Oxo-M (b). Images provided by Ann R. Rittenhouse, Rubing Zhao, Yen Bai, and Michael J. Sanderson; used with permission.

Sanderson, M. J., Bonventre, J. V., and Rittenhouse, A. R., unpublished results) will be described in detail in other publications. Further collaborative demonstrations of the utility of DBPLs bearing other head groups as substrates for a variety of animal, plant, and bacterial PLA enzymes *in vitro* and in cells are in progress and will be reported in due course.

METHODS

NMR and MS data can be found online in the Supporting Information.

6-(*p*-Methyl Red)amino hexanoic Acid (1). The sodium salt of *p*-methyl red (Acros) was converted to its acid form by treatment with 2 M HCl, followed by lyophilization. The acid form (6.8 g, 25 mmol) was then combined with NHS (*N*-hydroxysuccinimide, 4.4 g, 38 mmol) and EDCl (*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride, 7.3 g, 38 mmol) in DMF (60 mL). After stirring overnight, the reaction was concentrated under vacuum and washed (H₂O) to give the NHS-ester of *p*-methyl red. The NHS-ester (226 mg, 0.62 mmol) was further reacted with 6-aminocaproic acid (123 mg, 0.94 mmol) in a 2:1 solution of DMF/H₂O containing 7% DMAP. After 24 h, the reaction was acidified with 3 M HCl, then extracted with 10% MeOH in CH₂Cl₂. Solvents were removed *in vacuo*, and the product was precipitated from EtOH/H₂O as a deep red solid (200 mg, 77% yield, two steps).

1-*O*-(6-(*p*-Methyl Red)amino hexanoyl)-2-*O*-(6-(*p*-methyl red)amino hexanoyl)-*sn*-glyceryl Phosphatidylcholine (2). A solution of 6-(*p*-methyl red)amino hexanoic acid (40 mg, 0.11 mmol) and carbonyldiimidazole (25 mg, 0.16 mmol) in anhydrous DMSO (0.5 mL) was stirred under Ar for 30 min. This thick reaction mixture was transferred via syringe in 1 mL DMSO to a flask containing 4.5 mg (0.02 mmol) *sn*-glycero-3-phosphocholine (Bachem) and DBU (21 μL, 0.14 mmol) stirring in DMSO (0.25 mL) at 40 °C. After 24 h, the reaction was concentrated under vacuum and purified on SiO₂ (Flash Chromatography ASTM 230–400 Silica Gel) by elution with 65:25:4 CH₂Cl₂/MeOH/H₂O to give a red solid (7 mg, 40% yield).

1-*O*-(6-(*p*-Methyl Red)amino hexanoyl)-*sn*-glyceryl Phosphatidylcholine (3). Compound 2 (100 mg) was dissolved in MeOH (300 μL) and then treated with 1 mg (ca. 1400 U) PLA₂

(*Naja mossaibica mossaibica*, Sigma) in 600 μL of 0.1 M sodium borate buffer containing 0.1 M CaCl₂ and adjusted to pH 7.8. This solution was allowed to stir at 40 °C. Over 48 h, another 1 mg of PLA₂ was added in a total of 1.2 mL of 1:2 MeOH/buffer. After 72 h, the solvents were removed under a stream of Ar and compound 3 (56 mg, 90% yield) was collected as a red solid from a small SiO₂ column with 60:35:7 CH₂Cl₂/MeOH/H₂O.

1-*O*-(6-(*p*-Methyl Red)amino hexanoyl)-2-*O*-(12-Boc-amino dodecanoyl)-*sn*-glyceryl Phosphatidylcholine (4). To 65 mg (0.21 mmol) of 12-(Boc-amino)dodecanoic acid (Fluka) and TPSNT (23, 47) (76 mg, 0.20 mmol) stirring in distilled CH₂Cl₂ (2 mL) was added 30 μL (0.38 mmol) of 1-methylimidazole. The solution was stirred for 30 min and then added to a stirred solution of 3 (40 mg, 0.06 mmol) in 2 mL of CH₂Cl₂. After 72 h, the solvent was evaporated and product 4 (50 mg, 85% yield), a red solid, was isolated from SiO₂ with a step gradient of 1:9 MeOH/CH₂Cl₂; 65:25:4 CH₂Cl₂/MeOH/H₂O; then, 60:35:7 CH₂Cl₂/MeOH/H₂O.

1-*O*-(6-(*p*-Methyl Red)amino hexanoyl)-2-*O*-(12-(5-BODIPY-pentanoyl)aminododecanoyl)-*sn*-glyceryl Phosphatidylcholine (DBPC). Compound 4 (10 mg, 0.01 mmol) was dissolved in a solution of 1:2 TFA/CHCl₃ (1.5 mL, total). After stirring for 1 h, the solvent was removed *in vacuo* and the crude amine was used without purification in the next step. Freshly made 1 M TEAB buffer (600 μL, pH 8.5) was added to the crude amine (8.9 mg, 0.01 mmol), followed by a solution of 5 mg (0.01 mmol) of C₅-BODIPY(FL), SE (Molecular Probes) in 400 μL of DMF. The reaction was stirred for 24 h, at which time the solvents were removed with a stream of Ar. DBPC (9.5 mg, 78% yield) was isolated from a short SiO₂ column as a red solid, using a step gradient of 30% MeOH in CH₂Cl₂; then, 65:25:4 CH₂Cl₂/MeOH/H₂O.

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-(5-BODIPY-pentanoyl)-aminododecanoyl)-sn-glyceryl Phosphatidic Acid (DBPA). To a solution of 2 mg (0.002 mmol) of **4** in 300 μL of CHCl_3 was added 140 U of peanut PLD (Sigma) dissolved in 700 μL of 0.1 M sodium acetate buffer (pH 5.6) that contained 0.1 M CaCl_2 . The reaction was stirred vigorously at 35–40 °C for 2 days, at which time the solvents were removed by vacuum. The crude PLD product was purified as a red solid from SiO_2 with 20% MeOH in CH_2Cl_2 . The PLD product (ca. 4 mg) was dissolved in a 1:5 solution of TFA/ CH_2Cl_2 (1.2 mL, total). After stirring for 2 h, the solvents were removed, first under a stream of nitrogen and then *in vacuo*. The product (2 mg, 0.003 mmol) was then dissolved in 1 M TEAB buffer (600 μL), and 1.3 mg (0.003 mmol) of C_5 -BODIPY(FL), SE was added in 400 μL of DMF. The reaction was stirred for 6 h, and solvents were removed *in vacuo*. The product was isolated from SiO_2 with a step gradient elution of 1:9 MeOH/ CH_2Cl_2 ; then, 65:25:4 CH_2Cl_2 /MeOH/ H_2O + 1% TEA. DBPA was further purified by passage through Sephadex LH-20 in 1:9 MeOH/ CH_2Cl_2 . It was then passed through DOWEX 50WX8-200 (sodium form) to give DBPA as a red, solid sodium salt (1.6 mg, 25% yield, 3 steps).

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-Boc-amino-dodecanoyl)-sn-glyceryl Phosphatidylglycerol (5). Compound **4** (5 mg, 5.4 μmol) was dissolved in 1.3 mL of CHCl_3 . An excess (~250 μL , total) of glycerol, 97 mg of Amberlite IRC-50 beads, and 1000 U of *Streptomyces* sp. PLD(P) (Genzyme) were then added. After stirring for several days at 40–45 °C, no reaction was observed. Approximately 250 μL of buffer (0.2 M NaOAc and 0.08 M CaCl_2 , pH 5.6) and 25 U of additional PLD were then added. After stirring overnight, the reaction had proceeded almost to completion. The DBPG precursor **5** was purified on SiO_2 using 25% MeOH/ CH_2Cl_2 to give a red solid.

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-(5-BODIPY-pentanoyl)-aminododecanoyl)-sn-glyceryl Phosphatidylglycerol (DBPG). Compound **5** (2 mg, 2.2 μmol) was dissolved in a 50% solution of TFA in CH_2Cl_2 . After stirring for 2.5 h at room temperature, the solvents were removed under a stream of Ar. After 12 h under vacuum, the residue was dissolved in DMF (200 μL) containing C_5 -BODIPY(FL), SE (1.2 mg, 2.9 μmol). To this solution was added 400 μL of 1 M TEAB buffer (pH 8.4), and the reaction was stirred for 4 h. Solvents were removed with a stream of argon and then *in vacuo*. DBPG (1.7 mg, 68% yield, two steps), a red solid, was isolated from SiO_2 using 10% MeOH/ CH_2Cl_2 ; 20% MeOH/ CH_2Cl_2 ; then, 70:20:2 CH_2Cl_2 /MeOH/ H_2O .

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-Boc-amino-dodecanoyl)-sn-glyceryl Phosphatidylethanolamine (6). Amberlite IRC-50 beads (50 mg) and CHCl_3 (500 μL) were added to a vial containing dried **4** (6 mg, 6.5 μmol). Ethanolamine (10 μL , 164 μmol), then 200 U of *Streptomyces* sp. PLD(P) (Genzyme) in 400 μL of buffer (0.2 M NaOAc and 0.08 M CaCl_2 , pH 5.6) was added. The reaction was stirred at 40–45 °C. After 24 h, another 200 U of PLD was added, and then 200 U were added again after another 24 h. After stirring overnight, the solvents were stripped with a stream of nitrogen and DBPE precursor **6** was collected as a red solid from a SiO_2 column with 65:25:4 CH_2Cl_2 /MeOH/ H_2O .

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-Boc-amino-dodecanoyl)-sn-glyceryl Phosphatidylethanolamino-Fmoc Carbamate (7). To compound **6** (1.6 mg, 1.8 μmol) was added Fmoc-Cl (0.6 mg, 2.4 μmol) in 0.9 mL of dioxane, followed by 0.5 mL of a saturated solution of NaHCO_3 . After stirring for 24 h, the reaction was diluted with EtOAc and washed with water and brine. The solvents were removed, and the red, solid product **7** (1.5 mg, 78% yield) was purified on SiO_2 with a step gradient of 20% MeOH/ CH_2Cl_2 , then 65:25:4 CH_2Cl_2 /MeOH/ H_2O .

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-(5-BODIPY-pentanoyl)-aminododecanoyl)-sn-glyceryl

Phosphatidylethanolamino-Fmoc Carbamate (8). Compound **7** was dissolved in 1 mL of CH_2Cl_2 and stirred with 0.2 mL of TFA for 1 h at room temperature. Following removal of solvents by vacuum, the crude amine (3.1 mg, 3.1 μmol) was dissolved in 0.7 mL of 1 M TEAB buffer, followed by addition of 0.5 mL of DMF, then C_5 -BODIPY(FL), SE (1.4 mg, 3.4 μmol). The solution was concentrated by about half, after several minutes, with a stream of nitrogen, and another 0.5 mL of DMF was added to improve solubility. After stirring overnight, the reaction was stripped of solvent with nitrogen, and DBPE precursor **8** was purified on SiO_2 using a step gradient of 10% MeOH/ CH_2Cl_2 , 20% MeOH/ CH_2Cl_2 , and 65:25:4 CH_2Cl_2 /MeOH/ H_2O to give a red solid.

1-O-(6-(p-Methyl Red)amino)hexanoyl)-2-O-(12-(5-BODIPY-pentanoyl)-aminododecanoyl)-sn-glyceryl Phosphatidylethanolamine (DBPE). A 5% solution of piperidine in DMF was added to a vial containing a dried film of **8** (3.3 mg, 2.5 μmol). The solution was allowed to stir for 15 min and then evaporated under a stream of nitrogen. The product, DBPE, was collected from a short SiO_2 column as a red solid using 10% MeOH/ CH_2Cl_2 , then 65:25:4 CH_2Cl_2 /MeOH/ H_2O .

Fluorogenic Assays Using DBPC and DBPA with PLA₂. DBPC or DBPA was combined with either PC or Triton X-100 (reduced) at the relative concentrations given in the text. Carrier solvents (usually chloroform, or 65:25:4 chloroform/methanol/water) were evaporated under a stream of nitrogen and then *in vacuo* before buffer (0.1 M Tris, 0.1 M CaCl_2 , and 0.05% NaN_3 , pH 8.9) was added. The resulting suspension was sonicated for 30 s on a probe sonicator set at output = 3 and duty cycle = 50%, and then for an additional 2 min at the same sonicator settings with ice-bath cooling. Aliquots (0–100 μL) of the resulting micelle/vesicle solution were diluted to 100 μL total volume with buffer in a 96-well plate. Enzyme in buffer was added to initiate hydrolysis of the probe following a separate thermal pre-equilibration (2 min) at 37 °C of the enzyme and micelle/vesicle solutions. Fluorescence readings were recorded (λ_{ex} = 500 nm; λ_{em} = 530 nm) every 15 s for 5 min on a SpectraMax GeminiXS fluorescent plate reader. Plots of initial velocity (RFU s^{-1}) versus probe concentration were plotted and fit to the Michaelis-Menten equation using GraphPad Prism (GraphPad Software, Inc.) and were shown at various mole fractions of DBPA and DBPC in Triton mixed micelles (Figure 4).

Thioether Amide-PC Inhibitor Assay. To a 2 μg dried film of DBPC was added 2.2 mL of 0.04% (w/v) Triton X-100 (reduced) that had been prepared in 0.1 M Tris buffer, pH 8.8, containing 0.1 M CaCl_2 and 0.05% NaN_3 . Using a probe sonicator set to power output = 3 and duty cycle = 50%, we sonicated this suspension for 30 s at room temperature, then in an ice bath for 2 min. The resulting mixed micelle stock solution was 0.8 μM in DBPC, amounting to 0.25% (w/w) of the Triton micelles. Thioether amide-PC (Cayman Chemical) was aliquoted into test tubes from 1 ng μL^{-1} and 10 ng μL^{-1} stock solutions to give 0, 15, 46, 77, 152, 463, and 772 ng of inhibitor in 100 μL total of 0.04% (w/v) Triton X-100 (reduced) buffer. A total of 250 μL of the DBPC mixed micelle stock solution was added to each test tube of inhibitor. Each test tube was then sonicated as above for 30 min at room temperature and placed on ice for 2 min. When an approximate MW of 631 for Triton X-100 (reduced) was assumed, this gave thioether amide-PC/DBPC/Triton mixed micelle solutions of mole fraction $X_{\text{TEPC}} = 0.0001$ – 0.005 and a final concentration of DBPC in Triton of 0.18% (w/w), or 0.6 μM relative to solution. From each test tube, 100 μL of solution was placed in a black 96-well plate in duplicate. Separate wells containing 0.0013 U μL^{-1} bee venom PLA₂ were also prepared. After the plate was incubated at 37 °C for 2 min, 10 μL of bee venom PLA₂ solution was added to each well of DBPC/thioether amide-PC solution. Readings were recorded at 15 s intervals on a fluorescent plate reader during 5 min (λ_{ex} = 500 nm;

$\lambda_{em} = 530$ nm). Initial velocities were calculated from the slopes generated within the first 75 s of enzyme reaction. The $\log(X_{TEPC})$ was plotted against the initial velocities and fit to a sigmoidal curve for one-site competitive binding using GraphPad Prism (Figure 5).

Head Group Assay. Samples of DBPA, DBPC, DBPE, and DBPG were dried into separate vials and carefully weighed five times to get average masses. Stock solutions were generated by diluting the samples with 65:25:4 CH₂Cl₂/MeOH/H₂O. On the basis of the calculated average concentration of each sample, 5 μ g aliquots of each stock solution were withdrawn. To increase precision and decrease variability between experiments, the UV absorbance at 510 nm of each 5 μ g aliquot was measured and compared to the absorbance at 510 nm of its mother solution. Aliquot concentrations were then adjusted as necessary to bring each absorbance to its expected value. The 5 μ g aliquots were dried down to a film. To each resulting film was added a 0.76% (w/v) solution of Triton X-100 (reduced) in Tris buffer (0.1 M, containing 0.1 M CaCl₂ and 0.05% NaN₃, pH 8.8). These were sonicated as above to generate 1.2 μ M solutions of each probe as mixed micelles in Triton (0.02% (w/w) probe). Each solution was added to the wells of a black 96-well plate in duplicate. The wells were incubated and inoculated with 0.1 U well⁻¹ (70 ng well⁻¹) bee venom (Sigma), 0.6 U well⁻¹ LysoMaxS (Genecot), 0.5 U well⁻¹ (260 ng well⁻¹) *N. mossambica* venom (Sigma), 0.1 U well⁻¹ (5 μ g well⁻¹) bovine pancreas (Sigma), 0.02 U well⁻¹ *S. violaceoruber* (Sigma), or 20 ng well⁻¹ human Type V PLA₂ (Cayman Chemical) at 37 °C, and the fluorescence was measured as above. Slopes were obtained from 0–150 s, averaged, and compared (Figure 6).

Cell Assay. Images (Figure 7) were recorded from enzymatically dissociated SCG neurons from cPLA₂ +/- mice. Dissociated neurons from 8–16 week-old SCG were plated at a density of one-half ganglion per one 35 mm poly-L-lysine-coated glass bottom petri dish (no. 1.5, MatTek Corp.) and incubated at 37 °C in a 5% CO₂ environment in MEM. A 1 μ g μ L⁻¹ stock solution of DBPC in chloroform was prepared, divided into 12 μ L samples, dried with nitrogen, and stored at –80 °C. Aliquots were mixed with 44 μ L of PS in chloroform and redried under nitrogen. The DBPC-PS mixture was rehydrated with 1 mL of PBS, sonicated for 15 min on ice to generate liposomes, and used within a few hours of preparation. Cells cultured in Matek dishes were preincubated in HBSS (Hank's balanced salt solution) for 30 min. The DBPC-labeled liposomes were then added to the cell culture, yielding a final DBPC concentration of 0.01 μ g μ L⁻¹. Cells were incubated for 40 min at 37 °C, washed three times with HBSS to remove any adherent liposomes, and then viewed on a custom-built, video-rate confocal microscope with a 40 \times objective lens. An excitation wavelength of 488 nm was used, and emission spectra were collected with long pass filters (OGSIS) at 515 nm. After recording time-zero images (Figure 7a), cells were stimulated with Oxo-M (oxotremorine-M) or treated with HBSS for unstimulated control images (not shown). For each time point, images were collected at room temperature at 30 frames/s for 1 s using Video Savant (IO industries, Inc.) and directly written to a PC. Each set of 30 images was averaged to create an image for time intervals ranging from 1 to 9 min and analyzed using NIH ImageJ133. The image generated after 6 min stimulation is shown in Figure 7b.

Acknowledgment: We gratefully acknowledge Rubing Zhao, Yen Bai, Michael J. Sanderson, and Ann R. Rittenhouse for the cell images and associated experimental details. We thank the Center for Cell Signaling, a member of the Utah Centers of Excellence Program (1997-2002) and the NIH (Grants HL070231 and NS29632) for financial support of this research.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Balsinde, J., Winstead, M. V., and Dennis, E. A. (2002) Phospholipase A(2) regulation of arachidonic acid mobilization, *FEBS Lett.* 531, 2–6.
2. Aoki, J., Taira, A., Takanezawa, Y., Kishi, Y., Hama, K., Kishimoto, T., Mizuno, K., Saku, K., Taguchi, R., Arai, H. (2002) Serum lysophosphatidic acid is produced through diverse phospholipase pathways, *J. Biol. Chem.* 277, 48737–48744.
3. Moolenaar, W., van Meeteren, L., and Giepmans, B. (2005) The ins and outs of lysophosphatidic acid signaling, *BioEssays* 26, 870–881.
4. Aoki, J. (2004) Mechanisms of lysophosphatidic acid production, *Semin. Cell Dev. Biol.* 15, 477–489.
5. Kalyvas, A., and David, S. (2004) Cytosolic phospholipase A2 plays a key role in the pathogenesis of multiple sclerosis-like disease, *Neuron* 41, 323–335.
6. Niessen, H. W., Krijnen, P. A., Visser, C. A., Meijer, C. J., and Erik Hack, C. (2003) Type II secretory phospholipase A2 in cardiovascular disease: A mediator in atherosclerosis and ischemic damage to cardiomyocytes?, *Cardiovasc. Res.* 60, 68–77.
7. Sun, G. Y., Xu, J., Jensen, M. D., and Simonyi, A. (2004) Phospholipase A2 in the central nervous system: Implications for neurodegenerative diseases, *J. Lipid Res.* 45, 205–213.
8. Laye, J. P., and Gill, J. H. (2003) Phospholipase A2 expression in tumours: A target for therapeutic intervention?, *Drug Discovery Today* 8, 710–716.
9. Hiramatsu, T., Sonoda, H., Takanezawa, Y., Morikawa, R., Ishida, M., Kasahara, K., Sanai, Y., Taguchi, R., Aoki, J., Arai, H. (2003) Biochemical and molecular characterization of two phosphatidic acid-selective phospholipase A1s, mPA-PLA₁alpha and mPA-PLA₁beta, *J. Biol. Chem.* 278, 49438–49447.
10. Sun, Y.-X., Tsuboi, K., Okamoto, Y., Tonai, T., Murakami, M., Kudo, I., Ueda, N. (2004) Biosynthesis of anandamide and *N*-palmitoylethanolamine by sequential actions of phospholipase A2 and lysophospholipase D, *Biochem. J.* 380, 749–756.
11. Badiani, K., and Arthur, G. (1995) Evidence for receptor and G-protein regulation of a phosphatidylethanolamine-hydrolysing phospholipase A1 in guinea-pig heart microsomes: Stimulation of phospholipase A1 activity by DL-isoprenaline and guanine nucleotides, *Biochem. J.* 312, 805–809.
12. Abe, A., Hiraoka, M., Wild, S., Wilcoxon, S. E., Paine, R., III, and Shayman, J. A. (2004) Lysosomal phospholipase A2 is selectively expressed in alveolar macrophages, *J. Biol. Chem.* 279, 42605–42611.
13. Lauber, K., Bohn, E., Krober, S. M., Xiao, Y. J., Blumenthal, S. G., Lindemann, R. K., Marini, P., Wiedig, C., Zobywalski, A., Baksh, S., Xu, Y., Autenrieth, I. B., Schulze-Osthoff, K., Belka, C., Stuhler, G., Wesselborg, S. (2003) Apoptotic cells induce migration of phagocytes via caspase-3-mediated release of a lipid attraction signal, *Cell* 113, 717–730.
14. Masuda, S., Murakami, M., Takanezawa, Y., Aoki, J., Arai, H., Ishikawa, Y., Ishii, T., Arioka, M., Kudo, I. (2005) Neuronal expression and neurotogenic action of group X secreted phospholipase A2, *J. Biol. Chem.* 280, 23203–23214.
15. Hosono, H., Aoki, J., Nagai, Y., Bando, K., Ishida, M., Taguchi, R., Arai, H., Inoue, K. (2001) Phosphatidylserine-specific phospholipase A1 stimulates histamine release from rat peritoneal mast cells through production of 2-acyl-1-lysophosphatidylserine, *J. Biol. Chem.* 276, 29664–29670.
16. Bellini, F., and Bruni, A. (1993) Role of a serum phospholipase A1 in the phosphatidylserine-induced T cell inhibition, *FEBS Lett.* 316, 1–4.
17. Shinozaki, K., and Waite, M. (1999) A novel phosphatidylglycerol-selective phospholipase A2 from macrophages, *Biochemistry* 38, 1669–1675.
18. Valentin, E., Ghomashchi, F., Gelb, M. H., Lazdunski, M., and Lambeau, G. (2000) Novel human secreted phospholipase A(2)

- with homology to the group III bee venom enzyme, *J. Biol. Chem.* **275**, 7492–7496.
19. Reynolds, L. J., Washburn, W. N., Deems, R. A., and Dennis, E. A. (1991) Assay strategies and methods for phospholipases, *Methods Enzymol.* **197**, 3–23.
20. Yu, L., and Dennis, E. A. (1991) Thio-based phospholipase assay, *Methods Enzymol.* **197**, 65–75.
21. Feng, L., Manabe, K., Shope, J. C., Widmer, S., DeWald, D. B., and Prestwich, G. D. (2002) A real-time fluorogenic phospholipase A(2) assay for biochemical and cellular activity measurements, *Chem. Biol.* **9**, 795–803.
22. Hendrickson, H. S., Hendrickson, E. K., Johnson, I. D., and Farber, S. A. (1999) Intramolecularly quenched BODIPY-labeled phospholipid analogs in phospholipase A(2) and platelet-activating factor acetylhydrolase assays and *in vivo* fluorescence imaging, *Anal. Biochem.* **276**, 27–35.
23. Wichmann, O., and Schultz, C. (2001) FRET probes to monitor phospholipase A2 activity, *Chem. Commun.*, 2500–2501.
24. Picq, M., Huang, Y., Lagarde, M., Doutheau, A., and Nemoz, G. (2002) Synthesis of photoreactive phosphatidic acid analogues displaying activatory properties on cyclic AMP-phosphodiesterases. Photoaffinity labeling of an isoform of phosphodiesterase, *J. Med. Chem.* **45**, 1678–1685.
25. Ishihara, K., Kubota, M., Kurihara, H., and Yamamoto, H. (1995) Scandium trifluoromethanesulfonate as an extremely active acylation catalyst, *J. Org. Chem.* **117**, 4413–4414.
26. Tamura, Y., Fukuda, W., and Tormoi, M. (1994) Polymer-supported bases. XII. Regioselective synthesis of lysophospholipids using polymer-supported bicyclic amidines or guanidines, *Synth. Commun.* **24**, 2907–2914.
27. Pisch, S., Bomscheuer, U. T., Meyer, H. H., and Schmid, R. D. (1997) Properties of unusual phospholipids IV: Chemoenzymatic synthesis of phospholipids bearing acetylenic fatty acids, *Tetrahedron* **53**, 14627–14634.
28. Rich, J. O., and Khmelnskiy, Y. L. (2001) Phospholipase D-catalyzed transphosphatidylation in anhydrous organic solvents, *Biotechnol. Bioeng.* **72**, 374–377.
29. Somerharju, P. J., Virtanen, J. A., Eklund, K. K., Vainio, P., and Kinnunen, P. K. (1985) 1-Palmitoyl-2-pyrenedecanoyl glycerophospholipids as membrane probes: Evidence for regular distribution in liquid-crystalline phosphatidylcholine bilayers, *Biochemistry* **24**, 2773–2781.
30. Berg, O. G., Gelb, M. H., Tsai, M. D., and Jain, M. K. (2001) Interfacial enzymology: The secreted phospholipase A(2)-paradigm, *Chem. Rev.* **101**, 2613–2654.
31. Reynolds, L. J., Hughes, L. L., and Dennis, E. A. (1992) Analysis of human synovial fluid phospholipase A2 on short chain phosphatidylcholine-mixed micelles: Development of a spectrophotometric assay suitable for a microtiterplate reader, *Anal. Biochem.* **204**, 190–197.
32. Yu, L., Deems, R. A., Hajdu, J., and Dennis, E. A. (1990) The interaction of phospholipase A2 with phospholipid analogues and inhibitors, *J. Biol. Chem.* **265**, 2657–2664.
33. Lin, Y., Nielsen, R., Murray, D., Hubbell, W. L., Mailer, C., Robinson, B. H., Gelb, M. H. (1998) Docking phospholipase A2 on membranes using electrostatic potential-modulated spin relaxation magnetic resonance, *Science* **279**, 1925–1929.
34. Lichtenbergova, L., Yoon, E. T., and Cho, W. (1998) Membrane penetration of cytosolic phospholipase A2 is necessary for its interfacial catalysis and arachidonate specificity, *Biochemistry* **37**, 14128–14136.
35. Scott, D. L., White, S. P., Otwinowski, Z., Yuan, W., Gelb, M. H., and Sigler, P. B. (1990) Interfacial catalysis: The mechanism of phospholipase A2, *Science* **250**, 1541–1546.
36. Barlow, P. N., Lister, M. D., Sigler, P. B., and Dennis, E. A. (1988) Probing the role of substrate conformation in phospholipase A2 action on aggregated phospholipids using constrained phosphatidylcholine analogues, *J. Biol. Chem.* **263**, 12954–12958.
37. Wells, M. A. (1974) The mechanism of interfacial activation of phospholipase A2, *Biochemistry* **13**, 2248–2257.
38. Kuerschner, L., Ejsing, C. S., Ekroos, K., Shevchenko, A., Anderson, K. I., and Thiele, C. (2005) Polyene-lipids: A new tool to image lipids, *Nat. Methods* **2**, 39–45.
39. Dessen, A., Tang, J., Schmidt, H., Stahl, M., Clark, J. D., Seehra, J., Somers, W. S. (1999) Crystal structure of human cytosolic phospholipase A2 reveals a novel topology and catalytic mechanism, *Cell* **97**, 349–360.
40. Tan, W., Wang, K., and Drake, T. J. (2004) Molecular beacons, *Curr. Opin. Chem. Biol.* **8**, 547–553.
41. Deniz, A. A., Laurence, T. A., Dahan, M., Chemla, D. S., Schultz, P. G., and Weiss, S. (2001) Ratiometric single-molecule studies of freely diffusing biomolecules, *Annu. Rev. Phys. Chem.* **52**, 233–253.
42. Takakusa, H., Kikuchi, K., Urano, Y., Sakamoto, S., Yamaguchi, K., and Nagano, T. (2002) Design and synthesis of an enzyme-cleavable sensor molecule for phosphodiesterase activity based on fluorescence resonance energy transfer, *J. Am. Chem. Soc.* **124**, 1653–1657.
43. Kobayashi, T., Beuchat, M. H., Lindsay, M., Frias, S., Palmiter, R. D., Sakuraba, H., Parton, R. G., Gruenberg, J. (1999) Late endosomal membranes rich in lysobisphosphatidic acid regulate cholesterol transport, *Nat. Cell Biol.* **1**, 113–118.
44. Vreken, P., Valianpour, F., Nijtmans, L. G., Grivell, L. A., Plecko, B., Wanders, R. J., Barth, P. G. (2000) Defective Remodeling of Cardiolipin and Phosphatidylglycerol in Barth Syndrome, *Biochem. Biophys. Res. Commun.* **279**, 378–382.
45. Liu, L., Roberts, M. L., and Rittenhouse, A. R. (2004) Phospholipid metabolism is required for M1 muscarinic inhibition of N-type calcium current in sympathetic neurons, *Eur. Biophys. J.* **33**, 255–264.
46. Ren, J., Xiao, Y. J., Singh, L. S., Zhao, X., Zhao, Z., Feng, L., Rose, T. M., Prestwich, G. D., Xu, Y. (2006) LPA is constitutively produced by human peritoneal mesothelial cells and enhances adhesion, migration and invasion of ovarian cancer cells, *Cancer Res.*, in press.
47. de Rooij, J. F. M., Wille-Hazeleger, G., van Deursen, P. H., Serdijn, J., and van Boom, J. H. (1979) Synthesis of complementary DNA fragments via phosphotriester intermediates, *Recl. Trav. Chim. Pays-Bas* **98**, 537–548.
48. Sugiyama, M., Ohtani, K., Izuhara, M., Koike, T., Suzuki, K., Imamura, S., Misaki, H. (2002) A novel prokaryotic phospholipase A2. Characterization, gene cloning, and solution structure, *J. Biol. Chem.* **277**, 20051–20058.
49. Volwerk, J. J., Jost, P. C., de Haas, G. H., and Griffith, O. H. (1986) Activation of porcine pancreatic phospholipase A2 by the presence of negative charges at the lipid-water interface, *Biochemistry* **25**, 1726–1733.
50. Kinkaid, A., and Wilton, D. C. (1991) Comparison of the catalytic properties of phospholipase A2 from pancreas and venom using a continuous fluorescence displacement assay, *Biochem. J.* **278**, 843–848.
51. Pluckthun, A., and Dennis, E. A. (1985) Activation, aggregation, and product inhibition of cobra venom phospholipase A2 and comparison with other phospholipases, *J. Biol. Chem.* **260**, 11099–11106.
52. Singer, A. G., Ghomashchi, F., Le Calvez, C., Bollinger, J., Bezzine, S., Rouault, M., Sadilek, M., Nguyen, E., Lazdunski, M., Lambeau, G., Gelb, M. H. (2002) Interfacial kinetic and binding properties of the complete set of human and mouse groups I, II, V, X, and XII secreted phospholipases A2, *J. Biol. Chem.* **277**, 48535–48549.
53. Kim, K. P., Han, S. K., Hong, M., and Cho, W. (2000) The molecular basis of phosphatidylcholine preference of human group-V phospholipase A2, *Biochem. J.* **348**, 643–647.

Structural Characterization of *in Vitro* and *in Vivo* Intermediates on the Loading Module of Microcystin Synthetase

Leslie M. Hicks^{†,¶}, Michelle C. Moffitt^{†,¶}, Laura L. Beer[‡], Bradley S. Moore^{‡,§,*}, and Neil L. Kelleher^{†,*}

[†]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, [‡]Departments of Pharmacology & Toxicology, University of Arizona, Tucson, Arizona 85721, and [§]Scripps Institution of Oceanography and Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093

Toxic cyanobacterial waterblooms are of serious concern due to their production of potent hepatotoxins and neurotoxins in freshwater lakes and water reservoirs. Of the known toxins, the hepatotoxic microcystins are produced in various cyanobacterial genera, including *Microcystis*, *Anabaena*, *Nostoc*, *Oscillatoria*, and *Planktothrix* (1–5). Previous biochemical and genetic studies found that these cyclic heptapeptides are assembled on a large multienzyme complex *via* a mixed nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS) pathway, the first such pathway identified in a cyanobacterium. NRPSs and PKSs are large, multifunctional protein complexes that catalyze the synthesis of nonribosomal peptides and polyketides, respectively, from amino acid and small chain carboxylic acid monomers *via* a thioester template mechanism (Figure 1) (6, 7). The microcystin synthetase (*mcy*) gene cluster was cloned from *Microcystis aeruginosa* (Figure 2) (8–10) and more recently from strains of the genera *Planktothrix* (11) and *Anabaena* (12). Through comparison of the sequence data from the *mcy* gene clusters with other NRPS and PKS systems, a biosynthetic scheme for microcystin was proposed (8–10). Microcystin synthetase contains an adenylation–peptidyl carrier protein (A–PCP) loading didomain at the N-terminus of the McyG PKS that has been postulated to activate and load the starter unit phenylacetate for subsequent polyketide extension in the formation of the novel aromatic β -amino acid (2S,3S,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoic acid (Adda) residue (Figure 2) (9, 10). While the structure of Adda suggests priming with phenylacetate, *in vivo* feeding experiments do not support the direct involvement of this starter unit (13, 14). Hence, on

ABSTRACT The microcystin family of toxins is the most common cause of hepatotoxicity associated with water blooms of cyanobacterial genera. The biosynthetic assembly line producing the toxic cyclic peptide, microcystin, contains an adenylation–peptidyl carrier protein didomain (A–PCP) at the N-terminus of the initiator module McyG (295 kDa) that has been postulated to activate and load the starter unit phenylacetate for formation of the unusual aromatic β -amino acid residue, Adda, before subsequent extension. Characterization of the McyG A–PCP didomain (78 kDa) using ATP–PP_i exchange assays and mass spectrometry revealed that assorted phenylpropanoids are preferentially activated and loaded onto the PCP carrier domain rather than phenylacetate itself. For the first time, thioesters formed *in vivo* were detected directly using large molecule mass spectrometry. Additionally substrates were cleaved using a type II thioesterase for structural elucidation by small molecule mass spectrometry. Unprecedented features of the McyG A–PCP didomain include the *in vivo* acylation of the *holo* PCP with exogenous and endogenous substrates, along with the ability of the *apo* protein to retain the acyl-AMP intermediate during affinity purification. These results imply that phenylpropanoids are preferentially loaded onto the McyG PCP; however one carbon must be excised following extension of the starter unit with malonyl-CoA in order to generate the expected polyketide chain which leads us to ponder the novel biochemistry by which this occurs.

*To whom correspondence should be addressed.

E-mail: kelleher@scs.uiuc.edu.

E-mail: bsmoore@ucsd.edu.

[¶]These authors contributed equally to this work.

Received for review November 18, 2005
and accepted March 2, 2006

Published online March 17, 2006

10.1021/cb500007v CCC: \$33.50

© 2006 by American Chemical Society

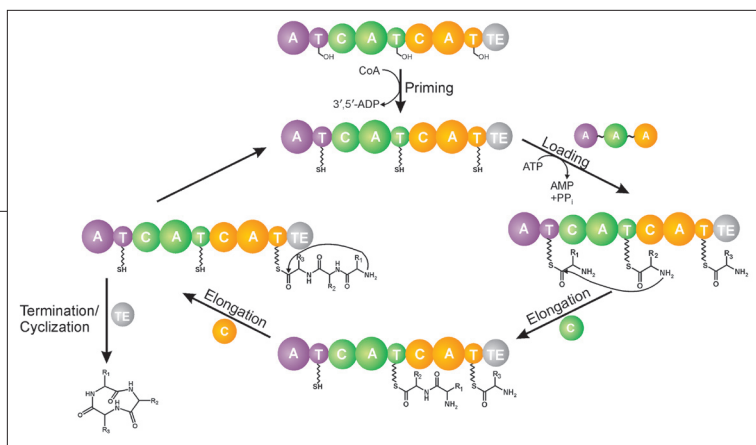


Figure 1. General mechanism of NRPS biosynthesis. Each module (denoted by color) contains the machinery to incorporate one monomer into the growing peptide chain. A minimal module contains an adenylation (A) domain, a condensation (C) domain, and a thiolation (T) domain (also known as peptidyl carrier protein or PCP) to which the intermediates are covalently tethered. The carrier protein domain contains an active site Ser to which a phosphopantetheine moiety is post-translationally loaded. This moiety then serves as the transporting group, moving the elongation intermediates between the catalytic centers. The A domains are responsible for recognizing and activating the requisite amino acid substrates and loading them onto the carrier protein domains. The C domains are responsible for catalyzing condensation, which forms the peptide bonds between the subunits. The final product composition depends on the number and arrangement of these modules, with the final product release step catalyzed by a thioesterase (TE) domain.

the basis of these incompatible *in vivo* data, we set out to investigate the substrate selectivity and specificity of the McyG didomain (McyG A-PCP).

Conventional methods to analyze the biosynthesis of NRPS/PKS complexes include quantification of individual intermediates present at stoichiometric levels by autoradiography and/or base hydrolysis of the thioester intermediates and analysis by radio-TLC (or high-performance liquid chromatography (HPLC)) (15–18). However, breakthroughs in electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) (19, 20) catalyzed a period of technology development that established a variety of instrumental modes for mass spectrometry (MS) to report on the molecular structure of small molecules, peptides, and intact proteins. Currently, the direct interrogation of enzyme intermediates by MS is illuminating a diverse array of chemistries used in nature (21). The development of hybrid quadrupole Fourier-transform mass spectrometry (Q-FTMS) instruments has enabled the detection of low-level components from complex mixtures and verification of their identity by high-resolution tandem MS (MS/MS) (22). The application of ESI coupled with Q-FTMS for the interrogation of thioester bound intermediates of the bacitracin, yersiniabactin, epothilone, and aminocoumarin pathways has further improved fundamental understanding of the mechanistic enzymology of these complex pathways (23–27). Probing of multiple modules in parallel has been achieved, along with substrate–activity relationships and quantification

of active site occupancy. Additionally, proteolysis and liquid chromatography on line with ion-trap MS applied to natural and engineered proteins in the prototypical PKS system, 6-deoxyerythronolide B synthase (DEBS), has recently been reported (28, 29).

By use of diverse types of measurement strategies and instruments, a growing number of labs are utilizing the potential of MS to directly observe enzyme-bound intermediates. In this study, we report the *in vitro* and *in vivo* processing of the microcystin loading module (McyG A-PCP) with assorted primer units. Using conventional ATP-PP_i exchange assays and a combination of MS technologies, including MALDI time-of-flight (TOF) MS, ESI-Q-FTMS, and small molecule structural elucidation by gas chromatography/electron impact MS (GC/EI-MS), novel NRPS biochemistry is revealed.

RESULTS AND DISCUSSION

While sequence analysis of the *mcy* gene cluster has identified the mixed NRPS/PKS assembly complex proposed for Adda biosynthesis (9, 10), the mechanism of starter unit incorporation remains elusive. Phenylacetate was suggested to be the starter unit for Adda biosynthesis since it correlates with the structures of these related cyclic peptides and recent analysis of their respective biosynthetic gene clusters did not suggest otherwise. Previous feeding experiments, however, did not directly support the incorporation of phenylacetate and instead indicated that the starter unit was derived from L-phenylalanine (13, 14). Because of these intriguing observations, we set out to determine the priming mechanism of microcystin synthetase to deduce whether starter unit modifications take place before or after tethering to the assembly line.

The McyG A-PCP Loading Domain. Sequence analysis revealed that the *M. aeruginosa* PCC7820 *mcyG* gene was 99% identical to those from the other *M. aeruginosa* strains. The deduced protein sequence was identical in its arrangement of catalytic domains (Figure 2, panel c) with each containing the conserved active site residues required for activity. The amino acid sequence of the McyG A domain was aligned with the sequence of the GrsA A domain, for which the crystal structure has been solved; however, the presence of additional amino acids around the active site of the McyG A domain which were absent in the GrsA sequence made it difficult to postulate the substrate specificity based on sequence alignments alone

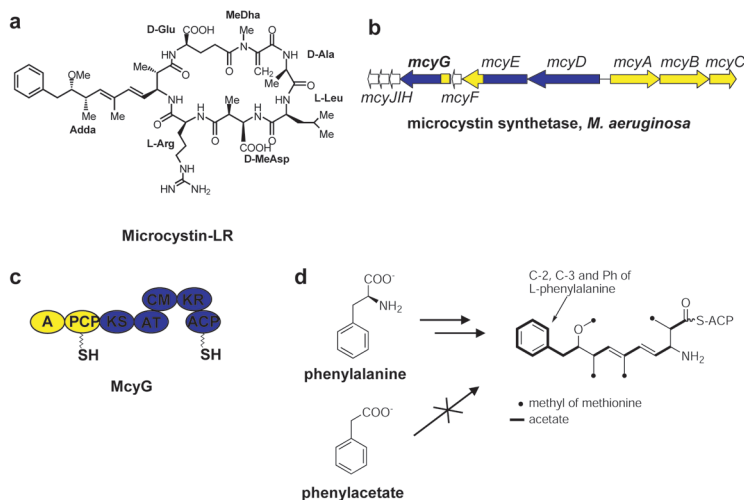


Figure 2. Overview of microcystin-LR structure, biosynthetic genes and proteins, and biosynthesis. **a)** Structure of microcystin-LR. Uncommon amino acid abbreviations: Adda (3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoic acid), MeDha (*N*-methyldehydroalanine), MeDhb (*N*-methyldehydrobutyrine), and *D*-MeAsp (*D*-erythro- β -methylaspartic acid). **b)** The microcystin (*mcy*) biosynthetic gene cluster from *M. aeruginosa*. The cluster contains 10 bidirectionally transcribed open reading frames comprised of six large multienzyme synthases/synthetases (McyA-E, G) that incorporate the amino acid and acyl-CoA precursors and four monofunctional proteins that are putatively involved in *O*-methylation (McyI), epimerization (McyF), dehydration (McyI), and localization (McyH). **c)** The hybrid NRPS (yellow) and PKS proteins (blue) encoded by *mcyG*. The McyG A-PCP didomain (yellow) is the focus of this study. Abbreviations: A (adenylation), PCP (peptidyl carrier protein), KS (ketosynthase), AT (acyltransferase), CM (C-methyltransferase), KR (ketoreductase), and ACP (acyl carrier protein). **d)** Origin of the carbons in the Adda subunit of microcystin.

(30, 31). The *apo* McyG A-PCP didomain (78 kDa) was successfully expressed as described in the Methods section and is shown in Supplementary Figure 1.

Characterization of Adenylation Domain Specificity.

To investigate the substrate specificity of the McyG didomain that is involved in the proposed loading of phenylacetate during the assembly of the Adda unit in the microcystin peptide, a range of substrates were tested against the *apo* McyG A-PCP didomain using the ATP-PP_i exchange assay (32). This enzyme accepted a wide range of phenylpropanoids (Figure 3), and the highest activity was clearly observed with *trans*-cinnamic acid. Other phenylpropanoids were utilized as possible substrates, including *D*-phenylalanine, *p*-coumarate (activities 30% and 17% relative to cinnamic acid), and to a lesser extent *L*-phenylalanine, phenylpyruvate, *D*-3-phenyllactate, and *L*-3-phenyllactate (less than 10% relative to cinnamic acid). Activity was less than 1% for the substrate phenylacetate along with benzoate, *L*-alanine, and water as negative controls.

In Vivo and *In Vitro* Formation of Holo McyG A-PCP.

In order to analyze substrate loading onto the McyG PCP, it was essential to generate the active *holo* form of the McyG PCP by transferring the phosphopantetheine moiety of CoA onto Ser 604. Since the *E. coli* strain used for heterologous expression of the McyG A-PCP protein does not contain a general phosphopantetheinyl transferase (PPTase)

that is suitable for post-translational modification of foreign PCPs (33), a bifunctional pHIS₈-based expression plasmid was constructed to generate the *holo* PCP *in vivo* by coexpressing the octahistidyl-tagged McyG A-PCP didomain with the Svp PPTase protein from *Streptomyces verticillus* ATCC15003 (34). Absence of the poly-histidine tag attached to the Svp protein enabled direct purification of the tagged *holo* McyG A-PCP didomain (*holo* A-PCP_{*in vivo*}) by

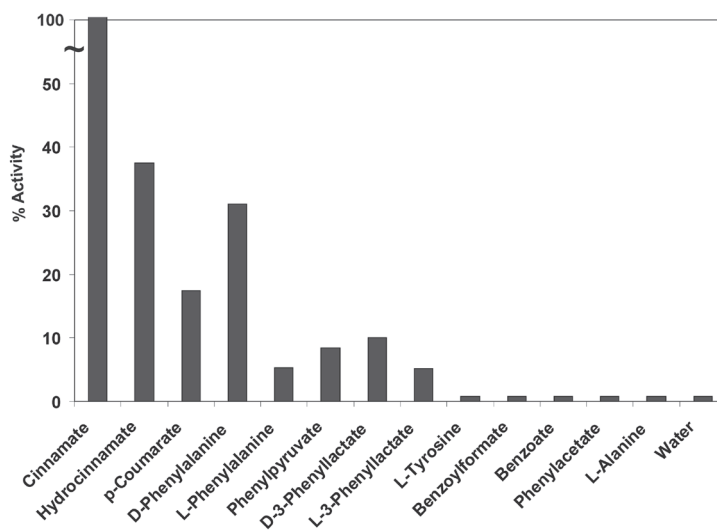


Figure 3. The McyG adenylation domain and its relative substrate-dependent ATP-PP_i exchange activities for a variety of substrates. Activity of the McyG didomain relative to cinnamic acid (100%).

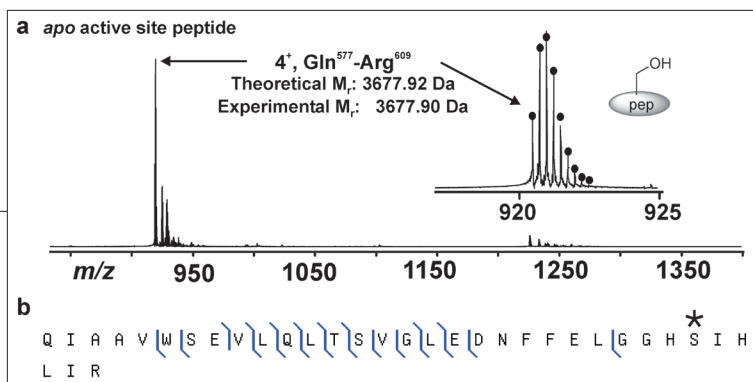


Figure 4. Identification of the tryptic peptide containing the McyG-PCP active site serine. a) Fourier-transform mass spectrum of the Q⁵⁷⁷-R⁶⁰⁹ tryptic peptide containing the *apo* McyG-PCP active site serine (Ser⁶⁰⁴). b) The fragment ion map correlates the observed MS/MS fragment ions (not shown) to the Q⁵⁷⁷-R⁶⁰⁹ peptide; the active site serine at position 604 is marked with an asterisk.

Ni²⁺-affinity chromatography. Alternatively, purified PPTases, Svp (34) or Sfp from *Bacillus subtilis* (35), were used to catalyze the post-translational modification of the *apo*-PCP *in vitro* to form *holo* A-PCP_{*in vitro*}.

Analysis of Intact McyG A-PCP by MALDI-TOF MS.

Phosphopantetheinylation of the McyG PCP *in vivo* and *in vitro* was confirmed by MALDI-TOF MS analysis. The *apo* A-PCP protein had the expected mass; however, the *holo* A-PCP_{*in vitro*} protein was found to be 511 Da greater than the mass of the *apo* protein, far greater than the anticipated 340 Da increase that would result from the attachment of the phosphopantetheine cofactor from coenzyme A. MALDI-TOF MS analysis of the *holo* A-PCP_{*in vitro*} protein revealed the *holo* enzyme (78604 Da) along with this larger species (78751 Da) previously observed in the *in vivo* generated A-PCP sample (Supplementary Table 1). These data suggested that the overweight *holo* protein may reflect a common acylated form that was modified not only *in vivo* but also *in vitro* through the copurification of an acyl-AMP bound to the enzyme that when converted from *apo* to *holo* autoacylates. In order to characterize this enzyme-bound intermediate at isotopic resolution, proteolysis combined with ESI-Q-FTMS was utilized.

Q-FTMS Characterization of McyG A-PCP.

HPLC separation of peptides generated by exhaustive trypsin digestion of the *apo* McyG A-PCP didomain followed by Q-FTMS analysis was performed in order to characterize the PCP active site (Figure 4). Initial experiments detected a peptide corresponding to Gln⁵⁷⁷-Arg⁶⁰⁹ in the *apo* didomain (Figure 4, panel a). The identity of this PCP active site containing peptide was confirmed by tandem mass spectrometry (data not shown), the fragment ions from which are correlated to the peptide primary sequence as seen in the graphical fragment map (Figure 4, panel b). Analysis of all HPLC fractions indicated that the peptide was present only in its *apo*-form, revealing that the *E. coli* PPTases (EntD or ACPS) had no detectable activity during McyG A-PCP overexpression.

Acylation of *in Vitro*-Generated *Holo* McyG A-PCP

Didomain. Purified *apo* A-PCP protein was incubated with Sfp and CoA in order to generate the *holo* A-PCP_{*in vitro*} protein and subsequently used to analyze loading of substrates *in vitro* (Figure 5). In the absence of substrate, a small amount of *holo* peptide was identified; however, the majority was found to have a M_r value 132.0 Da greater than the *holo* form (Figure 5, panel a). This observation is consistent with the MALDI-TOF MS results, which identified that a mixture of *holo* and acylated protein was generated from *in vitro* phosphopantetheinylation. On the basis of the +132.0 Da mass shift (Δm) and the structural similarity to the active phenylpropanoid compounds in the ATP-PP_i exchange assay, the substrate loaded onto the PCP was hypothesized to be either hydrocinnamate or benzoylformate (both would add 132.0 Da).

To elucidate the structure of the loaded substrate, the acyl-S-PCP protein was incubated with the surfactin thioesterase II enzyme (36), resulting in the release of the hydrolyzed substrate and intact *holo* A-PCP protein. After precipitation of the protein, analysis of the supernatant by negative ion mode ESI-MS revealed that the hydrolyzed substrate had an m/z of 149, which was identical to the hydrocinnamate control (data not shown). The retention time and ion fragmentation pattern resulting from GC/EI-MS of the substrate were also identical to the hydrocinnamate standard (Supplementary Figure 2), confirming that hydrocinnamyl-S-PCP is generated in *E. coli*. Here, the utility of unambiguous structural assignment of unexpected intermediates by small molecule mass spectrometry was demonstrated. Revisiting the ATP-PP_i exchange assay with these two substrates further confirmed hydrocinnamate as the *in vivo* acylating molecule, as hydrocinnamate was active and benzoylformate was not (Figure 3). Lastly, when *holo* A-PCP_{*in vitro*} protein was generated from the *apo* A-PCP protein in the absence of ATP, the hydrocinnamyl-S-PCP species is observed at >85% occupancy in repeated experiments (data not shown). We conclude that hydrocinnamate is adenylated *in vivo* to form hydrocinnamyl-AMP which is then retained within the adenylation active site during purification of the A-PCP didomain in a ~1:1 stoichiometry (37). Once the *holo* PCP is generated *in vitro*, the hydrocinnamyl substrate is then transferred from AMP to the thiol of the *holo* PCP, to form the hydrocinnamyl-S-PCP thioester. Only

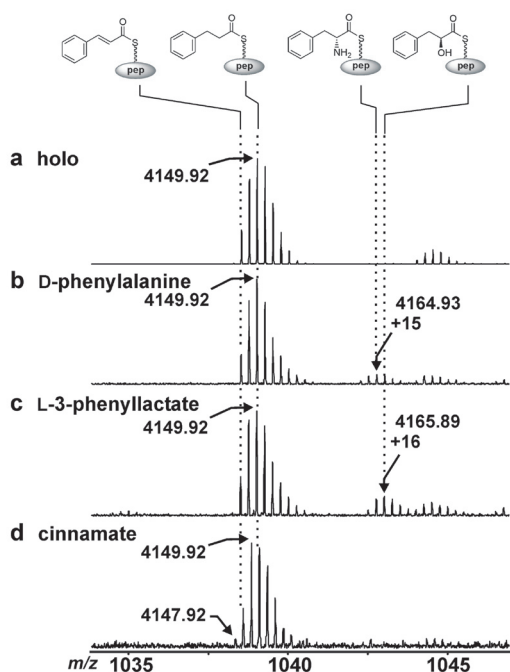


Figure 5. *In vitro* acyl-S-PCP peptide intermediates from McyG A-PCP generated with assorted phenylpropanoids: structures and mass shifts. Structures at the top of the figure from left to right show the cinnamyl- (+130 Da), hydrocinnamyl- (+132 Da), D-phenylalanyl- (+147 Da), and L-3-phenyllactyl-S-PCP (+148 Da) derived Q⁵⁷⁷-R⁶⁰⁹ peptides (pep). Mass shifts (*i.e.*, +XXX Da) are noted relative to the *holo* peptide. Shown are the *in vitro* loading on *holo* McyG-PCP (1035–1045 *m/z* region; 4+ ions) of a) hydrocinnamate derived from the growth medium, b) D-phenylalanine, c) L-3-phenyllactate, and d) cinnamate. In each case, the majority of the peptide has a MW of 4149.92, correlating to the acylation of the PCP with hydrocinnamyl-AMP following *in vitro* phosphopantetheinylation with Sfp.

a minor fraction (<5%) of the *in vitro* generated *holo* PCP contained the free thiol group, thereby limiting enzyme occupancy available for acylation.

Consistent with ATP-PP_i assays, when *holo* A-PCP_{*in vitro*} was incubated with ATP and substrates D-phenylalanine and L-3-phenyllactate, the *holo* PCP was converted to the acylated form (Figure 5, panels b and c). However, in each case, the hydrocinnamyl-S-PCP was still present as the major species. *Holo* A-PCP_{*in vitro*} was also incubated with ATP and cinnamate to generate cinnamyl-S-PCP (+130.0 Da); however identification was complicated by the presence of the

overlapping isotopic distribution from hydrocinnamyl-S-PCP (+132.0 Da) (Figure 5, panel d).

For a more robust loading to be observed, the *holo* A-PCP_{*in vitro*} reaction was incubated with surfactin thioesterase II enzyme to regenerate free, unacylated *holo* A-PCP_{*in vitro*} enzyme. Subsequent loading experiments with various substrates were carried out, the results of which are summarized in Table 1. As expected, phenylpropanoids D-phenylalanine, L-[ring-²H₅]-phenylalanine, L-3-phenyllactate, D-3-phenyllactate, phenylpyruvate, hydrocinnamate, and cinnamate were shown to load, with percent occupancies in a 20 min reaction varying from <5 to 60%, while phenylacetate did not load to any detectable occupancy in this assay. As a multichannel measurement approach with minimal bias, large molecule MS is proving to be a general approach for detection and discovery of noncovalently bound substrates and covalent intermediates on purified enzymes (21, 37, 38). For regeneration of *holo* carrier peptides, the type II TE domain allows gentle hydrolysis of the loaded species and re-acylation of the PCP domain by various substrates introduced *in vitro*.

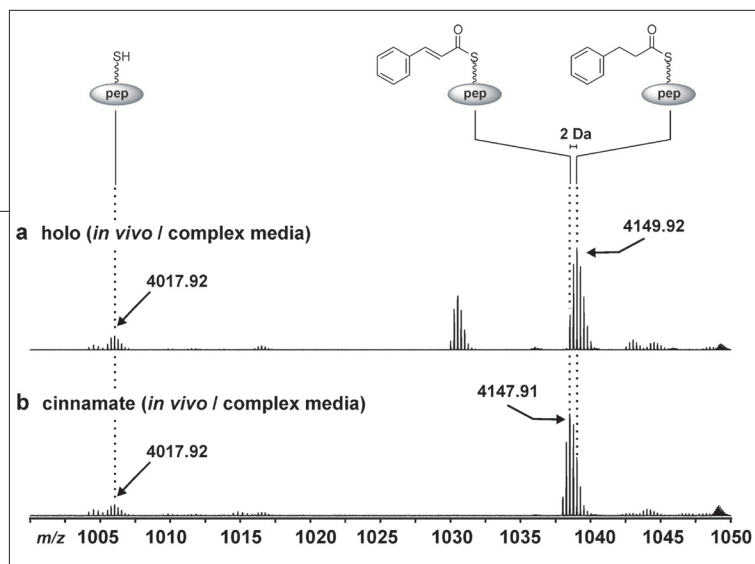
Acylation of *in Vivo* Generated *Holo* McyG A-PCP Didomain. *Holo* A-PCP_{*in vivo*} was similarly digested, fractionated, and analyzed by FTMS. Although the Gln⁵⁷⁷-Arg⁶⁰⁹ peptide was not detected in the apo form, the *holo*-form was identified and the majority of the peptide was further shifted by +132.0 Da, indicating formation of an acyl-S-PCP *in vivo* (Figure 6, panel a). The +132.0 species was assigned as the hydrocinnamyl-S-PCP based on the *in vitro* results reported above. This represents, as far as the authors know, the first direct observation of a thioester

TABLE 1. *In vitro* reloading of McyG A-PCP after TE II hydrolysis

Substrate	% occupancy
Cinnamate	20
Hydrocinnamate	50
D-Phenylalanine	40
L-[ring- ² H ₅]-Phenylalanine	5
Phenylpyruvate	<5
L-3-Phenyllactate	40
D-3-Phenyllactate	60
Phenylacetate	ND ^a

^aND, loaded species not detected.

Figure 6. *In vivo* acyl-S-PCP peptides from *holo* McyG A-PCP (complex medium): structures and mass shifts. Structures at the top of the figure from left to right show the *holo*-, cinnamyl- (+130 Da), and hydrocinnamyl-S-PCP (+132 Da) derived Q⁵⁷⁷-R⁶⁰⁹ peptides (pep) that correlate to the equivalent peaks in the mass spectra. Shown are the *in vivo* loading on *holo* McyG-PCP (1000–1050 *m/z* region; 4+ ions) of a) hydrocinnamate derived from the growth medium and b) supplemental cinnamate added during IPTG induction.



intermediate formed *in vivo* on an NRPS or PKS carrier domain, raising future possibilities for the detection of (rate limiting) intermediates covalently bound *in vivo* or stalled/misprimed synthetases from natural producing organisms for optimized combinatorial biogenesis of natural product analogues via substrate feeding during fermentation.

While hydrocinnamate is attached *in vivo* to the *holo* A-PCP when expressed in a TB growth medium, cinnamic acid had greater *in vitro* activity based on the ³²PP_i assay (Figure 3). Therefore, in order to determine if cinnamate can compete for PCP occupancy *in vivo*, it was added at 0.1 mM to a culture of *E. coli* upon induction of the *holo* enzyme. Here, substrate was available in the growth medium and could be actively transported into the cell (39). When the *E. coli* growth medium was supplemented with cinnamate, a –2 Da shift in the major species was easily detected by FTMS analysis correlating to formation of the cinnamyl-S-PCP *in vivo* (Figure 6, panel b). In a similar fashion, phenylacetate was added to the culture upon induction. However, formation of the phenylacetyl-S-PCP was not identified; rather, hydrocinnamyl-S-PCP was the major species detected. These results suggest that cinnamate is actively loaded onto the PCP *in vivo*, and despite the addition of the assumed starter unit phenylacetate, hydrocinnamate is preferentially loaded onto the PCP in direct correlation with the *in vitro* ATP-PP_i exchange assay results.

In order to further explore the *in vivo* loading properties of the McyG A-PCP didomain, the *holo* enzyme was expressed in the *E. coli* host grown in a defined, glucose-based medium. We suspected that hydrocinnamic acid, which is preferentially loaded *in vivo*, is derived from the Terrific Broth (TB) fermentation medium. By expressing the *holo* A-PCP protein in a defined glucose-based medium, it was then possible to create a growth environment devoid of these

“unnatural” *E. coli* phenylpropanoids. Although expression levels of the didomain protein were lower in the defined medium relative to expression in TB, sufficient quantities were available for characterization. Analysis of the resultant Gln⁵⁷⁷-Arg⁶⁰⁹ peptide did not yield the +132 Da shift correlating with hydrocinnamyl-S-PCP as previously observed in the TB-sample. Instead, a new +148 Da shift was detected (Figure 7, panel a). The mass of this acylated peptide correlates with the loading of 3-phenyllactate vs either phenylalanine or phenylpyruvate which are ~1 and 2 Da off, respectively. Upon addition of hydrocinnamic acid to the culture during induction, the PCP active site peptide was solely present as the hydrocinnamate loaded form (Figure 7, panel b). To investigate the origin of the 3-phenyllactate residue, L-[ring-²H₅]phenylalanine was administered upon induction to the culture grown in the defined medium. Two phenylalanine residues are present in the Gln⁵⁷⁷-Arg⁶⁰⁹ peptide, and analysis of the *holo* peptide showed the anticipated +10 Da isotope shift (Figure 7, panel c, left). On the other hand, the acylated peptide was shifted +15 Da (Figure 7, panel c, right), consistent with the incorporation of the two phenylalanine residues in the peptide chain and the phenylalanine-derived 3-phenyllactate unit. Using FTMS technology, we were able to identify some unusual characteristics of the McyG A-PCP didomain, which may not have been possible with traditional techniques. The resolution of the FTMS instrument was valuable for discrimination between various acylated species that were generated under specific growth conditions and differed in some cases by just 1 Da.

Utilization of MS Technologies To Elucidate Enzyme-Bound Intermediates. By employing a combination of MS technologies and *in vitro* biochemical assays, we establish through this study that the loading didomain of McyG has broad substrate selectivity for a variety of phenylpropanoids. Further, free

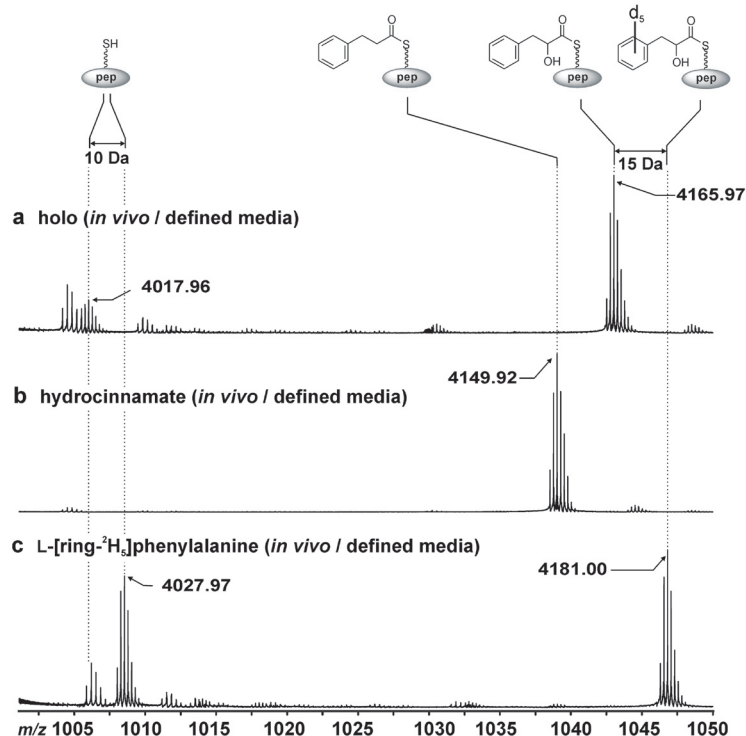


Figure 7. *In vivo* acyl-S-PCP peptides from *holo* McyG A-PCP (defined medium): structures and mass shifts. Structures at the top of the figure from left to right show the *holo*-, hydrocinnamyl- (+132 Da), and 3-phenyllactyl-S-PCP (+148 Da) derived Q⁵⁷⁷-R⁶⁰⁹ peptides (pep). Shown are the *in vivo* loading on *holo* McyG-PCP (1000–1050 *m/z* region; 4+ ions) of a) biosynthetic 3-phenyllactate, b) supplemental dihydrocinnamate added during IPTG induction, and c) biosynthetic 3-[ring-²H₅]phenyllactate derived *in vivo* from supplemental L-[ring-²H₅]phenylalanine. The +10 Da shift of the *holo*-PCP derived peptide in panel c is consistent with two [ring-²H₅]phenylalanine residues at positions 597 and 598, whereas the +15 Da shift of the 3-phenyllactyl-S-PCP species correlates to a third [ring-²H₅]phenylalanine-derived residue in the form of loaded 3-[ring-²H₅]phenyllactate.

the thioester carbonyl carbon. If the phenylpropanoid is chain extended with malonyl-CoA by the McyG PKS machinery, then the phenylalanine-derived carboxyl carbon would need to be excised, resulting in the cleavage of two C–C bonds and

phenylacetate is not a substrate and likely not the biosynthetic precursor of Adda. Results of ATP-PP_i exchange assays and MS demonstrated that substrate chain length was essential for adenylation activity of the McyG A domain, though substitutions at the α carbon were tolerated with little effect on activity. The results correlate with *Microcystis* feeding experiments (11) and suggest that the McyG A-PCP didomain has minimal editing function, consistent with previous analyses of other NRPS systems (40).

On the basis of these observations, a phenylpropanoid, and not phenylacetate, is the likely primer unit of the Adda polyketide. Preliminary data reveal that a phenylpropanoid primer unit is also incorporated into the Adda moiety of the structurally related cyanobacterial toxin, nodularin, by the NRPS module of NdaC, confirming this phenomenon is not restricted to McyG (41, M.C.M and B.S.M. unpublished data). A different phenylpropanoid, such as 3-phenyllactate which is derived from phenylalanine and was selectively activated *in vivo* from *E. coli*, may be the true starter unit. The processing of any phenylpropanoid once bound to the McyG synthetase, however, must undergo novel NRPS biochemistry involving the loss of

the formation of a new C–C bond. We are at present actively pursuing the product profile of the intact McyG protein by FTMS to determine the fate of various phenylpropanoid molecules upon PKS extension with malonyl-CoA.

Many structurally diverse bioactive natural products emanate from NRPS and PKS biosynthetic pathways. Thus, an astute understanding of how NRPSs and PKSs function at the molecular level is of high importance to further the ongoing effort to re-engineer these enzymes for efficient generation of novel therapeutics. Although substrate specificity can be predicted from the structure of the final product and the sequence of the biosynthetic gene cluster, this is based on reference to prototypical systems. Here, high-resolution MS allowed readout of *in vivo* and *in vitro* substrate preferences in a largely unbiased fashion. The ability to elucidate unexpected structures of enzyme-bound species in a discovery mode greatly complements more widespread biochemical techniques for determining substrate specificity and mechanism of thiotemplate biosynthesis, a process that increasingly defies highly stringent classification based on prototypical systems studied in detail to date.

METHODS

Bacterial Strains and Growth Conditions. *Escherichia coli* strain XL1-Blue was used for subcloning and grown in Luria broth liquid media or agar plates (Becton Dickinson). Unless otherwise stated, overexpression experiments were performed in Terrific Broth (TB) (42) using the expression strain *E. coli* BL21(DE3). *Microcystis aeruginosa* PCC7820 genomic DNA was provided by Professor R. E. Moore, University of Hawaii.

Design of McyG A–PCP Didomain Domain Expression

Vectors. The 7857 bp *mcyG* gene fragment was amplified from *M. aeruginosa* PCC7820 using DNA primers based on the 5′- and 3′-ends of the *mcyG* gene from *M. aeruginosa* strains PCC7806 (10) and K-139 (8, 9). High-fidelity PCR reactions were performed using *Pfu* Turbo polymerase as described by the manufacturer (Stratagene). Specific forward and reverse oligonucleotide primers (synthesized by Qiagen) were added to a final concentration of 0.2 μM. Amplified DNA was cloned directly into the pGEM vector (Promega) and sequence confirmed by specific primer walking.

The *mcyG* gene was amplified using the primers McyG-BamHI (5′-ggatccattccatcagttttggagg-3′) and GACP-SallRev (5′-aagctt-gctgacttattctgtaagattaatc-3′). The 2014 bp *mcyG* gene fragment encoding the NRPS A–PCP didomain was amplified using the primers McyG-BamHI (5′-ggatccattccatcagttttggagg-3′) and McyG-HindR2 (5′-ggaaaagctctgcatcacaataattgc-3′) and subcloned into pHIS₈ (43) using the restriction enzymes BamHI and HindIII. The resulting plasmid, pMAT701, was used for the expression and purification of *apo* McyG A–PCP didomain.

The 2014 bp *mcyG* gene fragment encoding the NRPS A–PCP didomain was also directionally ligated into pBM58, the pHIS₈-based vector which coexpresses Svp, the phosphopantetheinyl transferase protein from *Streptomyces verticillus* ATCC15003, from the T7 promoter (44). The resulting plasmid, pMAT692 (Svp), was used for the expression and purification of the *holo* McyG A–PCP_{*in vivo*} didomain.

Overexpression and Purification of His-Tagged McyG A–PCP.

For expression of McyG protein, 1.5 mL of an overnight culture of *E. coli* BL21(DE3), transformed with the appropriate vector, was inoculated into 75 mL of TB containing 50 μg mL⁻¹ kanamycin. The culture was grown at 37 °C to an OD₆₀₀ of between 0.4 and 0.6. Expression of the protein was then induced *via* the addition of IPTG at a final concentration of 0.5 mM, and cell growth continued for 5 h at 30 °C. Where necessary, the culture media was supplemented with 1 mg of cinnamate, L-[ring-²H₅]phenylalanine, or phenylacetate when protein expression was induced. Following expression, cells were harvested by centrifugation at 5000 rpm for 10 min and stored at –80 °C until use.

Alternatively, for the expression of *holo* McyG A–PCP_{*in vivo*}, 4.5 mL of an overnight culture of *E. coli* BL21(DE3), transformed with pMAT692 (Svp), cells was collected by centrifugation and the cell pellet washed twice with 4.5 mL of wash buffer (3 g of KH₂PO₄, 6 g of Na₂HPO₄, 0.5 g of NaCl per liter (pH 7.4)). The cells were then inoculated into 150 mL of modified M9 minimal media (pH 7.4) (45). The culture was grown at 37 °C to an OD₆₀₀ of 0.6. Expression of the protein was then induced *via* the addition of IPTG at a final concentration of 0.25 mM, and cell growth continued for 20 h at 30 °C. Where necessary, the culture medium was supplemented with 2 mg of hydrocinnamate when protein expression was induced. Alternatively, 2 mg of L-[ring-²H₅]phenylalanine was added 3 times at 2 h intervals after the cells were inoculated into the media.

For purification of the N-terminal His₈ tagged protein, cells were thawed in 15 mL of purification buffer (50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol) containing 10 mM imidazole. Cells were lysed by sonication, and the insoluble fraction separated from the soluble proteins *via* centrifugation at 13000 rpm for 30 min. The total

soluble fraction was applied to a column containing 0.5 mL of Ni-agarose (Qiagen) before washing three times with 5 mL of purification buffer containing 20 mM imidazole. The expressed protein containing the N-terminal His₈ tag was then eluted from the column with 5 × 250 μL of purification buffer containing 250 mM imidazole. Purified protein was then exchanged into Tris protein storage buffer (100 mM Tris (pH 7.2), 2 mM EDTA, 1 mM tris(2-carboxy-ethyl)phosphine hydrochloride (TCEP), 10% glycerol) using a PD-10 column (Pharmacia) and concentrated using a Microcon YM-50 centrifuge filter (Millipore). Protein was then used directly in assays or stored at –80 °C prior to use. All protein samples were analyzed by separation through a 7.5% or 10% SDS–polyacrylamide gel (BioRad) and visualized using Bio-Safe coomassie stain (BioRad). Amino acid numbering is based on the McyG sequence from *M. aeruginosa* PCC7806.

In Vitro ATP-PP_i Exchange Assays. Protein (0.75–10 μg) was combined with 1.5 mM substrate carboxylic acid and 2 mM sodium [³²P]pyrophosphate (PP_i) (activity of 500 nCi) in reaction buffer (75 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 10 mM ATP, 5 mM dithiothreitol (DTT)) at a final reaction volume of 100 μL and incubated for 45 min at room temperature. Reactions were terminated *via* the addition of 30 μL of charcoal solution (2.5% charcoal, 10% trichloroacetic acid, 20 mM Na₂HPO₄). The charcoal was pelleted by centrifugation at 13000 rpm for 5 min and washed twice with 200 μL of water. Charcoal was then resuspended in 150 μL of water, transferred to a scintillation vial with 2 mL of scintillant, and radioactivity counted.

In Vitro Phosphopantetheine Transfer and Substrate

Loading. Prior to substrate loading, the carrier domains were first phosphopantetheinylated (100 mM Tris (pH 7.2), 5–10 mM MgCl₂, 1–2 mM TCEP, 0.6 mg mL⁻¹ Svp/Sfp, 160 μM coenzyme A, and 27 μM *apo* McyG A–PCP protein) in 100 μL reactions by incubating for 3–4 h at room temperature or 30 °C. For subsequent loading of substrates onto the *holo* McyG PCP domain, 5–10 mM ATP and 1.5 mM substrate (D-phenylalanine, L-[ring-²H₅]phenylalanine, L-3-phenyllactate, D-3-phenyllactate, cinnamate, hydrocinnamate, or phenylpyruvate) were added to the reaction and incubation continued for 15 min to 1 h at room temperature or 30 °C. Samples were analyzed by MALDI-TOF MS (AZCC/SWEHSC proteomics core facility, University of Arizona). Alternatively, samples analyzed by FTMS were immediately digested using the protocol described below.

Enzymatic Hydrolysis of Substrates and Small Molecule Structural Elucidation. Covalent intermediates from the PCP domain of McyG A–PCP were hydrolyzed by the addition of a type II external thioesterase domain (SrfA-D) (36, 46) and incubation for 1–4 h at 30 °C. For small molecule characterization, proteins were precipitated by the addition of 50% TFA following incubation, and the supernatant was subsequently analyzed by negative ion mode ESI-MS (Micromass Quattro Tandem quadrupole/hexapole/quadrupole instrument) and/or GC/EI-MS (Micromass 70-VSE double focusing sector Instrument) for small molecule identification and structural validation. For McyG A–PCP reloading experiments, reactions were centrifuged in Microcon YM-50 centrifugal filter devices (Millipore) for 2 min at 9000 rpm and the ~30 μL of *holo* McyG A–PCP present in the supernatant was used for substrate loading experiments as described above.

Digestion of McyG A–PCP. Proteolysis was performed by the addition of TPCK-treated trypsin (Promega) to the target protein at protease to substrate ratios ranging from 1:5 to 1:10 w/w in 50 mM NH₄HCO₃, pH 7.8, and incubated at 30 °C for 5 min. Reactions were quenched by the addition of an equal volume of 10% formic acid (Acros) and applied to a wide-pore Jupiter C4 reversed-phase column (4.6 × 150 mm) (Phenomenex) with a linear gradient from 10% to 90% MeCN (0.1% TFA) over 60 min for fractionation/desalting prior to FTMS analysis. Samples were lyophilized before resuspension in 49% H₂O, 49% MeOH, and

2% formic acid for FTMS analysis. *Holo* and loaded active site containing HPLC fractions from 33 to 37 min were combined prior to FTMS analysis.

ESI-Q-FTMS and Reported Masses. ESI was used with a custom Q-FTMS instrument operating at 8.5 T (22). The ions were directed through a heated metal capillary, skimmer, quadrupole, and multiple ion guides into the ion cell ($\sim 10^{-9}$ Torr) of the FTMS. Scans were acquired every 1 s, and data were stored with a MIDAS data station (47) as 512k data sets. Spectra were calibrated externally using bovine ubiquitin (MW = 8559.62 Da), and theoretical isotopic distributions were generated using IsoPro v3.0 and fit to experimental data by least squares to assign the most abundant peak.

High-resolution mass spectrometry of large molecules results in isotopic distributions within the mass spectra, explanations of which have been described previously (48, 49). Briefly, while arrows in figures point to the most abundant isotopes, all molecular weight values (M_r) in this paper are reported as monoisotopic values, which refer to the molecular ion peak composed of the most abundant isotopes of the elements including the mass defect (*i.e.*, $C = 12.000000$, $N = 14.00307$, etc). Assignment of isotopic distributions to the corresponding enzyme intermediates involved correlating the experimental monoisotopic M_r values to the theoretical monoisotopic M_r values for the enzyme intermediates, with a maximum error of 20 ppm (Supplementary Table 2).

Accession Codes. The *mcyG* sequence, isolated from *Microcystis aeruginosa* PCC7820, reported in this paper has been deposited in GenBank under accession number AY910575.

Acknowledgment: We thank R. Moore for providing strains, C. Walsh for the SrfA-D plasmid, L. Miller for overexpression of the SrfA-D construct and providing Figure 1, R. Milberg for assistance with the GC/EI-MS analysis, the AZCC/SWEHSC proteomics core facility at the University of Arizona for MALDI TOF analyses (supported by NIEHS Grant ES06694 and NIH/NCI Grant CA023074-26), the DNA Sequencing Service at the University of Arizona for performing sequencing reactions, J. Becker for assistance, and J. Jez for critical reading of the manuscript. The Quattro and 70-VSE mass spectrometers were purchased in part with grants from the Division of Research Resources, National Institutes of Health (RR 07141 and RR 04648, respectively). Financial support is greatly appreciated from the Washington Sea Grant program (R/B-39) to B.S.M. and the University of Illinois and the National Institutes of Health (GM 067725) to N.L.K. L.M.H. is a recipient of an NSF Graduate Research fellowship. This article is dedicated in memory of Kenneth L. Rinehart at the University of Illinois.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Rinehart, K. L., Harada, K., Namikoshi, M., Chen, C., Harvis, C. A., Munro, M. H. G., Blunt, J. W., Mulligan, P. E., Beasley, V. R., Dahlem, A. M., and Carmichael, W. W. (1988) Nodularin, microcystin, and the configuration of Adda, *J. Am. Chem. Soc.* **110**, 8557–8558.
- Namikoshi, M., Yuan, M., Sivonen, K., Carmichael, W. W., Rinehart, K. L., Rouhiainen, L., Sun, F., Brittain, S., and Otsuki, A. (1998) Seven new microcystins possessing two L-glutamic acid units, isolated from *Anabaena* sp. strain 186, *Chem. Res. Toxicol.* **11**, 143–149.
- Beattie, K. A., Kaya, K., Sano, T., and Codd, G. A. (1998) Three dehydrobutyryne-containing microcystins from *Nostoc*, *Phytochemistry* **47**, 1289–1292.
- Sano, T., and Kaya, K. (1998) Two new (E)-2-amino-2-butenic acid (Dhb)-containing microcystins isolated from *Oscillatoria agardhii*, *Tetrahedron* **54**, 463–470.
- Meriluoto, J. A. O., Sandstrom, A., Eriksson, J. E., Remaud, G., Grey, C. A., and Chattopadhyaya, J. (1989) Structure and toxicity of a peptide hepatotoxin from the cyanobacterium *Oscillatoria agardhii*, *Toxicon* **27**, 1021–1034.
- Schwarzer, D., Finking, R., and Marahiel, M. A. (2003) Nonribosomal peptides: from genes to products, *Nat. Prod. Rep.* **20**, 275–287.
- Walsh, C. T. (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility, *Science* **303**, 1805–1810.
- Nishizawa, T., Asayama, M., Fujii, K., Harada, K., and Shirai, M. (1999) Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp., *J. Biochem. (Tokyo)* **126**, 520–529.
- Nishizawa, T., Ueda, A., Asayama, M., Fujii, K., Harada, K., Ochi, K., and Shirai, M. (2000) Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin, *J. Biochem. (Tokyo)* **127**, 779–789.
- Tillett, D., Dittmann, E., Erhard, M., von Dohren, H., Borner, T., and Neilan, B. A. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide–polyketide synthetase system, (2000) *Chem. Biol.* **7**, 753–764.
- Christiansen, G., Fastner, J., Erhard, M., Borner, T., and Dittmann, E. Microcystin biosynthesis in *Planktothrix*: genes, evolution, and manipulation, (2003) *J. Bacteriol.* **185**, 564–572.
- Rouhiainen, L., Vakkilainen, T., Siemer, B. L., Buikema, W., Haselkorn, R., and Sivonen, K. (2004) Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* strain 90, *Appl. Environ. Microbiol.* **70**, 686–692.
- Moore, R. E., Chen, J. L., Moore, B. S., Patterson, G. M. L., and Carmichael, W. W. (1991) Biosynthesis of microcystin-LR. Origin of the carbons in the Adda and Masp units, *J. Am. Chem. Soc.* **113**, 5083–5084.
- Rinehart, K. L., Namikoshi, M., and Choi, B. W. (1994) Structure and biosynthesis of toxins from blue-green algae (cyanobacteria), *J. Appl. Phycol.* **6**, 159–176.
- Chen, H., O'Connor, S. E., Cane, D. E., and Walsh, C. T. (2001) Epithilone biosynthesis: assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit, *Chem. Biol.* **8**, 899–912.
- Admiraal, S. J., Khosla, C., and Walsh, C. T. The loading and initial elongation modules of rifamycin synthetase collaborate to produce mixed aryl ketide products, *Biochemistry* **41**, 5313–5324.
- O'Connor, S. E., Chen, H., and Walsh, C. T. (2002) Enzymatic assembly of epithilones: the EpoC subunit and reconstitution of the EpoA-ACP/B/C polyketide and nonribosomal peptide interfaces, *Biochemistry* **41**, 5685–5694.
- Admiraal, S. J., Khosla, C., and Walsh, C. T. (2003) A switch for the transfer of substrate between nonribosomal peptide and polyketide modules of the rifamycin synthetase assembly line, *J. Am. Chem. Soc.* **125**, 13664–13665.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science* **246**, 64–71.
- Karas, M., and Hillenkamp, F. (1988) Laser desorption/ionization of proteins with molecular masses exceeding 10,000 daltons, *Anal. Chem.* **60**, 2299–2301.
- Kelleher, N. L., and Hicks, L. M. (2005) Contemporary mass spectrometry for the direct detection of enzyme intermediates, *Curr. Opin. Chem. Biol.* **9**, 424–430.
- Patrie, S. M., Charlebois, J. P., Whipple, D., Kelleher, N. L., Hendrickson, C. L., Quinn, J. P., Marshall, A. G., and Mukhopadhyay, B. (2004) Construction of a hybrid quadrupole/Fourier transform ion cyclotron resonance mass spectrometer for versatile MS/MS above 10 kDa, *J. Am. Soc. Mass Spectrom.* **15**, 1099–1108.
- Hicks, L., Weinreb, P., Konz, D., Marahiel, M. A., Walsh, C. T., and Kelleher, N. L. (2003) Fourier-transform mass spectrometry for detection of thioester-bound intermediates in unfractionated proteolytic mixtures of 80 and 191 kDa portions of Bacitracin A synthetase, *Anal. Chim. Acta* **496**, 217–224.

24. Mazur, M. T., Walsh, C. T., and Kelleher, N. L. (2003) Site-Specific Observation of Acyl Intermediate Processing in Thioesterase Biosynthesis by Fourier Transform Mass Spectrometry: The Polyketide Module of Yersiniabactin Synthetase, *Biochemistry* 42, 13393–13400.
25. McLoughlin, S. M., and Kelleher, N. L. (2004) Kinetic and Regiospecific Interrogation of Covalent Intermediates in the Nonribosomal Peptide Synthesis of Yersiniabactin, *J. Am. Chem. Soc.* 126, 13265–13275.
26. Hicks, L. M., O'Connor, S. E., Mazur, M. T., Walsh, C. T., and Kelleher, N. L. (2004) Mass Spectrometric Interrogation of Thioester-Bound Intermediates in the Initial Stages of Epothilone Biosynthesis, *Chem. Biol.* 11, 327–335.
27. Gameau, S., Dorrestein, P. C., Kelleher, N. L., and Walsh, C. T. (2005) Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A1 biosynthesis, *Biochemistry* 44, 2770–2780.
28. Hong, H., Appleyard, A. N., Siskos, A. P., Garcia-Bernardo, J., Staunton, J., and Leadlay, P. F. (2005) Chain initiation on type I modular polyketide synthases revealed by limited proteolysis and ion-trap mass spectrometry. *FEBS J.* 272, 2373–2387.
29. Schnarr, N. A., Chen, A. Y., Cane, D. E., and Khosla, C. (2005) Analysis of Covalently Bound Polyketide Intermediates on the 6-Deoxyerythronolide B Synthase by Tandem Proteolysis - Mass Spectrometry, *Biochemistry* 44, 11836–11842.
30. Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem. Biol.* 6, 493–505.
31. Challis, G. L., Ravel, J., and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, *Chem. Biol.* 7, 211–224.
32. Stachelhaus, T., and Marahiel, M. A. (1995) Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA, *J. Biol. Chem.* 270, 6163–6169.
33. Gehring, A. M., Lambalot, R. H., Vogel, K. W., Drucekhammer, D. G., and Walsh, C. T. (1997) Ability of *Streptomyces* spp. acyl carrier proteins and coenzyme A analogs to serve as substrates in vitro for *E. coli* holo-ACP synthase, *Chem. Biol.* 4, 17–24.
34. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) Cloning and characterization of a phosphopantetheinyl transferase from *Streptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin. *Chem. Biol.* 8, 725–738.
35. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a *Bacillus subtilis* Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases, *Biochemistry* 37, 1585–1595.
36. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases, *Proc. Natl. Acad. Sci. U.S.A.* 99, 14083–14088.
37. Miller, L. M., Mazur, M. T., McLoughlin, S. M., and Kelleher, N. L. (2005) Parallel Interrogation of Covalent Intermediates in the Biosynthesis of Gramicidin S Using High Resolution Mass Spectrometry, *Protein Sci.* 14, 2702–2712.
38. Burns, K. E., Xiang, Y., Kinsland, C. L., McLafferty, F. W., and Begley, T. P. (2005) Reconstitution and Biochemical Characterization of a New Pyridoxal-5'-Phosphate Biosynthetic Pathway, *J. Am. Chem. Soc.* 127, 3682–3683.
39. Diaz, E., Ferrandez, A., Prieto, M. A., and Garcia, J. L. (2001) Biodegradation of aromatic compounds by *Escherichia coli*, *Microbiol. Mol. Biol. Rev.* 65, 523–569.
40. Luo, L., Burkart, M. D., Stachelhaus, T., and Walsh, C. T. (2001) Substrate recognition and selection by the initiation module PheATE of gramicidin S synthetase, *J. Am. Chem. Soc.* 123, 11208–11218.
41. Moffitt, M. C., and Neilan, B. A. (2004) Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins, *Appl. Environ. Microbiol.* 70, 6353–6362.
42. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
43. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase, *Biochemistry* 39, 890–902.
44. Izumikawa, M., Cheng, Q., and Moore, B. S. (2006) Priming type II polyketide synthases via a type II nonribosomal peptide synthetase mechanism, *J. Am. Chem. Soc.* 128, 1428–1429.
45. Dorrestein, P. C., Zhai, H., McLafferty, F. W., and Begley, T. P. (2004) The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS, *Chem. Biol.* 11, 1373–1381.
46. Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated, *ChemBioChem* 5, 1290–1293.
47. Senko, M. W., Canterbury, J. D., Guan, S., and Marshall, A. G. (1996) A high-performance modular data system for Fourier transform ion cyclotron resonance mass spectrometry, *Rapid Commun. Mass Spectrom.* 10, 1839–1844.
48. Yergey, J., Heller, D., Hansen, G., Cotter, R. J., and Fenselau, C. (1983) Isotopic Distributions in Mass Spectra of Large Molecules, *Anal. Chem.* 55, 353–356.
49. McLafferty, F. W. (1994) High-Resolution Tandem FT Mass Spectrometry above 10 kDa, *Acc. Chem. Res.* 27, 379–386.

Measuring Picomolar Intracellular Exchangeable Zinc in PC-12 Cells Using a Ratiometric Fluorescence Biosensor

Rebecca A. Bozym[†], Richard B. Thompson^{†,*}, Andrea K. Stoddard[‡], and Carol A. Fierke[‡]

[†]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, 108 North Greene Street, Baltimore, Maryland 21201, and [‡]Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109

Zinc is the second most abundant “trace” element in the body. This metal ion is vital for normal cellular function as a cofactor in numerous enzymes (1), in transcription factors (2, 3), in the immune system (4), and in the reproductive system (5). In the brain, synaptically released zinc has physiological and perhaps pathological relevance (6–9); the level of free zinc ions after release may reach a range of 10–100 μM in the synaptic cleft (10). Zinc has also been shown to modulate the response of NMDA receptors at nanomolar concentrations (11, 12). Although zinc is essential for proper brain function, zinc may also operate as a neurotoxin. Added zinc ions are toxic to neurons (13); furthermore, dying neurons fill with free zinc following prolonged seizure and ischemic insult (6, 14, 15). In addition, while zinc ions at various levels (up to hundreds of micromolar) induce apoptosis in some systems (16), the membrane-permeant chelator tetrakis(2-pyridylmethyl)ethylenediamine (TPEN) also causes apoptosis, presumably due to excessive chelation of zinc. Finally, oxidative insults, including administration of nitric oxide, lead to the release of free zinc from intracellular stores (17–20) including the release of bound zinc from metallothionein (20, 21).

Although zinc plays important biological roles, little is known about the processes of distribution of this metal in the body or the incorporation of zinc into a variety of metalloproteins. Eukaryotic cells generally are rich in zinc with a total concentration in the range of 100 μM (1). However, the abundance of zinc ligands in cells, including metallothionein and other proteins, glutathione, histidine, cysteine, and diphosphate compounds (22, 23), ensures that the vast majority of cellular zinc is bound and not free. On the basis of the high affinity of certain zinc-sensitive transcription

ABSTRACT Zinc plays both physiological and pathological roles in biology, making it of increasing interest. To date, intracellular free zinc has been measured in cell types supplemented with or enriched in zinc, such as hippocampal neurons. Here we quantitatively image intracellular exchangeable zinc in an ordinary resting cell culture line (PC-12), using an excitation ratiometric fluorescent biosensor based on carbonic anhydrase (CA). Human CA II has a K_d of 4 pM for zinc and suffers no interference from millimolar calcium or magnesium ions. The CA-based biosensor was readily introduced into the cell by a novel approach: fusing a transactivator of transcription (TAT)-derived cell penetrating peptide to the CA molecule and adding it to the cells. Our results indicate that the resting concentration is approximately 5–10 pM in cytoplasm and nucleus. Interestingly, the tetrakis(2-pyridylmethyl)ethylenediamine (TPEN)–Zn complex and TPEN are both apoptogenic for this cell line.

*To whom correspondence should be addressed.
E-mail rthompso@umaryland.edu.

Received for review December 30, 2005
and accepted February 12, 2006

Published online March 10, 2006

10.1021/cb500043a CCC: \$33.50

© 2006 by American Chemical Society

factors, the concentration of free zinc in bacterial cells has been estimated as very small (femtomolar) (24). However, femtomolar free zinc equates to less than one atom of free zinc per bacterial cell and only a handful of ions in eukaryotic cells, which seems improbable in view of the high total concentration and abundant ligands with medium affinity (micromoles per liter to nanomoles per liter). If zinc complexed with labile ligands such as water or chloride is unavailable for insertion into the myriad of

transcription factors and apometalloenzymes following synthesis, then perform it must be carried and inserted by some other molecule, as has been shown for copper (25). However, copper plays a role in only a handful of enzymes, whereas zinc is a cofactor in hundreds of enzymes and transcription factors. Thus measuring the free (or rapidly exchangeable) zinc concentration in cells is central to understanding the biology of this metal ion.

Fluorescent indicators have been used to observe the intracellular free zinc under various conditions (26–29), but many of these studies were performed in cells known to be rich in free (or weakly complexed) zinc (e.g., hippocampal neurons and pancreatic islet cells). By comparison, the majority of resting cultured cells studied exhibit little cytoplasmic staining (30), except for perinuclear punctuate staining (31, 32), suggesting that cytoplasmic resting free levels are lower than the detection limits of these probes, which are generally not much under 1 nM. Some difficulties in measuring free zinc at nanomolar levels and below are widely appreciated, such as potential interference by Ca(II) and Mg(II). Less well understood are the influences of substantial (and variable) concentrations of medium strength ligands such as serum albumin, as well as contaminating zinc in the culture media.

To address these issues, we have been developing a fluorescent indicator system using apocarbonic anhydrase as a sensor transducer (33–36). Human carbonic anhydrase II (CA) exhibits picomolar affinity for zinc at pH 7.5, and its response to zinc is unaffected by Ca(II) at 10 mM or Mg(II) at 50 mM, levels signifi-

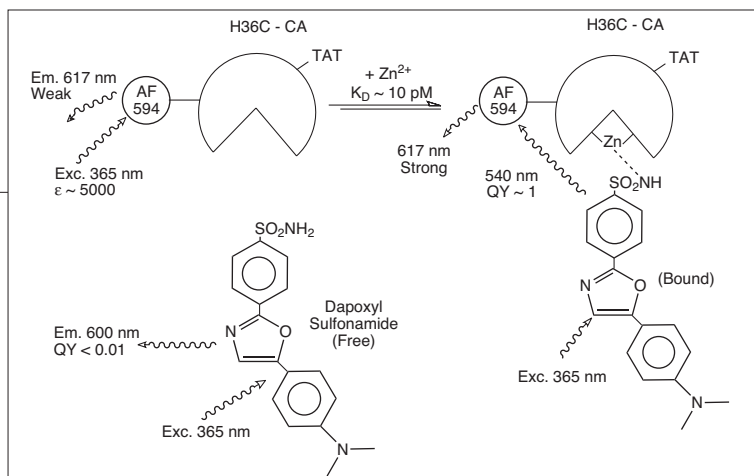


Figure 1. Schematic of ratiometric zinc determination with apoTAT-H36C-Alexa Fluor 594 carbonic anhydrase and Dapoxyl sulfonamide. In the absence of zinc, Dapoxyl sulfonamide does not bind to CA, therefore no FRET occurs to the fluorescent label on the protein and very weak emission at 617 nm is observed. In the presence of zinc, Dapoxyl sulfonamide binds to zinc, is excited with UV light, and FRET occurs from Dapoxyl to the Alexa Fluor, thus exciting the fluorescent label causing emission to occur at 617 nm. The emission when excited with 365 nm is normalized to the amount of labeled protein by directly exciting Alexa Fluor 594 with 543 nm and dividing the two intensities.

cantly higher than those encountered *in vivo* (34). The use of a biological molecule for the transducer allows facile modification of the system through site-directed mutagenesis (37). Single amino acid substitutions in the protein sequence can improve the selectivity, sensitivity, and response time of carbonic anhydrase (35), making it a more flexible and responsive indicator. For instance, a variant CA with a single amino acid mutation (E117A) exhibits a several-hundred-fold faster association rate constant for Zn(II) than wild-type CA, with only a 10-fold decrease in affinity (35). We have developed means to transduce metal binding (and thus concentration) as changes in fluorescence intensity, intensity ratio, anisotropy, or lifetime (33, 38, 39). The latter three lend themselves to easier calibration and use with microscopes for imaging the location of the free ion, since many labs are already equipped with fluorescence microscopes for calcium imaging.

To image intracellular zinc concentrations, we have adapted apocarbonic anhydrase for use as an excitation ratiometric fluorescent biosensor based on fluorescence resonance energy transfer (FRET) from a zinc-bound aryl sulfonamide to a fluorescent label on the protein (36). Zinc levels are measured by taking the ratio of intensities at two different excitation wavelengths (36). This approach (Figure 1) is based on FRET from Dapoxyl sulfonamide (whose binding is zinc-dependent) to the fluorescent label, Alexa Fluor 594 (AF594), covalently attached to the protein.

A stumbling block to the use of protein (or other macromolecule) biosensors in the cell is the difficulty of introducing them into cells. An elegant approach was

described wherein a protein-based fluorescence biosensor for calcium was expressed in the cell by transfecting the cell line with a gene for the fused construct comprising calmodulin flanked by GFP variants which serve as FRET donor and acceptor (40). While powerful, this approach is difficult to adapt to larger organisms which may be opaque, or tissues thereof, and control of expression is essential. Others have used microinjection, biolistic loading, and electroporation, but these may not be very gentle. We adapted the CA-based biosensor to intracellular use by attaching a transactivator of transcription (TAT) peptide to carbonic anhydrase which then induces cultured cells to take up the protein (41). Others have used TAT or similar cell penetrating peptides for sensing using TAT-decorated polymer nanoparticles named PEBBLES (42, 43) and nucleic acids in the form of molecular beacons or aptamers (44). We found that TAT-fused fluorescent-labeled CA was rapidly taken up by cells with no apparent ill effects in concentrations sufficient to permit fluorescence microscopy. Using this ratiometric biosensor, intracellular zinc ions were imaged and quantified in the picomolar range in PC-12 cells.

RESULTS AND DISCUSSION

Principle of the Ratiometric Zn(II) Biosensor.

Apocarbonic anhydrase has been adapted for use as an excitation ratiometric fluorescent biosensor based on fluorescence resonance energy transfer (FRET) from the zinc-bound Dapoxyl sulfonamide to Alexa Fluor 594 covalently attached to a single, engineered cysteine on CA (Figure 1) (36). Zinc levels are measured from the ratio of fluorescence intensities at two different excitation wavelengths (36). In the absence of zinc, Dapoxyl sulfonamide does not bind to apo-CA, so only the weak emission from free Dapoxyl sulfonamide is observed at 617 nm upon excitation at 365 nm. In contrast, Dapoxyl sulfonamide coordinates to the active site zinc of holo-CA, replacing water as the fourth ligand. This causes an enhancement of the fluorescent emission at 617 nm with UV excitation, due to FRET from bound Dapoxyl sulfonamide exciting the covalently attached Alexa Fluor. The change in emission when excited at 365 nm is normalized to the concentration of labeled CA by exciting the Alexa Fluor 594 directly (at 543 nm, where Dapoxyl sulfonamide does not absorb) and dividing the two intensities. In addition to being ratiometric, this approach has the added advantage that there is

practically no interference from the emission of Dapoxyl sulfonamide bound to adventitious carbonic anhydrase, lipids, or other proteins. For intracellular use, a TAT peptide is fused to the N-terminus of carbonic anhydrase, which allows the cell to carry the protein across cellular membranes and into the cytoplasm and nucleus (41). Dapoxyl sulfonamide penetrates the cell readily. This excitation ratiometric fluorescence method can be used for quantitatively imaging intracellular readily exchangeable zinc at picomolar levels. The advantages of excitation ratiometric methods for quantitatively imaging free metal ion concentrations such as calcium are well-known, including relative freedom from variations in excitation intensity, specimen thickness, and fluorophore concentration due to washout or bleaching (45).

Cellular zinc levels in the micrographs were calibrated by imaging wells containing known concentrations of free zinc with the CA-based indicators (Supplementary Figure 1). The mean fluorescence excitation intensity ratio for each well was used to plot the calibration curve from the microscope and compare it against the one acquired on the fluorometer (Figure 2). The curve from the microscope reveals an apparent K_d of 70 ± 15 pM whereas the fitted $K_{d,app}$ using the fluorometer is 137 ± 18 pM. We note that the optical trains of the microscope (using filters) and the fluorometer used for the cuvette measurements (using monochromators) are quite different,

such that agreement to within a factor of 2 is satisfactory. The ratio increases with increasing zinc concentration by almost 50%, which is slightly less than that achieved with the fluorometer. The response of the system is relatively insensitive to variations in protein and Dapoxyl sulfonamide concentrations (see Supplementary Figures 2 and 3). To visually compare

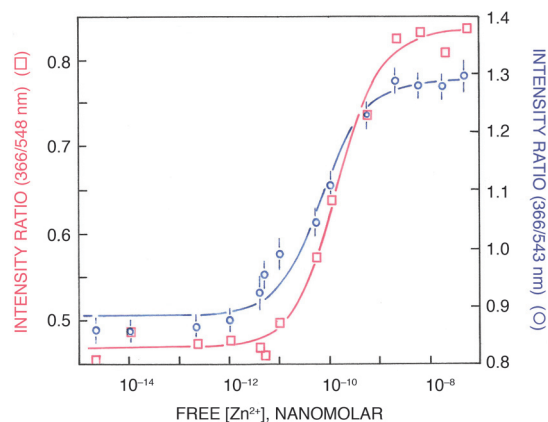


Figure 2. Calibration of zinc biosensor. Zinc-dependent excitation fluorescent ratio of 366 nm over 548 nm with emission at 617 nm of apo-H36C-AF594-CA and Dapoxyl sulfonamide performed on the microscope in a 1536-well plate (○) and on a steady-state fluorometer in a cuvette (□). The free zinc concentrations are maintained with NTA metal buffers. A single binding isotherm is fit to these data: $K_d = 70 \pm 15$ pM (○) and 137 ± 18 pM (□).

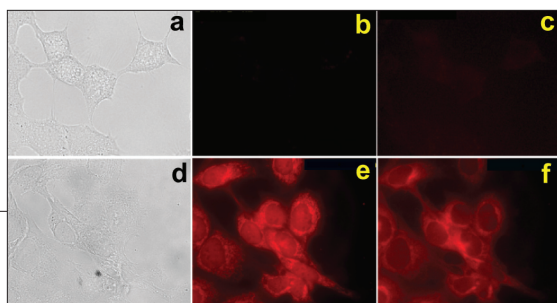


Figure 3. Comparison of intrinsic cellular fluorescence versus label fluorescence. These experiments were performed after addition of TAT-H36C-AF594-CA and Dapoxyl sulfonamide to PC-12 cells. Top row, unlabeled PC-12 cells: a) brightfield; b) excitation 543 nm, emission 617 nm; c) excitation 365 nm, emission 617 nm. Bottom row, PC-12 cells incubated with apo-TAT-H36C-AF594 CA and Dapoxyl sulfonamide: d) brightfield; e) excitation 543 nm, emission 617 nm; f) excitation 365 nm, emission 617 nm. All of the fluorescent images were taken with an exposure time of 500 ms with the exception of panel e, which was taken with an exposure time of 5 ms.

Biosensor. Brightfield and fluorescence images of PC-12 cells in the presence and absence of added apoTAT-H36C-AF594-CA and Dapoxyl sulfonamide were obtained to compare the level of background fluorescence with the fluorescence level of stained cells (Figure 3). PC-12 cells were examined in the absence of TAT-H36C-AF594-CA using an excitation wavelength of 365 nm (Figure 3, panel c) or 543 nm (Figure 3, panel b) and an emission wavelength of 617 nm with an exposure time of 500 ms. Basal fluorescence is slightly higher with UV excitation than with 543 nm excitation. PC-12 cells were incubated with apoTAT-H36C-AF594-CA and Dapoxyl sulfonamide for 15 min before images (Figure 3, panels d–f) were taken using the same excitation and emission wavelengths and conditions listed above unless otherwise noted. With the added TAT-H36C-AF594-CA and Dapoxyl sulfonamide, red fluorescence (617 nm) is observed for both excitation at 365 nm (Figure 3, panel f) and 543 nm (Figure 3, panel e). The fluorescence of the nuclei is brighter with excitation at 543 nm than at 365 nm, whereas the cytoplasm and processes stain similarly at the two excitation wavelengths, with brighter areas surrounding the nuclei.

In a control experiment, images of cells incubated with Dapoxyl sulfonamide alone (Figure 4, panels a–c) show almost no cellular or even background fluorescence with excitation at 543 nm and emission at 617 nm (Figure 4, panel b). In addition, only minimal background

fluorescence is present using excitation and emission wavelengths of 365 and 617 nm, respectively (Figure 4, panel c), with the cells barely visible. Dapoxyl sulfonamide will stain lipid bilayers with bright fluorescence emission peaking at around 504 nm, but evidently little of this emission is observed at 617 nm. To determine if the transducer protein can enter cells in the absence of the TAT peptide fusion, H36C-AF594-CA without the TAT peptide attached was incubated with PC-12 cells and Dapoxyl sulfonamide. The data indicate that there is no measurable protein uptake in the absence of the TAT peptide (Figure 4, panels d–f). Finally, to investigate whether the staining process would induce apoptosis or cause any other adverse effect, PC-12 cells were incubated with TAT-H36C-AF594-CA and Dapoxyl sulfonamide (separately and together). The cells survived for 2 weeks after addition of fluorophores and they did not stain with the Annexin V–Oregon Green conjugate, indicating that there was no prompt toxicity associated with either the labeled CA or Dapoxyl sulfonamide and that apoptosis was not induced (data not shown).

Probing PC-12 Cells with CA-Based Zinc

CA-Based Zinc

Measuring Rapidly Exchangeable Zinc Ions inside of PC-12 Cells. The free or rapidly exchangeable zinc level inside of PC-12 cells was measured by incubating the cells with apoTAT-H36C-AF594-CA and Dapoxyl sulfonamide and imaging the fluorescence excitation ratio. Under these conditions, the cells fluoresce

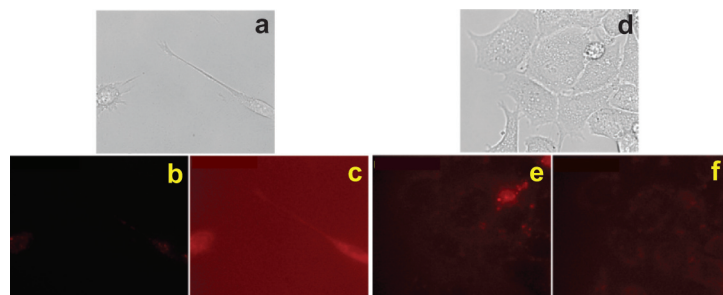


Figure 4. Fluorescence of cells stained with Dapoxyl sulfonamide alone and labeled CA lacking TAT peptide. a–c) Cells were incubated with 1 μ M Dapoxyl sulfonamide for 15 min. Images were taken at an excitation wavelength of b) 543 nm or c) 365 nm with an emission wavelength of 617 nm and an exposure time of 500 ms. Some background fluorescence is seen when the excitation wavelength is 365 nm although the cells are barely visible. d–f) Addition of H36C-AF594-CA to PC-12 cells. Labeled CA lacking the TAT peptide was added to PC-12 cells. Pictures were taken using excitation wavelengths of e) 543 nm and f) 365 nm, and an emission wavelength of 617 nm, using an exposure time of 500 ms. Cells did not stain with this protein although a slight red hue is seen most likely due to protein adhering to the cell culture dish.

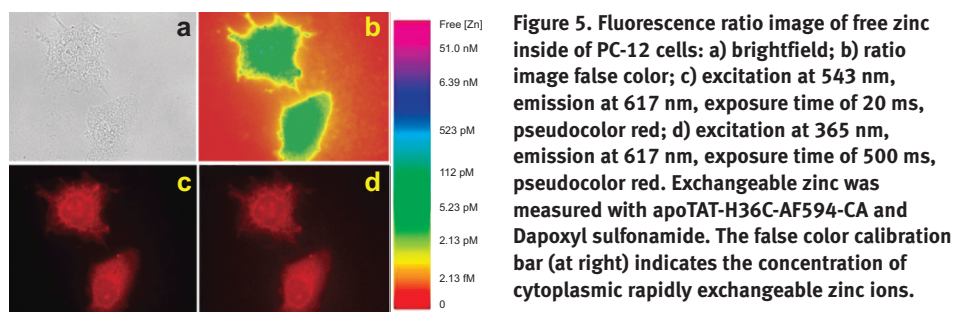


Figure 5. Fluorescence ratio image of free zinc inside of PC-12 cells: a) brightfield; b) ratio image false color; c) excitation at 543 nm, emission at 617 nm, exposure time of 20 ms, pseudocolor red; d) excitation at 365 nm, emission at 617 nm, exposure time of 500 ms, pseudocolor red. Exchangeable zinc was measured with apoTAT-H36C-AF594-CA and Dapoxyl sulfonamide. The false color calibration bar (at right) indicates the concentration of cytoplasmic rapidly exchangeable zinc ions.

brightly with excitation at 543 nm and emission at 617 nm (Figure 5, panel c) requiring an exposure time of only 20 ms. By comparison, the fluorescence with excitation at 365 nm is somewhat dimmer (Figure 5, panel d), needing an exposure time of 200–500 ms on average. The false color ratio image of the stained PC-12 cells (Figure 5, panel b), when compared to the microscope calibration standards, indicates an intracellular free zinc level of ~5 pM throughout the cells. For the majority of cells, the fluorescence excitation ratio is uniformly distributed throughout the cells, suggesting a relatively uniform free zinc concentration within the cell. These fluorescence ratios are observed within 15 min of staining and remain constant for hours. Under these conditions a few cells exhibit high levels of zinc, particularly in the nucleus, as well as perinuclear punctuate staining; however, these cells appeared unhealthy morphologically, exhibiting nuclear and cell membrane changes (Figure 6). The nuclear and perinuclear punctuate zinc levels are clearly above the dynamic range of our assay using this CA variant (*e.g.*, a few nanomolar or higher). We suspect that the apparent high levels of zinc indicate that such cells are dead or dying, perhaps as a result of apoptosis. It may be that elevated free zinc levels in cells indicate that the apoptosis process has begun.

The 5 pM value we observe is low, but not zero and well above the femtomolar levels proposed for prokaryotic cells (24). We note that if free levels were in the femtomolar regime (below the dynamic range of the wild-type protein), the cell cytoplasm would appear red (Figure 5, panel b). In order to confirm these

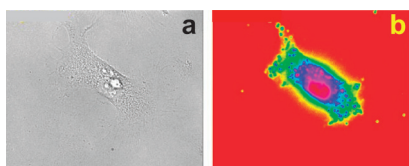


Figure 6. Ratiometric image of free zinc inside moribund PC-12 cell: a) brightfield and b) fluorescence ratio image of a morphologically unhealthy cell. Colors representing zinc concentrations are the same as in Figure 5.

results as well as further explore the kinetics of zinc equilibration with the biosensor (see below), we also attempted to calibrate our approach *in situ* using a variant of carbonic anhydrase with slightly lower affinity (K_d of 40 pM) but approximately 800-fold faster equilibration (46). Under comparable conditions, the apoTAT-E117A-H36C-Alexa Fluor 594-CA excitation ratio indicates a resting zinc concentration of approximately 5 pM (Figure 7, panel b), essentially identical to the value obtained using the wild-type protein.

Buffering of Intracellular Free Zinc. It might reasonably be asked how intracellular apocarbonic anhydrase concentrations probably in the nanomolar regime can be appreciably saturated by picomolar concentrations of free zinc. For the calibrations in the cuvette and microscope, this result is obtained by use of a metal ion buffer: a chelator is present which binds nearly all of the added metal ion leaving a small concentration (picomolar) free (47). The buffer acts such that binding of zinc by the apoprotein is offset by release from the chelator, maintaining the free level constant as long as the chelator is in substantial excess over the apoprotein, which is easy to arrange. We believe that substantially the same process ordinarily occurs in the cell, wherein known intracellular zinc chelators such as reduced glutathione, histidine, cysteine, and several

proteins such as metallothionein buffer the intracellular free zinc level essentially at the picomolar level despite the total cellular concentration being of the order of 100 μ M.

Our preliminary observations suggest that cells in distress and perhaps undergoing apoptosis exhibit rather higher free zinc levels and that loss of zinc homeostasis in the cell is an indication of cellular stress.

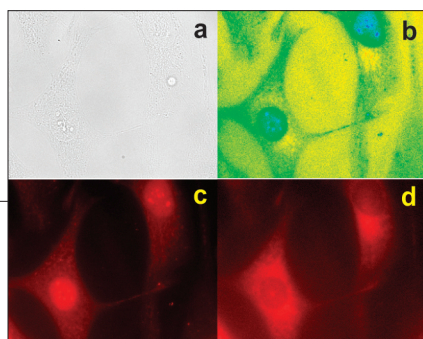


Figure 7. Measuring free zinc inside cells with a rapidly equilibrating CA variant: a) brightfield image; b) false color ratio image; c) excitation 543 nm, emission 617 nm, and exposure time of 50 ms, pseudocolored red; and d) excitation 365 nm, emission 617 nm, and an exposure time of 500 ms, pseudocolored red. Cells were stained with the apo-TAT-E117A-AF594 CA variant having faster zinc binding kinetics but slightly reduced affinity instead of wild type CA.

Zinc Equilibration Kinetics.

In vitro zinc ion binding kinetics of wild-type apo-CA II are quite slow, with association rate constants of 10^4 to 10^5 $M^{-1} s^{-1}$ and half-times for zinc dissociation of several months (37, 48). Therefore, we expected that the apoprotein would require hours to equilibrate with picomolar levels of intracellular free zinc. In fact the equilibration occurs in minutes. Additionally, the same

intracellular zinc concentration and rapid equilibration were found with the E117A variant, which has approximately an 800-fold faster association rate constant (35). These results indicate that we are measuring an equilibrium or steady state zinc concentration that is not limited by the zinc-binding kinetics of carbonic anhydrase. The rapid equilibration (minutes) of free zinc at picomolar levels with both proteins indicates that some catalysis of the binding process is occurring. The most likely explanation for this result is that the zinc ion binds to the apoprotein not as free zinc but as a zinc-chelate complex present at higher concentrations than picomolar. This suggestion is somewhat less surprising in view of the discovery that compounds such as 2,6-dipicolinate (49) or nitrilotriacetic acid (50) can catalyze the binding to or dissociation from carbonic anhydrase of zinc (51), and indeed the former is routinely used to remove zinc from the protein. We speculate that other aminocarboxylic acids normally present in the cell might also catalyze this reaction, especially since metallation of carbonic anhydrase and perhaps other apoproteins following polypeptide synthesis would otherwise be expected to take hours at picomolar levels of free zinc, which seems unlikely in view of the generally reduced stability of metal-free apoproteins. It may be that there are also zinc chaperone protein molecules analogous to the recently discovered copper chaperones (25) which specifically insert zinc ion into apoproteins, but we note that the much larger number and variety of zinc-containing proteins (including hundreds of zinc finger motifs in transcription factors) compared with the

relative handful of copper proteins suggests that either a much larger number of zinc chaperones might be needed or each chaperone inserts zinc into a variety of proteins. The parsimony of having a small molecule (perhaps a metabolite or precursor routinely present in abundance) perform the “chaperone” role is appealing.

***In Situ* Calibration of Zinc Indicator.** We attempted to calibrate the CA-based indicator inside the cells by bathing the cells in Neurobasal media whose free zinc level (approximately 10 nM) had been altered by addition of excess zinc ion to raise free zinc, or a chelator (NTA) to lower it, together with an ionophore (pyrithione) (52) to rapidly equilibrate zinc levels across the cell membrane. We found that pyrithione interferes in our method by competing with Dapoxyl sulfonamide binding to carbonic anhydrase (results not shown), so we replaced the ionophore with a different one used in zinc-selective electrodes, tetra-*n*-butylthiuram disulfide (TBTDs), which was satisfactory. However, we found that both high and low levels of zinc led to high nuclear zinc levels, perinuclear punctuate staining, and morphological changes suggesting toxicity, making these images difficult to interpret.

Chelation of Intracellular Zinc with TPEN. To further test the response of this sensor to changes in intracellular zinc concentration, we attempted to reduce intracellular free zinc ion concentrations by adding a high affinity, membrane-permeable chelator, TPEN, to PC-12 cells. As a control, TPEN complexed with zinc (TPEN-Zn) was added to PC-12 cells as well. Previously, Zalewski and others had demonstrated that TPEN induces apoptosis in cells (18, 53) (reviewed in (16)). Interestingly, PC-12 cells began rounding up after 4 h

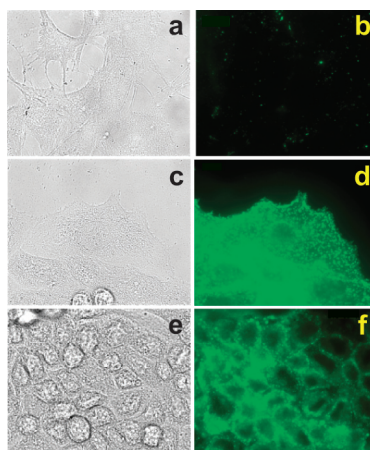


Figure 8. The zinc chelator TPEN and TPEN-Zn complex both induce apoptosis. Brightfield (panels a, c, and e) and fluorescence micrographs (panels b, d, and f) of Annexin V–Oregon Green stained PC-12 cells after incubation with TPEN and TPEN-Zn. PC-12 cells were incubated for 4 h with either media only (panels a and b), media with 1 μM TPEN (panels c and d), or 1 μM TPEN plus stoichiometric zinc (panels e and f).

of incubation with either TPEN or TPEN-Zn, a sign of apoptosis. Staining with the Annexin V–Oregon Green conjugate (which binds to exposed phosphatidylserine residues, an indicator of apoptosis) revealed that incubation with either TPEN or TPEN-Zn did in fact induce apoptosis (Figure 8). In the control dish (Figure 8, panels a and b), little or no fluorescence was observed when the Annexin V–Oregon Green conjugate was added to PC-12 cells. By comparison, PC-12 cells incubated with TPEN (Figure 8, panels c and d) or TPEN-Zn (Figure 8, panels e and f) show significant staining with the Annexin V–Oregon Green conjugate indicative of apoptosis (54). However, the intensity of the staining of the PC-12 cells is reduced upon incubation with TPEN-Zn compared to TPEN. These results suggest that TPEN does not induce apoptosis in cells exclusively (and perhaps not even mainly) by chelating zinc and certainly call into question its use as a reagent for this purpose.

Measuring Intracellular Free Zinc in CHO Cells. To assure that the results obtained above in PC-12 cells were not specific to that cell type, we also performed measurements in Chinese hamster ovary cells, a widely used cell line. The cells were grown in CD CHO A, a serum-free proprietary medium chosen in this case to minimize variations in zinc and zinc ligands likely to be present in serum-containing media. We found that healthy CHO cells generally exhibited zinc levels (Figure 9) that differed little from those of the PC-12 cells (compare Figure 5), although one of the cells in

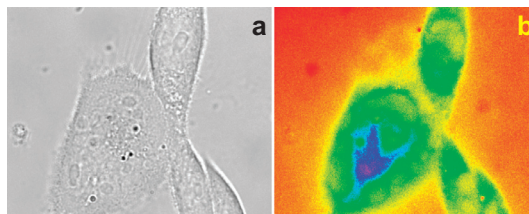


Figure 9. CHO cells have free-zinc levels similar to PC-12 cells: a) brightfield and b) false color fluorescence ratio micrographs of CHO cells stained with apoTAT-H36C-AF594-CA and Dapoxyl sulfonamide using conditions as described in the text.

the figure exhibits higher levels around the nucleus.

In summary, we have demonstrated a very sensitive method for quantitatively imaging low levels of intracellular zinc using fluorescence microscopy. The attachment of the TAT peptide to carbonic anhydrase facilitates the transfer of the transducer protein into the cells, bypassing any need for transfection or harsh injection strategies. The flexibility of using variants of the protein having different affinities, selectivities, and binding kinetics allows the use of this method in a variety of systems and to answer myriad questions. Moreover, the use of the TAT approach with other fluorescent protein biosensors such as those modified from bacterial periplasmic binding proteins (55, 56) suggests that sensors can be made for a large variety of intracellular analytes of interest. We hope that this new approach will lead to answers for many of the urgent questions now emerging in zinc biology.

METHODS

Carbonic Anhydrase II. Fused to the N-terminus of human carbonic anhydrase II (CA) is a His₆-tag, a bridging glycine, an 11 amino acid TAT peptide, and another glycine giving the following N-terminal amino acid sequence: MHHHHHHGYGRKKRRQRRRG (41, 57). The His₆-TAT-CA gene was constructed by three rounds of PCR amplification using the CA gene as a template with a series of oligonucleotides containing the necessary nucleotide sequences to encode the His₆-tag and the TAT peptide at the 5' end of the CA gene. The constructed His₆-TAT-CA gene was then inserted into pACA (58), the ampicillin resistant T7 expression plasmid that encoded the wild-type CA, replacing the entire wild-type gene using a Seamless Cloning Kit (Stratagene) to generate the plasmid pAHTC1. The new gene product was cloned into the vector through the introduction of a restriction enzyme site, *Eam*1104 I, which cleaves outside its recognition sequence and which incorporates ³²P-dCTP during PCR amplification, protecting pre-existing *Eam*1104 I sites. Following amplification, digestion with *Eam*1104 I and ligation in the presence of the enzyme resulted in the desired vector. Site-directed mutagenesis of pAHTC1 was then performed (QuikChange Site-Directed Mutagenesis Kit, Stratagene) to produce the CA mutants. The double mutant

C206S/H36C His₆-TAT-CA (pAHTC3) was constructed for site-specific labeling with fluorophores and the C206S/H36C/E117A His₆-TAT-CA triple mutant was prepared to increase the zinc equilibration rate of the protein (46). The plasmids were transformed into *E. coli* BL21 (DE3) cells and grown in rich induction medium at 37 °C (57). During the mid-log phase of cell growth, protein expression was induced by the addition of 0.25 mM isopropyl β-D-thiogalactopyranoside and 0.4 mM ZnSO₄ and the cells were incubated for 6 h at 30 °C and 225 rpm before harvesting. These His₆-TAT-CA proteins were isolated from the cell lysate by binding the histidine tag on the fusion protein to a Ni²⁺-charged chelating Sepharose fast flow resin (Amersham Biosciences) in 30 mM HEPES pH 8, 250 mM NaCl, 2 mM imidazole, washing extensively, and then eluting the CA protein with a 0.1–0.5 M imidazole gradient in 30 mM HEPES, pH 8, 250 mM NaCl. *E. coli* endotoxins were removed by running the protein over DEAE Sephacel resin (Amersham Biosciences) in 10 mM Tris SO₄, pH 8, 0.1 mM Zn SO₄, 1 mM DTT. The protein bound loosely to the resin and was eluted with 0.1–0.2 M (NH₄)₂SO₄ in the same buffer. Typical protein yields are 5–7 mg L⁻¹ of culture medium. The protein is conjugated with Alexa Fluor 594 maleimide (Molecular Probes) and stripped of zinc as previously described (49, 59). Dapoxyl sulfonamide was synthesized and

purified by silica gel column chromatography as previously described (60).

Cell Culture. PC-12 cells, a rat pheochromocytoma cell line, were grown in Neurobasal-A (without phenol red) with 2% B-27 supplement and 0.5 mM L-glutamine (all Gibco, Invitrogen Corp.) (61) on 35 mm glass bottom dishes (World Precision Instruments) at 37 °C with 5% CO₂. The glass bottom dishes are convenient for focusing and using the Olympus 60×/1.2 NA oil objective on the microscope. PC-12 cells were stained by immersion in an isotonic medium stripped of zinc (60) containing 1 μM Alexa Fluor 594-labeled apo-TAT-H36C-CA and 1 μM Dapoxyl sulfonamide (dissolved in DMSO) for 20 min. Afterward the dish was rinsed three times with zinc-free isotonic saline and then Neurobasal-A medium was added back to the cells. The Annexin V–Oregon Green 488 conjugate (Molecular Probes) was used to detect apoptosis. CHO cells were grown in CD CHO-A medium (Gibco, Invitrogen Corp.) with 8 mM L-glutamine under the same conditions as the PC-12 cells. The same staining procedure was also used in measuring the level of rapidly exchangeable zinc in CHO cells.

Fluorescence Microscopy. Cells were photographed using a Nikon Eclipse TE300 epifluorescence microscope with a D540-25X or a D360-40X filter in the excitation light path, 400DLCP dichroic, and D630-30M barrier filter (all Omega), through a Olympus 60×/1.2 NA oil objective or a Nikon Plan Fluor 100× oil/1.3 NA objective with a Cooke Sencis QE cooled CCD camera. Images were captured using IPLab software (Scanalytics, Inc.). For microscope calibration, Alexa Fluor-labeled apo-TAT-H36C-CA and Dapoxyl sulfonamide were dissolved in a series of NTA-MOPS zinc buffers with controlled concentrations of free (or rapidly exchangeable) zinc ion, formulated using the MINEQL program (Environmental Research Software), added into the wells of an inverted 1536-well plate, and photographed using the same optical filters, using a Nikon Plan Apo 4×/0.2 NA objective. In order to determine the fluorescence ratio for each well, the images for each excitation wavelength were divided and an oval region of interest was used to calculate the mean ratio for each well. Steady-state fluorescence spectra were obtained on a Spectronics AB-2 fluorometer with the same zinc ion buffers.

Acknowledgment: The authors wish to thank Chris Frederickson for many fruitful discussions and Krystyna Grycznska for some figures. The authors also wish to thank the National Institute of Biomedical Imaging and Bioengineering (1 RO1 EB03924 to R.B.T.) for support. Some of these data were reported in preliminary form at the SPIE Conference on Advanced Biomedical and Clinical Diagnostic Systems.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Vallee, B. L., and Falchuk, K. H. (1993) The biochemical basis of zinc physiology, *Physiol. Rev.* **73**, 79–118.
- Berg, J. M., and Shi, Y. (1996) The galvanization of biology: a growing appreciation for the roles of zinc, *Science* **271**, 1081–1085.
- O'Halloran, T. V. (1993) Transition metals in control of gene expression, *Science* **261**, 715–725.
- Fraker, P. J., and Telford, W. G. (1997) A reappraisal of the role of zinc in life and death decisions of cells, *Proc. Soc. Exp. Biol. Med.* **215**, 229–236.
- Costello, L. C., and Franklin, R. B. (1998) Novel role of zinc in the regulation of prostate citrate metabolism and its implications in prostate cancer, *The Prostate* **35**, 285–296.
- Frederickson, C. J., Hernandez, M. D., and McGinty, J. F. (1998) Translocation of zinc may contribute to seizure-induced death of neurons, *Brain Res.* **480**, 317–321.
- Frederickson, C. J. (1989) The neurobiology of zinc and of zinc-containing neurons, *Int. Rev. Neurobiol.* **31**, 145–238.
- Frederickson, C. J., and Bush, A. I. (2001) Synaptically released zinc: Physiological functions and pathological effects, *BioMetals* **14**, 353–366.
- Choi, D. W., and Koh, J. Y. (1998) Zinc and brain injury, *Annu. Rev. Neurosci.* **21**, 347–375.
- Vogt, K., Mellor, J., T Ong, G., and Nicoll, R. (2000) The actions of synaptically-released zinc at hippocampal mossy fiber synapses, *Neuron* **26**, 187–196.
- Howell, G. A., Welch, M. G., and Frederickson, C. J. (1984) Stimulation-induced uptake and release of zinc in hippocampal slices, *Nature* **308**, 736–738.
- Paoletti, P., Ascher, P., and Neyton, J. (1997) High-affinity zinc inhibition of NMDA NR1-NR2A receptors, *J. Neurosci.* **17**, 5711–5725.
- Yokoyama, M., Koh, J., and Choi, D. W. (1987) Brief exposure to zinc is toxic to cortical neurons, *Neurosci. Lett.* **71**, 351–355.
- Tonder, N., Johansen, F. F., Frederickson, C. J., Zimmer, J., and Diemer, N. H. (1990) Possible role of zinc in the selective degeneration of dentate hilar neurons after cerebral ischemia in the adult rat, *Neurosci. Lett.* **109**, 247–252.
- Koh, J. Y., Suh, S. W., Gwag, B. J., He, Y. Y., Hsu, C. Y., and Choi, D. W. (1996) The role of zinc in selective neuronal death after transient global cerebral ischemia, *Science* **272**, 1013–1016.
- Truong-Tran, A. Q., Carter, J., Ruffin, R. E., and Zalewski, P. D. (2001) The role of zinc in caspase activation and apoptotic cell death, *BioMetals* **14**, 315–330.
- Cuajungco, M. P., and Lees, G. (1998) Nitric oxide generators produce accumulation of chelatable zinc in hippocampal neuronal perikarya, *Brain Res.* **799**, 118–129.
- Aizenman, E., Stout, A. K., Hartnett, K. A., Dinely, K. E., McLaughlin, B., and Reynolds, I. J. (2000) Induction of neuronal apoptosis by thiol oxidation: putative role of intracellular zinc release, *J. Neurochem.* **75**, 1878–1889.
- Jiang, D., Sullivan, P. G., Sensi, S. L., Steward, O., and Weiss, J. H. (2001) Zn²⁺ induces permeability transition pore opening and release of pro-apoptotic peptides from neuronal mitochondria, *J. Biol. Chem.* **276**, 47524–47529.
- Bossy-Wetzel, E., Talantova, M. V., Lee, W. D., Scholzke, M. N., Harrop, A., Mathews, E., Gotz, T., Han, J., Ellisman, M. H., Perkins, G. A., and Lipton, S. A. (2004) Crosstalk between nitric oxide and zinc pathways to neuronal cell death involving mitochondrial dysfunction and p38-activated K⁺ channels, *Neuron* **41**, 351–365.
- Maret, W. (2000) The function of zinc metallothionein: A link between cellular zinc and redox state, *J. Nutr.* **130**, 1455S–1458S.
- Perkins, D. J. (1953) A study of the effect of amino acid structure on the stabilities of the complexes formed with metals of group II of the periodic classification, *Biochem. J.* **55**, 649–652.
- Prasad, A. S., and Oberleas, D. (1968) Zinc in human serum: evidence for an amino acid-bound fraction, *J. Lab. Clin. Med.* **70**, 1006.
- Finney, L. A. and O'Halloran, T. V. (2003) Transition metal speciation in the cell: insights from the chemistry of metal ion receptors, *Science* **300**, 931–936.
- Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C., and O'Halloran, T. V. (1999) Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase, *Science* **284**, 805–808.
- Frederickson, C. J., Kasarskis, E. J., Ringo, D., and Frederickson, R. E. (1987) A quinoline fluorescence method for visualizing and assaying histochemically reactive zinc (bouton zinc) in the brain, *J. Neurosci. Methods* **20**, 91–103.
- Budde, T., Minta, A., White, J. A., and Kay, A. R. (1997) Imaging free zinc in synaptic terminals in live hippocampal slices, *Neuroscience* **79**, 347–358.

28. Canzoniero, L. M. T., Sensi, S. L., and Choi, D. W. (1997) Measurement of intracellular free zinc in living neurons, *Neurobiol. Dis.* **4**, 275–279.
29. Gee, K. R., Zhou, Z. L., Ton-That, D., Sensi, S. L., and Weiss, J. H. (2002) Measuring zinc in living cells. A new generation of sensitive and selective fluorescent probes, *Cell Calcium* **31**, 245–251.
30. Chang, C. J., Jaworski, J., Nolan, E. M., Sheng, M., and Lippard, S. J. (2004) A tautomeric zinc sensor for ratiometric fluorescence imaging: application to nitric oxide-induced release of intracellular zinc, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1129–1134.
31. Walkup, G. K., Burdette, S. C., Lippard, S. J., and Tsien, R. Y. (2000) A new cell-permeable fluorescent probe for Zn^{2+} , *J. Am. Chem. Soc.* **122**, 5644–5645.
32. Burdette, S. C., Walkup, G. K., Spingler, B., Tsien, R. Y., and Lippard, S. J. (2001) Fluorescent Sensors for Zn^{2+} based on a fluorescein platform: synthesis, properties, and intracellular distribution, *J. Am. Chem. Soc.* **123**, 7831–7841.
33. Thompson, R. B., and Jones, E. R. (1993) Enzyme-based fiber optic zinc biosensor, *Anal. Chem.* **65**, 730–734.
34. Thompson, R. B., Maliwal, B. P., and Fierke, C. A. (1999) Selectivity and sensitivity of fluorescence lifetime-based metal ion biosensing using a carbonic anhydrase transducer, *Anal. Biochem.* **267**, 185–195.
35. Fierke, C. A., and Thompson, R. B. (2001) Fluorescence-based biosensing of zinc using carbonic anhydrase, *BioMetals* **14**, 205–222.
36. Thompson, R. B., Cramer, M. L., Bozym, R., and Fierke, C. A. (2002) Excitation ratiometric fluorescent biosensor for zinc ion at picomolar levels, *J. Biomed. Opt.* **7**, 555–560.
37. Hunt, J. A., and Fierke, C. A. (1997) Selection of carbonic anhydrase variants displayed on phage: aromatic residues in zinc binding site enhance metal affinity and equilibration kinetics, *J. Biol. Chem.* **272**, 20364–20372.
38. Thompson, R. B., and Patchan, M. W. (1995) Lifetime-based fluorescence energy transfer biosensing of zinc, *Anal. Biochem.* **227**, 123–128.
39. Thompson, R. B., Maliwal, B. P., and Fierke, C. A. (1998) Expanded dynamic range of free zinc ion determination by fluorescence anisotropy, *Anal. Chem.* **70**, 1749–1754.
40. Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., and Tsien, R. Y. (1997) Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin, *Nature* **388**, 882–887.
41. Schwarze, S. R., Ho, A., Vocero-Akbani, A., and Dowdy, S. F. (1999) *In vivo* protein transduction: Delivery of a biologically active protein into the mouse, *Science* **285**, 1569–1572.
42. Clark, H. A., Hoyer, M., Philbert, M. A., and Kopelman, R. (1999) Optical nanosensors for chemical analysis inside single living cells. 1. Fabrication, characterization, and methods for intracellular delivery of PEBBLE sensors, *Anal. Chem.* **71**, 4831–4836.
43. Webster, A., Compton, S. J., and Aylott, J. W. (2005) Optical calcium sensors: development of a generic method for their introduction to the cell using conjugated cell penetrating peptides, *The Analyst* **130**, 163–170.
44. Heyduk, E., and Heyduk, T. (2005) Nucleic acid-based fluorescence sensors for detecting proteins, *Anal. Chem.* **77**, 1147–1156.
45. Tsien, R. Y. (1989) Fluorescent probes of cell signaling, in *Annual Review of Neuroscience*, pp 227–253, Annual Reviews, Inc.: Palo Alto, CA.
46. Kiefer, L. L., Paterno, S. A., and Fierke, C. A. (1995) Hydrogen bond network in the metal binding site of carbonic anhydrase enhances zinc affinity and catalytic efficiency, *J. Am. Chem. Soc.* **117**, 6831–6837.
47. Nuccitelli, R. (1994) *A Practical Guide to the Study of Calcium in Living Cells*, Vol. 40, p 364, Academic Press: New York.
48. Henkens, R. W., and Sturtevant, J. M. (1968) The kinetics of the binding of $Zn(II)$ by apocarbonic anhydrase, *J. Am. Chem. Soc.* **90**, 2669–2676.
49. Hunt, J. B., Rhee, M. J., and Storm, C. B. (1977) A rapid and convenient preparation of apocarbonic anhydrase, *Anal. Biochem.* **79**, 614–617.
50. Thompson, R. B., Zeng, H. H., Loetz, M., and Fierke, C. (2000) Issues in enzyme-based metal ion biosensing in complex media, in *In-vitro Diagnostic Instrumentation* (Cohn, G. E., Ed.), pp 120–127, SPIE: San Jose, CA.
51. Pocker, Y., and Fong, C. T. O. (1983) Inactivation of bovine carbonic anhydrase by dipicolinate: Kinetic studies and mechanistic implications, *Biochemistry* **22**, 813–818.
52. Zalewski, P. D., Forbes, I. J., and Betts, W. H. (1993) Correlation of apoptosis with change in intracellular labile $Zn(II)$ using zinquin [(2-methyl-8-p-toluenesulphonamido-6-quinolyloxy)acetic acid], a new specific fluorescent probe for $Zn(II)$, *Biochem. J.* **296**, 403–408.
53. Zalewski, P., Forbes, I. J., and Giannakis, C. (1991) Physiological role for zinc in prevention of apoptosis (gene-directed death), *Biochem. Int.* **24**, 1093–1101.
54. Yao, P. M., and Tabas, I. (2000) Free cholesterol loading of macrophages induces apoptosis involving the Fas pathway, *J. Biol. Chem.* **275**, 23807–23813.
55. Li, Q. Z., and Cass, A. E. G. (1991) Periplasmic binding-protein based biosensors. 1. Preliminary study of maltose binding-protein as sensing element for maltose, *Biosens. Bioelectron.* **6**, 445–450.
56. Looger, L. L., Dwyer, M. A., Smith, J. J., and Hellinga, H. W. (2003) Computational design of receptor and sensor proteins with novel functions, *Nature* **423**, 185–190.
57. Krebs, J. F., and Fierke, C. A. (1993) Determinants of catalytic activity and stability of carbonic anhydrase II as revealed by random mutagenesis, *J. Biol. Chem.* **268**, 948–954.
58. Nair, S. K., Calderone, T. L., Christianson, D. W., and Fierke, C. A. (1991) Altering the mouth of a hydrophobic pocket: structure and kinetics of human carbonic anhydrase II mutants at Val-121, *J. Biol. Chem.* **266**, 17320–17325.
59. Thompson, R. B., Maliwal, B. P., Felliccia, V. L., Fierke, C. A., and McCall, K. (1998) Determination of picomolar concentrations of metal ions using fluorescence anisotropy: biosensing with a “reagentless” enzyme transducer, *Anal. Chem.* **70**, 4717–4723.
60. Thompson, R. B., Jr., W. O. W., Maliwal, B. P., Fierke, C. A., and Frederickson, C. J. (2000) Fluorescence microscopy of stimulated $Zn(II)$ release from organotypic cultures of mammalian hippocampus using a carbonic anhydrase-based biosensor system. *J. Neurosci. Methods* **96**, 35–45.
61. Price, P. J., and Brewer, G. J. (2001) Serum-free media for neural cell cultures, in *Protocols for Neural Cell Culture* (Federoff, S., and Richardson, A., Eds.), 3rd ed., pp 255–264, Humana Press: Totowa, NJ.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California-Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroiyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Read, Write, and Forget Science

We at *ACS Chemical Biology* strive to educate scientists at all levels. When we attend meetings, we work hard to talk to all of you whether you are a student, a postdoctoral fellow, or an independent researcher. In addition to current research, we often discuss other topics that are of interest such as science policy, funding, and education. Although research discussions predominated at the recent annual meetings of the American Chemical Society (ACS) and American Society for Biochemistry and Molecular Biology (ASBMB), the topic of education also permeated the sessions and hallway banter. It is clear the ACS, ASBMB, and many other societies are working hard to enhance science education, but they continue to face new challenges that threaten the pipeline of young scientists entering the field. The latest challenge comes from developments in K–12 education.

A recent *New York Times* article titled “Schools cut back subjects to push reading and math” (1) discussed a disturbing trend to narrow the K–12 curriculum, placing greater emphasis on reading and math at the expense of other topics such as science and history. This article interprets a recent report from the Center on Education Policy (<http://www.cep-dc.org/>), which has been studying federal, state, and local implementation of the No Child Left Behind Act (NCLB). It found that 71% of the nations 15 000 school districts are cutting back on subjects, including science, to make more time for reading and math. This practice was implemented because the 2002 federal law requires annual exams in only these two subject areas at the expense of encouraging young scientists, historians, artists, or musicians. In an effort to increase student understanding to a level that allows them to pass the standardized tests, some schools are going as far as requiring 2–3 hours of math or reading a day. For those who love these subjects, this is welcomed news—no more balancing chemical equations or dissecting frogs.

Do the benefits of this intensive schooling outweigh the benefits of a broader education? The Secretary of Education believes that this “return to basics” approach solidifies student understanding and allows low-performing schools the opportunity to meet the federal standards. Some schools include history and science reading in an effort to retain other subjects in the narrow curriculum.

Others in education, including some scientists, feel that this focused approach may rob future generations of essential scientific and historical knowledge that will subsequently impact their future prospects for jobs and higher education. In a recent *New York Times* letter to the Editor, Randi Weingarten, President of the United Federation of Teachers, commented on the impact of “return to basics” instruction on New York City students. A 2005 study of elementary school teachers in New York City showed that 85% of them spent less than 2 hours a week teaching science or social studies. Test scores did improve for their students, but the failure rate for eighth-grade social studies students increased from 62% in 2001–2002 to 81% in 2003–2004 (2). Data on scientific proficiency were not reported.

If the new narrowed curriculum is helping students catch up on their reading and math, what is happening to the students who are *not* struggling in these subjects? Are they allowed to skip the extra hours of reading and math to enhance their understanding of science and social studies? There is little information on this group so it is unclear how the

“return to basics” program will impact the gifted and talented students. The program could potentially limit student inquiry and stifle creativity.

This “return to basics” trend may also impact science policy and global competitiveness. The students in K–12 are our future lawmakers. If they lack a basic understanding of science, will we be able to rely on them to make sound decisions regarding fiscal policy for research programs or to provide proper regulatory oversight of scientific research? For many of us, the thought of someone who doesn’t understand the concept of a molecule, a cell, or a chemical reaction legislating policies on drugs and stem cells is disturbing. Furthermore, if students are not exposed to scientific innovation, will they know how to participate in a culture of scientific discovery? Even President Bush acknowledged in his State of the Union address that “we need to encourage children to take more math and science and to make sure those courses are rigorous enough to compete with other nations”. Through NCLB, the President is giving “... early help to students who struggle in math” but so far, little is being done at the federal level to enhance student understanding of science.

Although we at *ACS Chemical Biology* may not directly impact K–12 education, our future success will depend on the proper education of our students. These are our future readers, authors, reviewers, and editors. We urge you to speak out on these educational trends and support your society’s efforts to shape education policy to guarantee that our future scientists are not left behind.

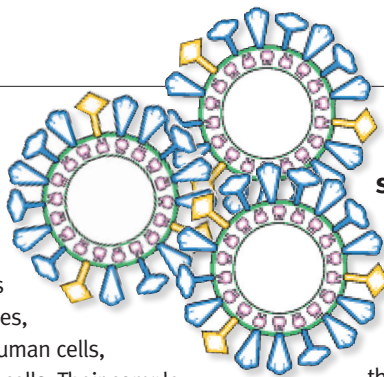


Evelyn Jabri
Executive Editor

1. Dillon, Sam (March 26, 2006) Schools cut back subjects to push reading and math, *New York Times*.
2. Weingarten, Randi (March 27, 2006) To the Editor: Test prep isn't education, *New York Times*.

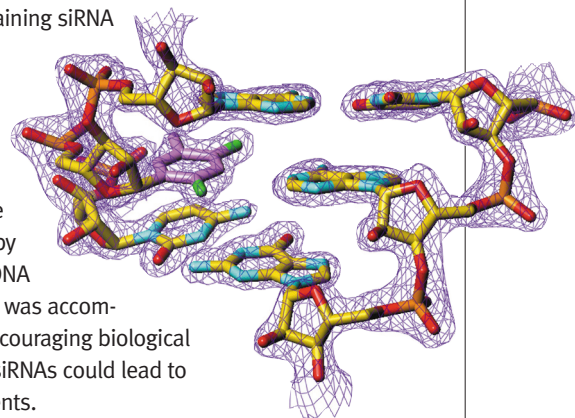
Sweeter Vaccines

Polysaccharides coat the surfaces of bacteria, viruses, parasites, and human cells, including cancer cells. Their complex structures are often specific to particular organisms or cell types, making them enticing candidates for vaccine development. However, carbohydrate-based vaccine development has been hindered by the notorious difficulty of carbohydrate synthesis and the lack of effective vaccine delivery platforms for polysaccharide antigens. Recent progress in polysaccharide synthesis has facilitated the creation of vaccine candidates, alleviating one hurdle. Now, based on recent success with the use of immunostimulating reconstituted influenza virosomes (IRIVs) for vaccine delivery of synthetic peptides, Liu *et al.* (p 161) have used IRIVs as an integrated carrier and adjuvant for the tetrasaccharide epitope on the lipophosphoglycans that decorate the surface of leishmania parasites. The researchers synthesized the tetrasaccharide in a linear fashion, incorporating a thiol-based linker for conjugation to either a phospholipid or carrier protein. The conjugates were formulated into IRIVs and used to immunize mice. Analysis by ELISA and immunofluorescence revealed that IgG antibodies were generated and that they cross-reacted with leishmania parasites. These results suggest that IRIVs have potential as an effective vaccine delivery platform for carbohydrate antigens and pave the way for the design of a wide range of carbohydrate-based vaccines.



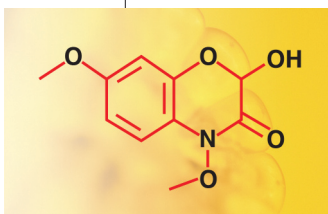
siRNA ...with Fluoride

Short interfering RNAs (siRNAs) are small duplex RNA strands that silence gene expression, offering a promising therapeutic strategy for a wide range of diseases. However, siRNAs suffer from chemical instability and susceptibility to endogenous nucleases, hampering their clinical potential. Chemically modified siRNAs could overcome these obstacles, provided they maintain gene silencing activity and are recognized by the RNA-induced silencing complex (RISC), a complex of enzymes that orchestrates the gene silencing process. Xia *et al.* (p 176) have synthesized and characterized siRNAs containing 2,4-difluorotoluyl nucleosides (rF). The researchers found that while rF-containing duplexes were less thermally stable than unmodified duplexes, they were more resistant to endonuclease cleavage in serum, indicating encouraging biostability properties. The crystal structure of an rF-containing siRNA duplex showed local geometric and thermodynamic changes, largely due to the loss of hydrogen bonds sustained in base pairs containing the rF nucleoside. However, activity studies revealed that rF-containing duplexes were capable of silencing the firefly luciferase gene, and analysis of the cleavage products by a modified 5'-RACE (rapid amplification of cDNA ends) technique suggested that the silencing was accomplished through the RISC mechanism. The encouraging biological properties demonstrated by these modified siRNAs could lead to development of siRNA-based therapeutic agents.



Clever Corn

Young maize seedlings exude an organic compound, 2-hydroxy-4,7-dimethoxybenzoxazin-3-one (HDMBOA), that has been implicated as an inhibitor of the plant pathogen *Agrobacterium tumefaciens* (*A. tumefaciens*). Upon sensing certain environmental signals, *A. tumefaciens* expresses the virulent genes VirA-VirG, which induce the transfer of tumorigenic DNA from the bacteria into plant cells. The wide variety of plants affected by *A. tumefaciens* and its unique ability to insert DNA into eukaryotic cells render antibiotics against the bacteria of interest to both the agricultural and medical communities.



Mareš *et al.* (p 165) have conducted structural studies of HDMBOA and revealed that, while HDMBOA is stable in the acidic pH environment of the soil, it rapidly decomposes to an unusual *o*-imidoquinone, (3*Z*)-2,2-dihydroxy-*N*-(4-methoxy-6-oxocyclohex-2,4-dienylidene)acetamide, at the neutral pH of the bacterial cytoplasm. Furthermore, activity studies suggest that the *o*-imidoquinone inhibits *vir* gene expression in *A. tumefaciens*. Because the *vir* genes function as the bacteria's sensing mechanism, this effectively masks the presence of the maize seedling to *A. tumefaciens*. The authors propose that HDMBOA cleverly functions as a "pro-drug" to deliver the *o*-imidoquinone to the bacterial cytoplasm. These findings redefine current understanding of this class of antibiotics and could facilitate the design of novel antibiotics for use in the agricultural and medical fields.

Published online April 21, 2006.

10.1021/cb6001575 CCC: \$33.50

© 2006 by American Chemical Society

Spotlight

Silencing the Monkey

RNA interference (RNAi) is a gene silencing technology that has become an invaluable tool for exploring biological processes and has tantalizing therapeutic potential for a variety of disease targets, especially those that have been challenging to pursue using conventional drug discovery methods. Apolipoprotein B (ApoB) is an essential protein in the regulation of cholesterol, but because it is a large, lipid-associated protein, it has been difficult to target for drug discovery. Zimmerman *et al.* (*Nature*, advance online publication, March 26, 2006, doi: 10.1038/nature04688) have demonstrated that systemic delivery of small interfering RNA (siRNA) silences ApoB in non-human primates and lowers cholesterol levels in clinically significant amounts.

While local delivery of siRNAs is well established, there are few reports of systemic delivery of siRNAs and none in non-rodent species. Systemic delivery of the APOB-specific siRNA, denoted siApoB-2, was accomplished by encapsulating siApoB-2 in stable nucleic acid lipid particles (SNALPs) and administering the formulation by intravenous injection to cynomolgus monkeys. To evaluate the effectiveness of this approach, the pharmacokinetics, efficacy, and safety of a single dose of SNALP-formulated siApoB-2 was assessed. A dose-dependent reduction in liver APOB mRNA levels of up to 90% was observed 48 h after treatment, and blood levels of ApoB-100 protein, total cholesterol, and LDL were reduced by up to 78%, 62%, and 82%, respectively. Remarkably, these effects lasted through 11 days post-treatment. No effects on complement activation, delayed coagulation, pro-inflammatory cytokine production, or changes in hematology parameters were observed, indicating that the treatment was well tolerated and non-toxic.

This ground-breaking study is the first to demonstrate the effectiveness of systemic delivery of RNAi-mediated silencing in non-human primates, substantiating a potential new strategy for treating cholesterol-related and perhaps numerous other diseases. **EG**

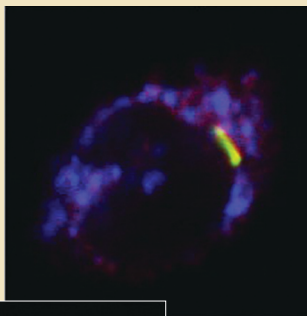


Image courtesy of Getty Images

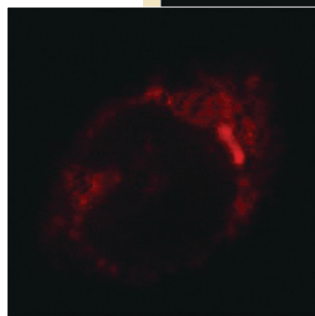
The Sun Takes Its Toll

Mammalian immune cells ward off infection by sensing environmental cues and switching on the proper genes for fighting off the infection. Toll-like receptors (TLRs) function early in the pathway by sensing microbial antigens. Many of the genes downstream of the TLRs have remained elusive in humans. A recent report by Liu *et al.* (*Science* 2006, 311, 1770–1773) demonstrates a new cog in the immune machine and an interesting link between sunlight, hormones, and immunity.

Using DNA microarrays, human genes were screened to find mRNAs that are upregulated when TLR receptors are activated with a ligand. Among the short list of candidates were the vitamin D receptor and Cyp27B1, an enzyme that converts the inactive vitamin D pro-hormone into the active form. This result suggested that vitamin D metabolism may bridge a gap between TLRs and the cell's ability to kill pathogens. Indeed, when cells infected with *Mycobacterium tuberculosis* were treated with active vitamin D, an antimicrobial peptide, cathelicidin, was switched on and bacterial counts decreased. Since Cyp27B1 converts pro-hormone to vitamin D, the authors reasoned that media levels of the pro-hormone may play a role in modulating the



Reprinted with permission from AAAS



immune response. They found that levels of the pro-hormone were vastly different between bovine and human sera. As a consequence, the TLR-coupled boost in mRNAs regulated by vitamin D was nearly absent in human cells cultured in bovine serum. The assay also revealed that the level of the pro-hormone in human serum from different donors is wide-ranging enough to show

dramatically different immune response. For example, serum from African-Americans, on average, had significantly less pro-hormone. Since the synthetic pathway for pro-hormone depends upon sunlight, darker skin indirectly casts a shadow on the TLR-dependent immune response. This may explain the higher rates of tuberculosis infection and progression among this ethnic group.

This study highlights one likely pathway among TLRs, vitamin D metabolism, and mRNAs that are expressed during an immune response. The striking pro-hormone data suggest that one's ability to produce vitamin D may contribute to susceptibility to microbial infection. **JU**

Published online April 21, 2006 • 10.1021/cb600147p CCC \$33.50

© 2006 by American Chemical Society

Botox on the Brain

Botulinum neurotoxin A (BoNT/A) is a key ingredient in the anti-wrinkle material Botox. BoNT/A is also prescribed for relief of severe pain and is a potential biological warfare agent. This toxin selectively binds to active neurons, and blocks exocytosis and neurotransmitter release. It has been proposed that BoNT/A receptors are composed of proteins and gangliosides, a group of glycosphingolipids found in the neuronal cell membrane. Now Dong *et al.* (*ScienceExpress*, published online March 16, 2006, doi:10.1126/science.1123654) show that BoNT/A enters nerve cells via the synaptic vesicle protein 2 (SV2).

The authors found that receptors for BoNT/A are localized to secretory vesicles and are exposed during exocytosis. The neurotoxin molecules interact with SV2, a conserved integral membrane glycoprotein with 12 putative transmembrane. BoNT/A bound to the largest luminal loop of three isoforms of SV2 but not to those of other synaptic vesicle proteins indicating that this neurotoxin-vesicle protein interaction was specific. Notably, glycosphingolipids bound the toxin

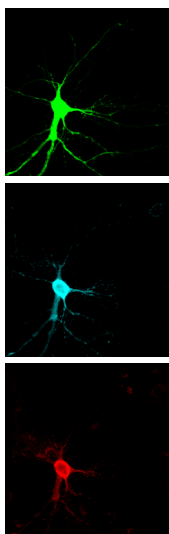


Image reprinted with permission from AAAS

with low affinity and promoted the formation of stable BoNT/A-SV2 complexes. These data indicate that both proteins and lipids are important for the selective entry of BoNT/A into neuronal cells.

Dong *et al.* also generated a loss of function cell model to determine if the cellular uptake of BoNT/A was dependent on SV2 proteins. The neurotoxin did not enter these SV2 knockout hippocampal neurons. However, entry could be restored by expressing three isoforms of the vesicle protein. Furthermore, mice lacking SV2B, one isoform of SV2, were less susceptible to BoNT/A. Collectively, these data show that SV2 is the receptor for BoNT/A.

Active neurons expose more SV2 receptors during exocytosis and neurotransmitter release. The authors propose that BoNT/A, by binding SV2 and shutting down exocytosis prevents the exposure of more receptors in the poisoned cell. This leaves toxin molecules free to enter other active neurons. This mechanism of action specifically and efficiently shuts down neuronal communication in the poisoned organism. **EJ**

Closing the CNS Gap

Central nervous system (CNS) injuries often result in tissue damage that does not heal properly. Several factors contribute to inadequate repair of CNS tissue, including scar tissue formation, gaps in nervous tissue from phagocytosis of dying cells, and failure of neurons to initiate axonal extension. Ellis-Behnke *et al.* (*PNAS* 2006, 103, 5054–5059) have now developed novel nanobiomedical materials, called self-assembling peptide nanofiber scaffolds (SAPNS), that promote regeneration of axons when injected at sites of CNS trauma.

SAPNS consist of alternating positive and negative L-amino acids that form highly hydrated scaffolds in physiological solutions. Use of SAPNS as biomaterials provides several advantages over currently available polymer biomaterials. First, the individual nanofibers are approximately 10 nm in diameter, which

is similar in scale to the native extracellular matrix and thus may facilitate cell growth, migration, and differentiation. In addition, degradation of SAPNS leaves natural L-amino acids that can potentially be used by the surrounding tissue. Finally, because SAPNS are synthetic materials and appear to be immunologically inert, biological contamination and tissue rejection issues prevalent in other biomaterials are avoided.

To test the ability of the SAPNS to promote CNS wound healing, the researchers generated a lesion in the optic tract of the hamster midbrain, resulting

in a tissue gap and loss of vision. When RADA-1, a SAPNS that is known to support neuronal growth and development, was injected at the site of injury, significant axonal growth through the tissue gap was observed and, remarkably, vision was restored. The authors hypothesize that interaction of SAPNS with the extracellular matrix promotes cell migration into the lesion area and creates an environment conducive to

axon regeneration. By overcoming some of the obstacles in CNS wound healing previously thought to be impenetrable, this groundbreaking work launches a new method to treat CNS trauma. **EG**

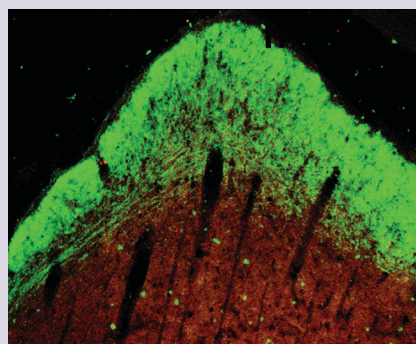


Image reprinted with permission from PNAS.

New Light on the Lightning Bug

Fireflies don't just fascinate six year olds. Scientists have shown that their bright yellow light allows the insects to communicate, whether to attract members of the opposite sex or to warn predators to stay away. The biochemistry behind the bioluminescence comes from oxidation of the small molecule luciferin by the enzyme luciferase; light is emitted upon relaxation of excited oxyluciferin to its ground state. Remarkably, a single amino acid change in luciferase, serine 286 to asparagine (S286N), changes the emission color from yellow-green to red, but the molecular basis for the color change has not been clear. Now, Nakatsu *et al.*

(*Nature* 2006, 440, 372–376) have uncovered a structural explanation for this fascinating phenomenon.

The researchers use X-ray crystallography to investigate structural differences between wild-type and mutant luciferase enzymes. Three crystal structures of wild-type luciferase, the first complexed with its natural reactant, the second with a stable analog of the high energy intermediate, 5'-O-[N-(dehydroLuciferyl)-sulfamoyl]adenosine (DLSA), and the third with its natural products, provided snapshots of the reaction pathway; changes in luciferase structure reflect movement of amino acids as the reaction progresses. Although the structures with the reactant and the products were essentially identical, the DLSA structure revealed significant movement of isoleucine 288, close to serine 286, creating a hydrophobic pocket in the active site. The authors suggest that the hydrophobic environment prevents energy loss from the excited state of oxyluciferin, resulting in yellow-green bioluminescence.

To investigate the basis for the color change, the structure of luciferase (S286N) complexed with DLSA was compared with that of the wild-type enzyme. The mutant enzyme does not exhibit the conformational changes seen in the wild-type system, and thus no hydrophobic pocket is created. The authors hypothesize that this results in energy loss from the excited state, manifested by lower energy red light emission. EG

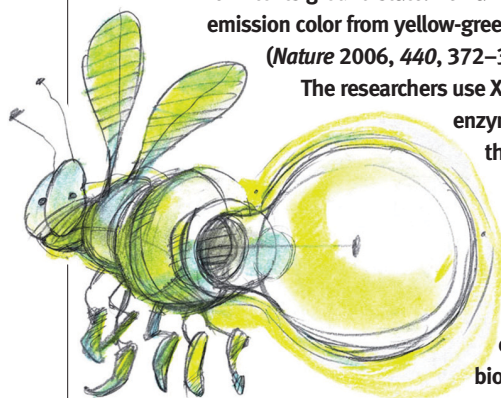


Image courtesy of Getty images

Fortuitous Promiscuity

Labeling reagents for visualizing biological interactions come in all shapes and sizes. Proteins, peptides, small molecules, and inorganic nanoparticles have all been used to label proteins, but each method must tackle many issues, including labeling specificity, speed, stability, tag size, and toxicity. In addition, it has been especially difficult to target specific proteins in the context of living cells. Lin *et al.* (*JACS* 2006, 128, 4542–4543) have developed a method for labeling proteins *in vitro* and on the surface of live cells using the enzyme transglutaminase.

Transglutaminases (TGases) catalyze amide bond formation between glutamine (Q) and lysine residues, serving to cross-link proteins in diverse cellular processes including

cell migration and apoptosis. The authors exploit this activity using guinea pig TGase (gpTGase), which is specific for its glutamine-containing substrate but can be quite promiscuous for its amine-containing substrate. This promiscuity allows for a variety of amine-containing labeling reagents, such as biotin or fluorescein derivatives, to be ligated to peptide substrates of gpTGase (called “Q-tags”) genetically fused to the target protein of interest.

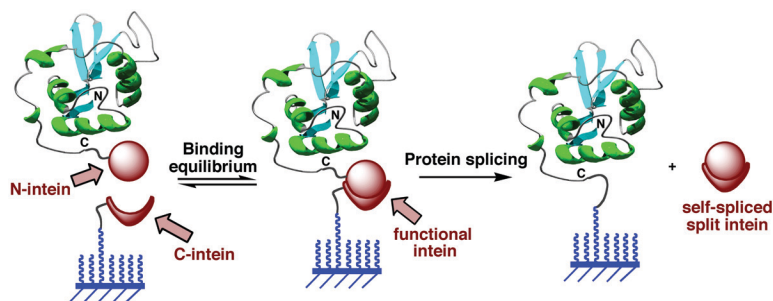
To demonstrate the utility of this approach in live cells, two proteins, cyan fluorescent protein and the epidermal growth factor receptor, were Q-tagged, expressed on the surface of HeLa cells, and labeled. The labeling was demonstrated to be site-specific, enzyme-dependent, and non-toxic and did not affect the

activity of the proteins. In addition, this approach was used to attach a photoaffinity probe to the NFκB transcription factor p50 *in vitro*. This allowed the researchers to examine the effects of reagents other than DNA on the homodimerization of p50, a distinct advantage over the currently favored gel-shift assay. Though there are disadvantages to using gpTGase, such as the inability to label intracellular proteins and the fact that gpTGase specificity for the Q-tag is lower than that of other enzymes (such as biotin ligase) for their peptide substrates, this work opens the door to using transglutaminase-introduced probes to investigate protein trafficking, protein–protein interactions, and protein conformational changes in the cellular context. EG

Protein Capture... Without a Trace

Protein microarrays are an attractive method for investigating protein interactions due to the small amount of protein required and the large number of interactions that can be probed on a single glass microscope slide. However, methods for attaching proteins to the microarray surface suffer from technological challenges. Non-covalent attachment methods leave proteins attached at random orientations or may not be resilient enough to last through subsequent assays, and covalent attachment strategies can be inefficient or leave capture reagents attached to the surface. Kwon *et al.* (*Angew. Chem., Int. Ed.* 2006, 45, 1726–1729) have developed a traceless capture ligand approach for creating protein microarrays.

The traceless capture method is based on protein trans-splicing, a naturally occurring process wherein a segment of a protein, called an intein, excises itself and rejoins the remaining portions. The authors use the DnaE intein from cyanobacteria,



Reprinted with permission from *Angewandte Chemie, International Edition*

which contains two fragments, the N-intein (I_N) and the C-intein (I_C), that spontaneously self-assemble to form a functional splicing domain. In the study, the I_N fragment was fused at the DNA level to the C-termini of two proteins, maltose binding protein and enhanced green fluorescent protein, and the I_C fragment was immobilized onto a glass slide. Spotting of the protein solutions on the glass slide should allow for the intein fragments to interact, ligate the protein to the slide, and splice themselves out into solution. Indeed, detection of the proteins was dependent on

the presence of I_N in the protein and I_C on the slide, suggesting that the proteins covalently attached to the surface by this mechanism. In addition, the authors demonstrated that I_N -containing fusion proteins from soluble cellular fractions of protein expressed in *E. coli* and from crude *in vitro* transcription/translation reactions could be immobilized using this approach, expanding its utility to proteins in biological mixtures. EG

Resisting the Resistance

Antibiotic resistance is a growing health concern across the globe. Many bacteria acquire resistance through genetic mutations that circumvent the antibiotic's mechanism of action. Antibiotics that possess multiple modes of action are therefore more likely to resist resistance. Nisin, a common food preservative and member of a group of antimicrobial peptides called lantibiotics, utilizes several mechanisms to kill bacteria, including forming pores in the membrane, disrupting cell wall biosynthesis, and inhibiting the outgrowth of bacterial spores. Nisin A contains five cyclic

thioethers, called lanthionines and methyllanthionines, that are created by addition of cysteine to dehydroalanine and dehydrobutyryne. It is believed that the enzyme NisC is responsible for formation of the rings, but the cyclase activity of NisC has not previously been demonstrated. Li *et al.* (*Science* 2006, 311, 1464–1467) report the *in vitro* enzymatic synthesis of nisin A and the crystal structure of NisC. Their findings confirm the cyclase activity of NisC and provide insight into the mechanism of

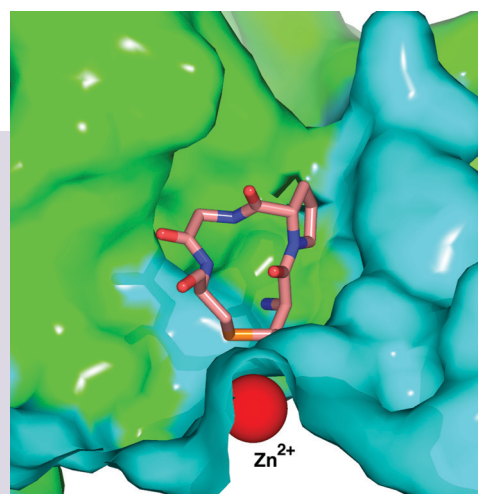


Image reprinted with permission from AAAS.

biosynthesis of this important antibiotic. Sequencing of the nisin gene cluster several years ago revealed that the

(continued on page 120)

Untranslated Roadblock

The sequences found at each end of a typical eukaryotic mRNA lack protein-coding potential and instead play important roles in gene regulation. Recently, these untranslated regions (UTRs) have garnered extra attention as more sequence motifs are implicated in



Image courtesy of Getty Images

RNA stability and the efficiency of translation. A new study by Babendure *et al.* (*RNA*, published online March 15, 2006, doi: 10.1261/rna.2309906) takes on an old UTR question: how does RNA secondary structure at the 5' end of an RNA affect the translation efficiency of the message?

The authors used a remarkably simple cell-based assay where

translation efficiency is monitored by a ratio of two different colors of fluorescent protein. Red fluorescent protein was encoded after a standard 5' UTR, while the UTR of green fluorescent protein (GFP) was flavored with various hairpins of RNA secondary structure. These RNA elements were engineered between the GFP mRNA transcription start site and the start codon for protein synthesis. Structural elements that reduced translation efficiency in living cells were identified by a change in the red to green signal ratio. Both the stability of the inserted RNA helix and its position could be easily altered and assayed. Hairpin RNAs at the extreme 5' end of the RNA had the most profound effect on translation. Also, a surprisingly sharp drop in translation occurred as the stability of the hairpin pushed into the $-25 \text{ kcal mol}^{-1}$ range. Some of the rules developed from this

assay stand in contrast to earlier studies with *in vitro* translation systems, thus highlighting the need to caution and comparison. These synthetic hairpin elements were compared with natural UTR sequences that have strong predicted secondary structures near the 5' end of the mRNA. These natural 5' UTRs also imparted lower translational efficiency.

This paper represents a rigorous next step toward understanding yet another RNA-mediated control mechanism. In addition to secondary structures, ligand-induced RNA tertiary structures such as those found in bacterial riboswitches may also regulate protein synthesis in eukaryotes. Finally, the authors suggest that small engineered RNAs might be used to alter the structural character of natural or designer UTRs, thereby providing a mechanism for cells to tinker with translation. **JU**

Resisting the Resistance, *continued*

peptide precursor to nisin A is synthesized via translation and then post-translationally modified to contain the cyclic thioethers. However, the *in vitro* reconstitution of nisin A, as well as that of other lantibiotics, has been challenging due to their structural complexity and the requirement for access to NisC and its substrate. To obtain NisC, the authors cloned the enzyme from *Lactococcus lactis* and expressed it in *Escherichia coli*. In order to generate the NisC substrate, an *L. lactis* strain was engineered to contain the genes that code for the

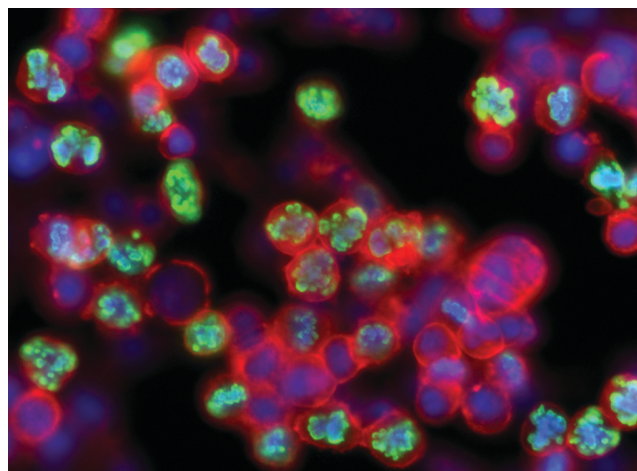
precursor peptide, the dehydratase that creates dehydroalanine and dehydrobutyrine, and the transporter that secretes the dehydrated peptide. Treatment of the secreted dehydrated peptide with NisC, followed by removal of the leader peptide with trypsin, generated active nisin A, as shown by a clear zone of inhibition when spotted on a lawn of an indicator strain.

The crystal structure of NisC was determined and found to contain an α -toroidal core structure that is topologically similar to bacterial cellulases, endoglucanase, terpenoid cyclases, and the β subunit of

protein farnesyl transferase. In addition, a zinc ion is strategically located in the center of the toroid bound by two conserved cysteines and a histidine, where it assists activation of the nisin cysteine thiol for intramolecular nucleophilic attack on the dehydroalanine or dehydrobutyrine. This structure provides insights into the possible roles, such as posttranslational modification of cysteine residues, of structurally similar proteins found in a variety of organisms including humans but whose functions are currently unknown. **EG**

Genome-Wide Interference

RNA interference (RNAi), where specific RNA fragments suppress gene expression and result in loss of function phenotypes, has revolutionized the exploration of gene function in mammals. In an effort to adapt RNAi technology to genome-wide exploration, The



Reprinted with permission from Cell

RNAi Consortium (TRC) has been created to generate genome-scale short hairpin RNA (shRNA) libraries and develop high throughput screening methods to characterize these libraries in mammalian cells. To this end, Moffat *et al.* (*Cell* 2006, 124, 1283–1298) have created the initial portion of a lentiviral shRNA library, denoted TRC1, and screened a subset of this library in human colon cancer cells.

TRC1 contains 100,000 arrayed shRNA constructs that target 22,000 human and mouse genes.

Several challenges in high throughput RNAi screening were addressed in the design and screening of TRC1. First, introduction of shRNA into cells was accomplished using lentiviral vectors, which can transduce a wide range of cell types including primary and nondividing cells. Second, variable effectiveness of different shRNA constructs and off-transcripts effects were limited by including five shRNA constructs per gene and requiring that at least two shRNAs induce a similar phenotype in order for the

effects of the gene to be considered. Finally, a semiautomated protocol was developed to adapt the library to high throughput phenotypic assays.

As a test of the utility of TRC1, a subset of the library was screened to identify regulators of mitosis. One

hundred genes were identified that significantly altered the mitotic index in human colon cancer cells. Further experiments revealed that although some of these genes also affected mitosis in human fibroblasts, others did not, pointing to potential cancer targets.

The TRC1 library is an initial realization of an effort to annotate the genome using RNAi. Information generated from this project will contribute to a global understanding of the genetics of biological processes. **EG**

UPCOMING CONFERENCES

First Annual Meeting

Japanese Society for Chemical Biology
May 8–9, 2006
Tokyo, Japan

Yale Chemical Biology Symposium

May 12, 2006
New Haven, CT

71st Symposium: Regulatory RNAs

May 31–June 5, 2006
Cold Spring Harbor Laboratory, NY

Spotlights written by Eva Gordon, Evelyn Jabri, and Jason Underwood

Nicholas R. Cozzarelli (1938–2006)

Carlos Bustamante*

Howard Hughes Medical Institute, and Departments of Molecular and Cell Biology, Physics, and Chemistry, University of California, Berkeley, Berkeley California 94720

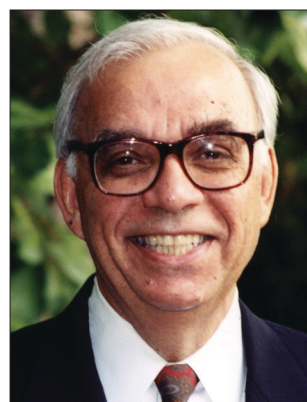
Nicholas R. Cozzarelli, Professor of Molecular and Cell Biology at the University of California at Berkeley, a star scientist of first magnitude, a colleague of uncompromising integrity, a great teacher, and warm friend, died on Sunday March 19, 2006, of complications associated with the treatment of lymphoma. Nick was born in New Jersey, the son of immigrants from southern Italy. His father was a shoemaker and his mother worked editing phone books. As a young man, Nick obtained a full scholarship to attend Princeton where he graduated with an A.B. Magna Cum Laude in 1960. He then entered Yale University with the intention of becoming a lawyer, but became instead interested in medical school, which he attended from 1960 to 1961. It was during this period that he realized that his true vocation was research and thus he applied and was admitted to graduate school at Harvard University where he obtained a Ph.D. in Biochemistry in 1966. Following graduate studies, Nick joined Arthur Kornberg's laboratory at Stanford, where he became engaged and passionately interested in DNA replication.

The discovery of topoisomerases in the early 1970s sparked a host of experiments to understand how these molecules controlled the topological state of DNA molecules. Soon it became evident that these important enzymes were involved in many other central cellular processes besides their function as regulators of the degree of DNA supercoiling. Topoisomerases facilitate the opening of the double helix during DNA replication and transcription, thus favoring the

translocation of these enzymes along the DNA. Moreover, these enzymes are also essential to allow the proper segregation of chromosomes prior to cell division.

Many ingenious mechanisms had initially been proposed for topoisomerases, but they proved to be incorrect. Nick and co-workers demonstrated instead that DNA was passed through enzyme-bridged breaks. This allowed these enzymes not just to change supercoiling, the activity which first defined topoisomerases, but also to disentangle and decatenate DNA. The Cozzarelli lab also showed that decatenation, not supercoiling or relaxation, was the vital function of topoisomerases III and IV. Cozzarelli and co-workers demonstrated that type II topoisomerases act, in a way, as Maxwellian demons, removing the topological links far past equilibrium, using the energy of ATP hydrolysis. They also uncovered the mechanism underlying this feat. Topoisomerase IV has yet additional specificity: the Cozzarelli lab showed that it removes (+), but not (–) supercoils, so that it can promote the movement of the replication fork without eliminating the (–) supercoils, which are essential for chromosome organization.

Early, during his initial involvement with topoisomerases, Nick Cozzarelli realized that the number and type of possible products of the topoisomerase catalyzed reactions were restricted by powerful mathematical theorems that dictate the topological transformations of closed space curves possessing connectedness, such as circular DNA. What followed was his elegant unification of biochemical and mathematical methods to determine the conformation



Nicholas R. Cozzarelli

*To whom correspondence should be addressed:
E-mail: carlos@alice.berkeley.edu.

Published online April 21, 2006
10.1021/cb600153e CCC \$33.50
© 2006 by American Chemical Society

and topology of higher-order DNA structures such as supercoils, catenanes, and knots. In this effort, Nick Cozzarelli realized that these mathematical restrictions (called topological invariants) made it possible also to infer the possible molecular mechanisms used by enzymes that catalyzed the modification of these DNA topological states. Using this approach, his group was able to suggest possible enzymatic mechanisms for these topology-acting catalysts.

The most recent set of contributions by Nick and co-workers concerns the global structure of chromosomes. This subject had deep roots in Nick's past studies. He and co-workers first found that free (–) supercoiling in bacteria plays an unexpected and vital role in DNA compaction. They showed that many DNA gyrase mutants failed not at initiation or elongation of replication, as was initially expected, but because decatenation of DNA by topoisomerase IV was blocked. They then showed that supercoiling increases the rate of decatenation by many orders of magnitude. They also studied the mechanism of DNA condensation by condensins and found again that loops of supercoiled DNA played a critical role in this process. In addition, Nick and co-workers showed that, in bacteria, topological domains are much smaller than that which was presumed to be the case over the past 30 years. These domains have been found to be, on average, only 10 kb in size. Moreover, the Cozzarelli laboratory discovered that the domain barriers are not DNA sequence-specific and are, in fact, quite transient. Nick and his students then went on to characterize, through genetic and biochemical screens, four proteins that are involved in maintaining the topologically closed domains of a bacterial chromosome.

Because of their importance in DNA metabolism, topoisomerases have become the target of numerous antibiotics and cancer treatment drugs. Here, Nick Cozzarelli's group played an important role in the deciphering of the mechanism of

inhibition of key antibiotics and anticancer agents, including quinolones such as ciprofloxacin and coumerins such as novobiocin, both inhibitors of DNA gyrase, an essential type II topoisomerase in Gram-positive and Gram-negative microorganisms. Today, about two-thirds of all cancer chemotherapy regimens contain a topoisomerase inhibitor. Thus, unlike the work of many scientists, Nick's academic endeavors were shown to have a direct impact in the treatment of diseases.

Although trained originally as a biologist and a biochemist, Nick Cozzarelli saw before many that scientific research is a playground where many disciplines converge and where understanding often requires bridging apparently disconnected realms of knowledge. He saw clearly the need to break the language and cultural barriers that still separated biologists from physicists and mathematicians in the last quarter of the twentieth century. Thus, because of his own frustrating experiences attempting to communicate with mathematicians, Nick formed in 1985 the Program in Mathematics and Molecular Biology (PMMB). This program got biologists, mathematicians, and physicists together for a number of years and has been the model for subsequent programs of science in which people try to reach across the scientific interface.

Nick Cozzarelli's leadership in science is widely recognized in the scientific community. Elected in 1989 to the National Academy of Sciences, Nick became Editor-in-Chief of the *Proceedings of the National Academy of Sciences* in 1995. In this position he transformed a respected publication with an arcane submission process into the premier journal for the publication of first rate scientific articles in biology and other disciplines. Nick Cozzarelli is broadly regarded as the best Editor of the *Proceedings*.

Nick's interest in fostering the development of science also led him to play

a number of important roles in various Academy committees. Thus, he was the chair of the National Academy of Sciences Section on Biochemistry, a member of the Academy's Committee on the Funding of Young Investigators, a member of the Committee on strengthening the linkages between science and the mathematical sciences, and the committee on research standards and practices to prevent the destructive application of biotechnology, among others. Finally, it is important to mention his leadership in the promotion of a freer access to the scientific literature. The journal that he edited, the *Proceedings of the National Academy of Sciences*, has taken a leadership role in this regard. He was also an important contributor to the establishment of PubMed Central and the Public Library of Science.

When I moved to Berkeley in 1998, Nick and I talked about a number of problems that were just then becoming accessible to analysis by the methods of single molecule manipulation that my students and I had been developing at the University of Oregon. What ensued were some of the most exciting years of my scientific career. Endowed with a keen sense of what is important, Nick was a mind constantly on the move, thinking of new ways of using these powerful new methods to study important problems in biochemistry. Working with Nick was one of the privileges that I cherished upon joining the faculty at Berkeley. Our meetings, coded by our students either as a "Cozzamante" or a "Bustarelli" meeting, to indicate whether they were to be held in Nick's or in my laboratory, respectively, led to some of the most exciting science and publications of my career. It was in this small gatherings that I learned to admire Nick Cozzarelli's scientific insight, his unmatched ability to ask the crucial questions about any problem, and his intellectual honesty which only accepts one adjective: uncompromising. There, in these often-heated

scientific discussions, we became good colleagues first and then good friends. He was a tough, critical evaluator of his own science and that of his colleagues. Our common Italian background was a cultural reference that we enjoyed sharing with one another. He enjoyed life, the good cuisines of the world, the good wine, art, music, and the theater. Having acquired a classical education, he never used it to show off but to make instead incisive, appropriate comments that benefited from the wisdom of the old culture.

His premature departure robs us from one of the true creative talents in quantitative experimental biology. Throughout his almost two-year battle against the disease, Nick Cozzarelli taught his friends and students one final lesson. Hemingway has written that the “true measure of courage is to display grace under pressure”. Those of us that worked close to him witnessed, often in disbelief, his dedication to work until the very end, never uttering a com-

plaint even when it was plainly obvious that he was in pain, uncomfortable, or suffering. He refused to stop a meeting or to interrupt an argument because of his condition. In one of the opportunities in which I drove him to Stanford for his chemotherapy session, our scientific discussion got so animated that we passed the exit for the Dumbarton Bridge by 20 miles, before we realized it! In fact, those of us that worked closely with him until the very end often had the impression that he was treating his own condition as an inconvenience, a nuisance of secondary importance that should not distract us from our tasks at hand. It was his last lesson, his lesson on courage. It was his way to remind us how much was still left to do in the lab, his way to tell us, true to his Latin background, “*ars longa, vita brevis*”. Today, I can only reflect in these so many lessons about science and life that I learned from Nick Cozzarelli. I will miss greatly my friend.

Discovering the Building Blocks of RNA Interference



Of all the tools biologists and chemists have in their kits, few have revolutionized these disciplines as much as RNA interference (RNAi). Also known as RNA silencing or gene silencing, this technique knocks down the function of targeted genes when double-stranded RNA called small interfering RNA (siRNA) enters cells. RNAi has been on the scene for just over a mere decade, but it has already transformed the power that researchers wield over genes. Rather than taking the costly and lengthy route of engineering knockout animals to understand a gene's function, investigators can now quickly and cheaply accomplish that feat by using RNAi to decrease any given gene's mRNA. Researchers are also gradually harnessing the benefits of dialing down the influence of genes involved in diseases. Even with its widespread use, scientists know little about the natural processes in cells that guide gene silencing. Phillip Zamore, a biochemist and developmental biologist at the University of Massachusetts Medical School in Worcester, MA, is intensively detailing RNAi's major players and determining how they orchestrate gene expression. These efforts are gradually elucidating the mechanistic aspects of RNAi.

Putting the Pieces Together

Zamore was born in 1963 in East Flatbush, a neighborhood in Brooklyn, NY. His father, who was a lawyer, and his mother, who is a speech pathologist, encouraged an early and enduring love for reading and writing. Another of Zamore's childhood passions was for the Lego building toy, an interest that still remains with him today. "My approach to biology is still the same approach I took to Lego as

a child—how the pieces go together and why they do or don't work when they go together in different ways," he says. "It's a very reductionist approach."

In grade school, Zamore indulged this exploratory attitude by spending nearly all his free time asking myriad questions of his school's science teacher. Later, in junior high school, he regularly exercised his curiosity in the natural world through science classes and after-school clubs. When he was 16, Zamore attended a summer program in Roswell Park, NY, geared toward nurturing young scientists. There, he did experiments to examine how glycosylation affects cultured human cells' response to interferon. It was the first time that Zamore had ever lived on his own away from home, and he reveled in his newfound freedom. "I associated being grown up with doing science," he says.

By the time he began his undergraduate studies at Harvard University (Cambridge, MA), a school he picked for its liberal arts offerings in addition to its storied reputation, Zamore says that he was certain that he wanted to be a scientist. Though Zamore's parents were supportive, they were often puzzled by his scientific bent. "Where I came from, nobody placed a high premium on science. It was important in life to read the classics of American literature, but knowing what a ribosome is was not something that was valued," he says. "My father struggled to read *Scientific American* to figure out why I was so interested."

Seeking to maintain a well-rounded background, Zamore completed his undergraduate studies in 1985 with a Bachelor of Arts degree. He then took a job as a technician in molecular biologist Michael Green's lab

Profiles provide insights into the lives, backgrounds, career paths, and futures of scientists who serve as Experts on ACS Chemical Biology's online Ask the Expert feature. Readers are encouraged to submit questions to the Experts at www.acschemicalbiology.org. The editors will post the most interesting exchanges on the website.

while his girlfriend, now his wife, completed her undergraduate degree. Green encouraged Zamore to pursue graduate studies in his lab at Harvard, so for the next six years, Zamore investigated the function of U2AF, a protein integral for RNA splicing (1). He completed his doctoral degree in 1992.

Later that year, Zamore accepted a position at developmental geneticist Ruth Lehmann's lab at Massachusetts Institute of Technology's (MIT) Whitehead Institute in Cambridge, MA. Zamore worked with Lehmann for the next two years to investigate the molecular mechanisms behind how pole plasm, a specialized cytoplasm, predisposes cells in *Drosophila*'s posterior end to develop into germ cell progenitors. Then, unexpectedly, Lehmann announced that she was moving her lab to New York University in Manhattan. Unable to follow Lehmann because of family obligations, Zamore continued his work in the lab of James Williamson, an MIT biophysicist. When Williamson accepted a position at The Scripps Research Institute in La Jolla, CA, Zamore again moved his project, this time to MIT biochemist David Bartel's lab.

"At that time, it seemed like a baroque arrangement to be accreting principle investigators as they moved. But in fact, it was the single most important professional event for me because it gave me three different perspectives on how to do science, each associated with different techniques," says Zamore. "Now, in my lab, I have several different approaches that I use."

Investigating RNAi

In 1999, Zamore accepted a position at the University of Massachusetts Medical School. He chose this job over others because of the school's collegial atmosphere, he says. "The idea here not only is that science is fun, but that it's more fun when you do it with people you like," Zamore notes.

During the last few months in his post-doctoral position, Zamore discovered that

he and a fellow postdoc, Tom Tuschl, shared a common interest in the emerging field of RNAi. During a group meeting Zamore presented a paper written by Andrew Fire, now at Stanford University (California) School of Medicine and University of Massachusetts Medical School's Craig Mello that gave the first details on RNAi as observed in the nematode, *Caenorhabditis elegans* (2). "Then Tom got up and said, 'What a coincidence, I was going to talk about RNAi too,'" Zamore says.

Tuschl's plans were to set up a system to study the newly discovered phenomenon *in vitro* by recapitulating RNAi in a cell extract. "We needed a Lego set for RNAi—a cell extract that reproduced the process in a test tube so we could take it apart," says Zamore. Since a recent paper had shown that RNAi could take place in *Drosophila* embryos, Zamore suggested using them for their model system. The two researchers partnered to create a cell extract system to recapitulate RNAi. The results from this collaborative project were published in 1999 (3).

"By the time I showed up at the University of Massachusetts, everyone knew I wasn't going to work on anything but this," he says. "RNA silencing was already my passion."

The phenomenon captured his interest for a multitude of reasons, he adds. Not only does RNAi seem to be a pivotal force for fighting viruses in plants and insects, but all eukaryotic cells use self-produced double-stranded RNA pieces, called microRNAs (miRNA), to regulate gene expression. miRNA appears to play a large role in steering development of all multicellular organisms, making sure that development proceeds normally. "If you want to understand how organisms develop complex body patterns with many different cell types that are highly specialized, you have to understand miRNAs," Zamore says, "especially if you want to understand how they do it in such a robust manner."

miRNA seem to play an important role during development by selectively turning down the expression of genes. miRNA also appear to dampen the expression throughout a cell's life of so-called junk DNA, which makes up about half of the mammalian genome. "It prevents 50 percent of the genome from monopolizing important machinery, from making proteins in a nonproductive way by silencing it. You don't have eukaryotic life without it," Zamore says.

Zamore's research continues to expand our mechanistic understanding of RNAi. In 2001, he and his colleagues published findings indicating that cells use adenosine triphosphate to fuel gene silencing (4). "It was one of the first indications we had that this wasn't a simple, passive process, but an active, ordered pathway of building an RNAi machine to do RNAi," Zamore says. The same study revealed that cells seem to have a quality control process that prevents energy from being wasted on constructing needless RNAi machinery. "To our surprise, siRNAs that lacked 5' phosphates didn't get incorporated into RNAi machinery," he adds. "Cells seem to ask a question: Are these double-stranded RNAs really for RNAi? If they're junk, then cells don't want to build the machinery."

In 2003, his group published an explanation for how the RNAi machinery chooses one strand of siRNA over the other to use as a template for knocking down gene expression (5). Their findings suggest that thermodynamic features embedded in one strand of siRNA, but not the other, mark it for entry into the RNA-induced silencing complex (RISC). "siRNAs seem to have a lot of information embedded in them," says Zamore. "We discovered very rapidly that we could predetermine which strand is incorporated into RISC." Similar thermodynamic information seems to accomplish the same selective process for miRNA, the paper added, chalking up another similarity between the two gene silencing pathways.

"It was fun because we had to think hard about what kinds of stereospecific modifications might affect how the strands interact with RISC."

In a more recent paper, published in 2005, Zamore's team investigated what happens immediately after the RNAi machinery decides which strand of the duplex siRNA to use and which to discard. The prevailing view has been that an unknown ATP-dependent helicase then unwound the two strands before the passenger strand was destroyed. However, Zamore and his colleagues propose that, instead, the core enzyme of RISC, Argonaute2, cleaves the soon-to-be-discarded passenger strand during RISC loading (6). The key to these new findings, Zamore notes, was his team's ability to create siRNA that contained a single, non-natural phosphodiester bond. "It was fun because we had to think hard about what kinds of stereospecific modifications might affect how the strands interact with RISC," he says.

Targeting disease

Zamore's RNAi work also extends outside the university. In 2000, he, Tuschl, and Bartel, along with Phil Sharp of MIT and Paul Schimmel of The Scripps Research Institute, founded Alnylam Pharmaceuticals, based in Cambridge, MA. Zamore currently serves on the company's scientific advisory board. Alnylam's mission is to find new ways to treat human diseases using RNAi. Zamore notes that gene silencing could offer a fresh opportunity to effectively treat many currently incurable ailments, ranging from cystic fibrosis to avian flu. "In theory, you can turn off nearly any gene with expression that leads to disease, whether it's a human or viral gene, a mutant gene or just a wild-type gene that's expressed at too high a level or in the wrong place or the wrong time," he says.

Recently, Zamore adds, Alnylam published findings demonstrating that an RNAi-derived therapeutic could knock down an endogenous, clinically relevant gene in primates (7). The newly developed drug targets a gene that codes for the protein

apolipoprotein B (apoB), which metabolizes cholesterol. When investigators delivered the therapeutic agent to cynomolgus monkeys, they found that it silenced apoB mRNA by as much as 90 percent, which in turn lowered circulating LDL cholesterol concentrations by more than 80 percent.

With every new experiment, Zamore and his colleagues find that new questions arise about the mechanism of RNAi and chart the course for further mechanistic studies on RNAi-mediated gene silencing. With such uncharted territory spreading out before him, Zamore adds that he expects to be steeped in RNAi for the long haul. "If it remains fun, that's all that really matters. It has to be fun, or I won't do the work," he says. "I think we'll be at it for awhile."

—Christen Brownlee, Science Writer

REFERENCES

1. Zamore, P. D., Patton, J. G., and Green, M. R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF, *Nature* 355, 609–614.
2. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature* 391, 806–811.
3. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999) Targeted mRNA degradation by double-stranded RNA *in vitro*, *Genes Dev* 13, 3191–3197.
4. Nykänen, A., Haley, B., and Zamore, P. D. (2001) ATP Requirements and small interfering RNA structure in the RNA interference pathway, *Cell* 107, 309–321.
5. Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. D. (2003) Asymmetry in the assembly of the RNAi enzyme complex, *Cell* 115, 199–208.
6. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P., and Zamore, P. D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into ago2-containing RNAi enzyme complexes, *Cell* 123, 607–620.
7. Zimmerman, T. S., Lee, A. C. H., Akinc, A., Bramlage, B., Bumcrot, D., Fedoruk, M. N., Harborth, J., Heyes, J. A., Jeffs, L. B., John, M., Judge, A. D., Lam, K., McClintock, K., Nechev, L. V., Palmer, L. R., Racie, T., Röhl, I., Seiffert, S., Shanmugam, S., Sood, V., Soutschek, J., Toudjarska, I., Wheat, A. J., Yaworski, E., Zedalis, W., Koteliansky, V., Manoharan, M., Vormlocher, H. P., and MacLachlan, I. (2006) RNA-mediated gene silencing in non-human primates, *Nature*, published online Mar 26, <http://dx.doi.org/10.1038/nature04688>.

The Stochastic Nature of Gene Expression Revealed at the Single-Molecule Level

Xin Chen and Paul S. Cremer*

Department of Chemistry, Texas A&M University, 3255 TAMU, College Station, Texas 77843

Every child is different. Even identical twins can be readily differentiated through subtle but noticeable differences in appearance and personality. Similarly, phenotypic differences can be observed for individual *Escherichia coli* cells, even if they have identical genomes. This diversity cannot be related back to inherent genetic variation. A major factor in such differences at the cellular level can be, however, related to variability in gene expression (1, 2), which is intrinsically stochastic when a low copy number of molecules is involved. Until very recently, most of our knowledge about gene expression has been gleaned through ensemble measurements where the underlying stochastic nature of the process can be easily masked by population averages. To fully understand stochastic events, it is necessary to study them at the single-molecule level. In the case of gene expression, this means following transcription, translation, and the production of proteins at the single-molecule level. In two recent publications, X. Sunney Xie and co-workers have followed single-protein expression events *in vivo* in an elegant set of studies (3, 4).

In the first study (3), a Tsr–Venus fusion protein was used as both a gene reporter and fluorescent signal. Venus (5), which is a type of yellow fluorescent protein, was fused with a membrane protein Tsr (Figure 1, panel a). The gene encoding Tsr–Venus was spliced into *E. coli* to replace the native *LacZ* gene. Fusion to Tsr positioned the fluorescent Venus at the cell surface, which significantly limited its

diffusion and improved signal sensitivity. *In vivo* single-molecule protein detection was achieved in real time by taking epifluorescence measurements every 3 min after applying a short photobleaching pulse (Figure 1, panel b). Each signal burst represented no more than a few Tsr–Venus molecules, and the peak heights were quantized, corresponding to the number of protein molecules that were present. Only nascently inserted proteins generated fluorescence signals, as photobleaching eliminated the response of previously observed molecules.

In the second study (4), enzymatic amplification was exploited to detect the *in vivo* production of the enzyme, β -galactosidase (β -gal). β -Gal was used to catalyze the hydrolysis of a synthetic substrate (FDG), which generates a fluorescent product (Figure 2, panel a). Single-molecule detection has long been demonstrated with commercial optical setups *in vitro* (6, 7); however, to apply this method *in vivo*, one has to face a practical challenge because fluorescein is continuously and efficiently pumped out of living cells and diffuses away. To circumvent this difficulty, an ingenious lab-on-a-chip method was employed to confine single cells inside enclosed micron-sized chambers in a microfluidic device (Figure 2, panel b). The chambers not only offered microscale confinement, but also parallelism, allowing multiple cells to be monitored simultaneously. The fluorescence from the chamber increased as the hydrolysis reaction proceeded. The detected reaction rate

ABSTRACT Two recent papers have monitored the real-time synthesis of proteins *in vivo* at the single-molecule level. The work was done by two separate methods: fluorescent protein labeling and enzymatic amplification. Statistical analysis of the data reveals the inherent stochastic nature of gene expression.

*To whom correspondence should be addressed.
E-mail: cremer@mail.chem.tamu.edu.

Published online April 21, 2006
10.1021/cb600129z CCC: \$33.50
© 2006 by American Chemical Society

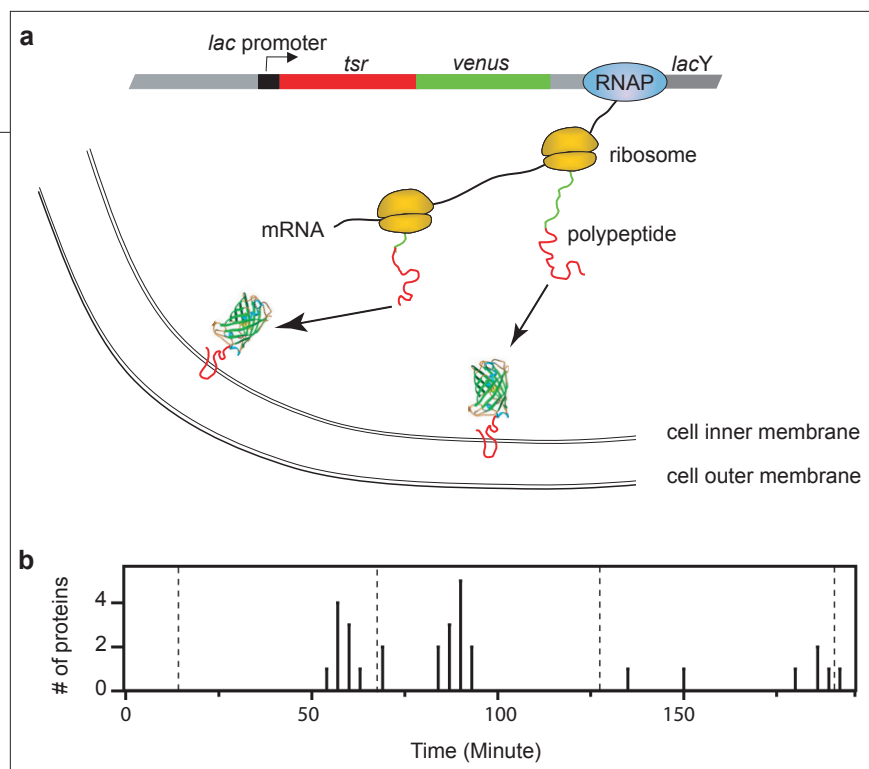


Figure 1. Gene expression of a green fluorescent protein detected at the single-molecule level. a) The gene encoding Tsr–Venus is expressed by a standard transcription and translation process. The nascently formed protein is then inserted into the inner membrane of *E. coli*. Venus was chosen over more traditional species such as green fluorescence protein because of its fast maturation time, ~4 min. Since the maturation is probably a stochastic event itself, the time resolution is inherently limited by the maturation event (5). b) Mature Tsr–Venus molecules are detected as individual burst events by fluorescence microscopy. The fluorescence signal was obtained every 3 min after photobleaching previously inserted Tsr–Venus molecules. The duration of this experiment was limited by the cells’ resistance to photodamage under these conditions. The vertical axis represents the number of proteins synthesized in a 3-min time period, and the vertical dashed lines mark cell division events. Adapted from ref 3.

was quantized, since the number of β -gal proteins present in the chamber was an integer number. The steps were therefore due to nascently synthesized β -gal, and the height of each step was proportional to the number of proteins that were made (Figure 2, panel c).

The reporter proteins in the two studies, Tsr–Venus and β -gal, were both expressed under highly repressed conditions and were monitored in real time with a resolution on the scale of minutes. The most important result in both experiments was bursts in protein production, which demonstrated that gene expression is an occasional event and that a few proteins are produced nearly simultaneously by such events, consistent with theoretical predictions (8, 9). Two key

parameters to describe such behavior, the burst size and frequency, correspond to how productive a single expression event is and how often such events occur. These data can be easily determined using Xie’s methods. For the Tsr–Venus assay, each event produced about 4.2 proteins on average and occurred about 1.2 times per cell cycle. For β -gal, the burst sizes were larger (7.8 proteins/event), but less frequent (0.16 events per cell cycle) in the *E. coli* cells employed. Other types of cells, such as *Saccharomyces cerevisiae* (yKT32) and embryonic mouse stem cells, also showed similar characteristics, albeit with different burst sizes and frequencies (4).

In addition to the average burst size and frequency, even richer information can be

extracted from a statistical treatment of the temporal evolution profile of burst events. For example, the burst size was not uniform but varied from burst to burst, representing the fluctuation in productivity of a single expression event as well as reflecting its stochastic nature. The distributions measured in the two studies were well-fit by exponential and geometric distributions, respectively. Both distributions are simple statistical functions, which assume total randomness in event occurrence. This, in turn, suggested that the productivity of an expression event fluctuated randomly. Such a finding leads to a very simple, yet fundamental, question. Which step(s) in gene expression can account for the randomness of the process? For the Tsr–Venus project (3), the authors addressed this question by measuring the copy number of mRNAs encoding Tsr–Venus. The average number was about 1.0 copy per cell cycle, which matched the burst frequency within experimental error. This strongly suggested that there was only a single mRNA copy for each expression event, at least under the experimental conditions explored. Therefore, the measured fluctuation was not likely coming from transcription. Instead, the authors suggested that it probably can be attributed to the fluctuating number of ribosomal binding events of the mRNA. However, any step after transcription may contribute to the fluctuation. This would also include protein folding, incorporation into the membrane, and maturation of the Venus probe.

Additional information might be gleaned from the temporal evolution profile by looking at the exact timing of a burst event within the cell cycle or the correlation of burst events with one another. Although unlikely to be generally true, a completely random burst distribution would imply that the probability of expression for a particular protein is not affected by extrinsic parameters and that the particular stage of the cell cycle is not the deciding factor. As other protein expression

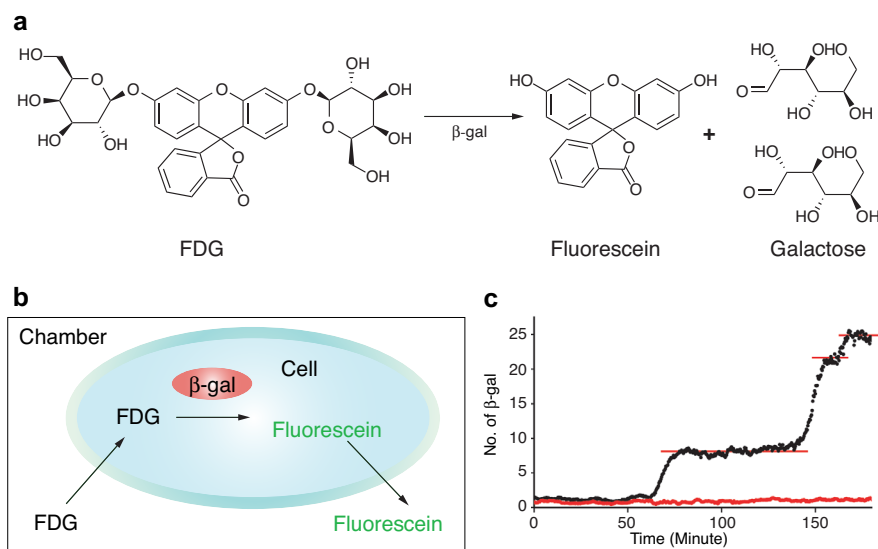


Figure 2. Gene expression of β -gal detected at the single-molecule level through enzymatic amplification. **a**) The enzyme β -gal cleaves FDG to create fluorescein and two galactose molecules. **b**) Schematic diagram of a single-cell, single-chamber apparatus for β -gal measurements. Each cell was treated with β -lactam antibiotics to increase the permeability of its membrane. This treatment enabled the facile diffusion of FDG into the cell, which would otherwise have been greatly attenuated. **c**) The reaction rate (number of proteins expressed) vs time plot shows discrete production events. The height of each step corresponds to the number of nascently synthesized β -gal molecules. Adapted from ref 4.

events are explored by this methodology, it will remain to be seen if such correlative behavior can be found or not.

Steady-state population analysis can also be conducted, since many cells/chambers can be monitored simultaneously. The protein distribution over a population of cells depends not only on burst size and frequency, but also on protein partitioning during cell division and the correlation between protein expression and cell division. In the β -gal study (4), a gamma-function distribution should be followed assuming equipartitioning between the two daughter cells and no other correlations, as was the case.

The studies reviewed here are among the first single-molecule gene expression experiments (10, 11) providing statistical information on stochastic gene expression events, which is very difficult, if not impossible, to obtain through ensemble

averages. They also point to a very promising future in this subfield, as both methods can be extended, in principle, to multireporter systems. For example, one might genetically label one gene with green fluorescent protein and another one with yellow fluorescent protein, in the same cell (12, 13). The advantage of two reporters at the single-molecule level is not merely parallelism, but rather the rich information carried in temporal pair-correlations between expression events for different genes. It is reasonable to hypothesize that expression of two genes coding for two proteins with cooperative functions might have a strong positive correlation, while two independent genes might be weakly correlated. Combining the two assays developed by the Xie laboratory may also be uniquely informative. Yellow (or green) fluorescent protein labeling essentially monitors the existence of a protein, and enzymatic

amplification monitors its activity. The difference between these two types of assays, therefore, could potentially differentiate between production and activation. This would be intriguing, as activation processes such as post-translation modification should be stochastic as well.

REFERENCES

- Novick, A., and Weiner, M. (1957) Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci. U.S.A.* 43, 553–566.
- Maloney, P. C., and Rotman, B. (1973) Distribution of suboptimally induced beta-d-galactosidase in *Escherichia coli*—enzyme content of individual cells. *J. Mol. Biol.* 73, 77–91.
- Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X. S. (2006) Stochastic gene expression in live cells, one protein molecule at a time. *Science* 311, 1600–1603.
- Cai, L., Friedman, N., and Xie, X. S. (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature* 440, 358–362.
- Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 20, 87–90.
- Rotman, B. (1961) Measurement of activity of single molecules of beta-d-galactosidase. *Proc. Natl. Acad. Sci. U.S.A.* 47, 1981–1991.
- Rondelez, Y., Tresset, G., Tabata, K. V., Arata, H., Fujita, H., Takeuchi, S., and Noji, H. (2005) Microfabricated arrays of femtoliter chambers allow single molecule enzymology. *Nat. Biotechnol.* 23, 361–365.
- Berg, O. G. (1978) Model for statistical fluctuations of protein numbers in a microbial-population. *J. Theor. Biol.* 71, 587–603.
- Rigney, D. R. (1979) Note on the kinetics and stochastics of induced protein-synthesis as influenced by various models for messenger-RNA degradation. *J. Theor. Biol.* 79, 247–257.
- Levsky, J. M., Shenoy, S. M., Pezo, R. C., and Singer, R. H. (2002) Single-cell gene expression profiling. *Science* 297, 836–840.
- Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* 123, 1025–1036.
- Raser, J. M., and O'Shea, E. K. (2004) Control of stochasticity in eukaryotic gene expression. *Science* 304, 1811–1814.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002) Stochastic gene expression in a single cell. *Science* 297, 1183–1186.

Micromanagement: A Role for MicroRNAs in mRNA Stability

Sarah F. Roush[†] and Frank J. Slack^{‡,*}

[†]Department of Molecular Biophysics and Biochemistry and [‡]Department of Molecular, Cellular and Developmental Biology, Yale University, P.O. Box 208103, New Haven, Connecticut 06520-8103

ABSTRACT Small, inhibitory RNA molecules called microRNAs cause large decreases in target protein levels through a post-transcriptional mechanism. Until recently, it was believed this mechanism operated almost exclusively at a step in translation. However, new work has revealed that microRNAs have a second, post-transcriptional mechanism that accelerates the rate of deadenylation, the initial step of mRNA decay.

One of the most sought after answers in microRNA (miRNA) research is the mechanism by which these small, inhibitory molecules post-transcriptionally regulate their mRNA targets (for review, see ref 1). MiRNAs are ~22 nt, non-coding RNAs that are present in most eukaryotes and have the potential to post-transcriptionally regulate a large number of genes (2, 3). In animals, miRNAs bind to partially complementary sites in the 3' UTRs of their target mRNAs, leading to significant decrease in the targets' protein levels (4, 5). Early research indicated that miRNAs regulate their targets through translational inhibition without lowering mRNA levels (4, 6). The point in translation in which the regulation occurs is currently being debated, but appears to happen before the completion of the polypeptide chain (6–12). More recently, the idea that miRNAs function only through translational inhibition has been challenged by studies indicating that miRNAs can cause decreases in target mRNA levels (12–14).

Two recent studies indicate that miRNA binding triggers target mRNA degradation through an increased rate of deadenylation from the poly(A) tail, normally the rate-limiting step in mRNA degradation (12, 14). In a report by Giraldez *et al.* (14), a miRNA expressed at the beginning of zebrafish zygotic transcription, miR-430, was shown to facilitate the clearing of many maternal mRNAs. These target mRNAs exhibited rapid deadenylation mediated by miR-430 and its target sites in the 3' UTRs of the mRNAs. The deadenylation was not a result of blocked

transcription, as an antisense morpholino oligonucleotide hybridized to the translational start site of a GFP reporter did not cause as much mRNA deadenylation and decay in the absence of the miRNA as in the presence of the miRNA. The fact that the miRNA could still cause accelerated deadenylation while translation was inhibited indicated that the miRNA did not need translation to mediate the poly(A) tail removal. In a second report by Wu *et al.* (12), mammalian cell culture showed similar results to that of Giraldez *et al.* (14). However, the study conducted by Wu *et al.* (12) not only demonstrates rapid deadenylation and decay of the target mRNA but also provides evidence for the intriguing idea that miRNAs can use two independent mechanisms of post-transcriptional regulation: translational inhibition and mRNA degradation through accelerated deadenylation (Figure 1).

To begin their work, Wu *et al.* (12) utilized an established promoter-reporter system to measure mRNA decay (15). The basic reporter contained an inducible *c-fos* promoter and β -globin gene (BG). Two copies of the miR-125b binding site, referred to as miRNA response element 1 (mRE1), from the 3' UTR of miR-125b's target mRNA, *lin-28* (16), were inserted into the 3' untranslated region (UTR) of the basic reporter mRNA. This modified reporter (BG+2E1) and the miR-125b gene were transiently transfected into 293T human embryonic kidney cells, where it has previously been shown that miR-125b can down-regulate a luciferase reporter mRNA containing several copies of miR-125b mRE 1 or 2 (16). In the present

*To whom correspondence should be addressed.
E-mail: frank.slack@yale.edu.

Published online April 21, 2006
10.1021/cb600138j CCC: \$33.50
© 2006 by American Chemical Society

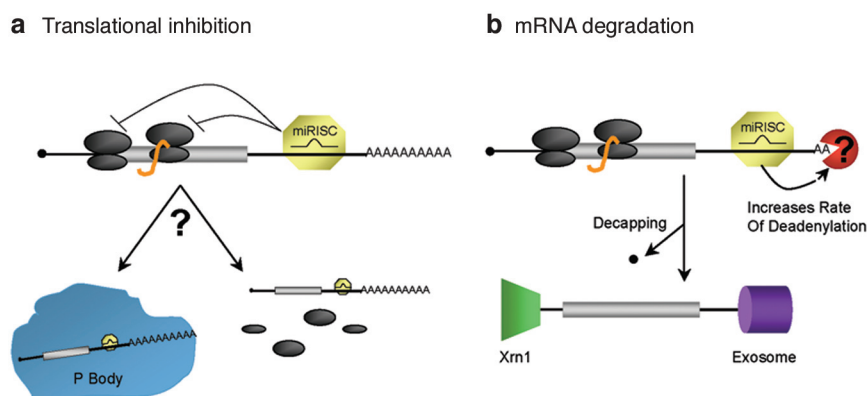


Figure 1. Two mechanisms of miRNA post-transcriptional regulation. **A** miRNA induced silencing complex (miRISC) loaded with its miRNA binds with imperfect complementarity to the 3' UTR of its target where it performs one or both of the following roles: **a)** Translational inhibition. Recent studies show that parts of the miRNA repression complex are present in processing bodies (P body) (20). Generally, P bodies are areas of mRNA degradation that lack translational machinery (21, 22); however, it seems that miRNA target mRNAs do not appear to be degraded at the same rate as other mRNAs in the P bodies (11). Another study suggests that miRNAs cause ribosomes to fall off the target mRNA during translation (10). **b)** mRNA degradation. As described in Wu *et al.* (12), miRNAs lead to accelerated removal of the poly(A) tail and, as a consequence, lead to cap removal followed by target mRNA degradation by the nuclease Xrn1 and the exosome complex. The red circle represents an associated deadenylase whose identity is unknown.

study, the presence of miR-125b led to an acceleration of reporter mRNA poly(A) tail deadenylation and decay. RT-PCR revealed that the degradation did not occur through a siRNA-like endonucleolytic cleavage within the mRE1 site, as occurred in a reporter containing a synthetic, perfectly complementary miR-125b binding site. Using susceptibility to a 5' exonuclease as an assay, it was then shown that deadenylation preceded removal of the 5' cap. In similar experiments, when the *let-7a* gene and a luciferase reporter containing *lin-28* mRNA's *let-7* binding elements (L7) were transiently transfected into 293T cells, there were also accelerated reporter mRNA deadenylation and decay.

To identify other possible miR-125b targets, a mouse genome microarray was performed using P19 cells. Undifferentiated P19 cells, which produce miR-125b when they differentiate into neurons, were mock-transfected or transfected with chemically synthesized miR-125b. Cytoplasmic RNA from these cells was used to probe the

microarray. Twenty-two putative targets of miR-125b were identified, all containing at least one possible miR-125b binding element. Possible miR-125b binding elements were selected from the 3' UTRs of two putative targets, Ajuba and MAPK kinase 7, and inserted into reporter vectors. These elements were found to accelerate the rate of reporter mRNA deadenylation and decay in the same mRNA decay and deadenylation assays used for mRE1 and L7.

The most intriguing experiments indicate that miRNAs may have two independent mechanisms by which they can down-regulate target expression. The first set of experiments demonstrated that a loss of translation was not what led to the accelerated target poly(A) tail loss and mRNA decay observed in the earlier miR-125b and BG+2E1 experiments. A large, 40 nt stem-loop was added into the 5' UTR of BG+2E1, which eliminated translation by specifically blocking translation initiation. Blocking translation in the absence of

miR-125b caused a small increase in the target mRNA's deadenylation and decay rates. However, rapid rates of deadenylation and decay were observed with both a block in translation and the presence of miR-125b. Importantly, the rapid rates of the mRNA decay were close to the rates observed when miR-125b was present and translation was not blocked, thus, indicating translation was not necessary for miR-125b's ability to accelerate deadenylation (Figure 2, panel b).

Reciprocal experiments showed that translational repression by miR-125b was not dependent on the presence of a poly(A) tail. These experiments used a luciferase reporter containing multiple mRE1 sites in its 3' UTR and a poly(A) tail or a histone-derived stem-loop in place of the poly(A) tail (Figure 2, panel c). Histone stem-loops mediate translation by using the stem-loop binding protein (SLBP) instead of a poly(A) tail and poly(A)-binding protein (PABP) (17). When these reporters were transfected with the miR-125b gene, no decrease in mRNA levels was seen, compared to the polyadenylated reporter, which had a drop in mRNA levels, as expected. However, a key observation was made upon calculation of the actual levels of luciferase protein, which showed that the level of translational inhibition by the miRNA was the same for both reporters. Therefore, even though removal of the poly(A) tail and the subsequent drop in mRNA levels was an important component of miR-125b regulation through mRE1, a large portion of the regulation was also due to translational repression that was independent of the poly(A) tail. Together, both of these regulatory mechanisms produced an even greater effect than either alone could do (Figure 2, panel d).

The idea that miRNAs may have two distinct functional mechanisms helps to reconcile previous results where evidence from different studies pointed to different types of post-transcriptional regulation, such as in the cases of *lin-4* and *let-7* (6, 13, 18).

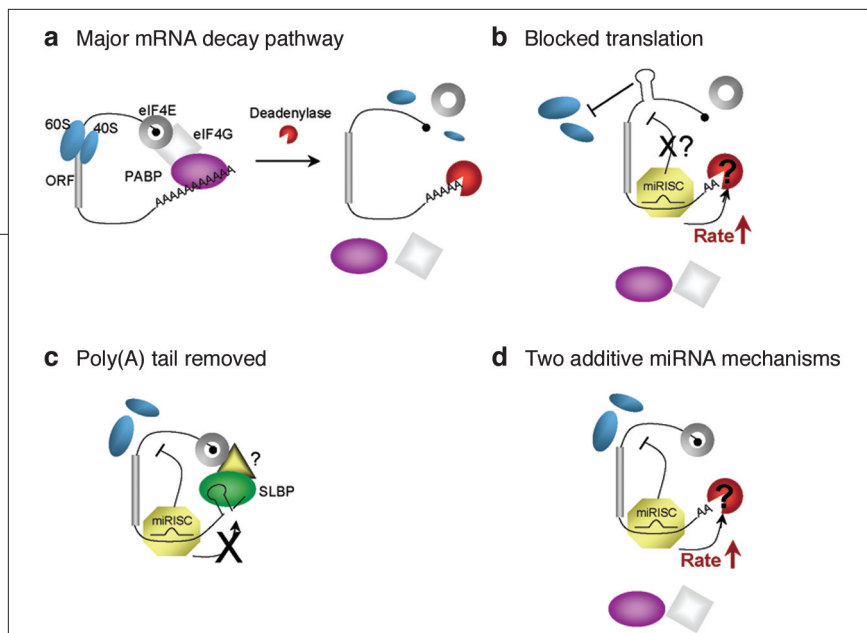


Figure 2. Additive effects of miR-125b repression. a) Major pathway of mRNA decay. The pathway begins with removal of the poly(A) tail and releasing of poly(A)-binding protein (PABP) followed by dissociation of the cap-binding proteins, eIF4E and eIF4G, which prevents further translation and leads to mRNA decay (19). b) Blocked translation. Blocking translation initiation by inserting a strong stem-loop structure reveals that the accelerated rate of deadenylation is not due to impaired translation, and hypothetically prevents any repression through translational inhibition. c) Poly(A) tail removed. Replacing the poly(A) tail with a histone-derived stem-loop prevents miRNA-mediated deadenylation and mRNA decay but does not prevent repression by translational inhibition. The triangle represents a hypothetical interacting partner. d) Two additive miRNA mechanisms. The two mechanisms can act additively to enhance repression of the target.

However, it also makes defining the exact nature of how miRNAs inhibit translation more complex. Recently, other reports have suggested that miRNAs may inhibit translation at initiation (10, 11) and/or elongation (10). If so, it could be expected that translational inhibition could be due to the loss of PABP binding as would result from the replacement of the poly(A) tail with a histone stem-loop (19). PABP binds to the poly(A) tail and interacts with the 5' cap binding protein eIF4G, resulting in mRNA circularization and aiding translation initiation (Figure 2, panel a) (19). However, the results of Wu *et al.* (12) indicate that miRNA translational inhibition is independent of the poly(A) tail and PABP loss or at a step when PABP and SLBP function similarly (12).

Continued research into miRNA cap and poly(A) tail dependence and their associated proteins, along with the development of novel inhibitors for specific translational states, such as hippuristanol (10), will further elucidate possible miRNA mechanisms of action. Perhaps when the two mechanisms of miRNA are completely separated and understood, synthesized

miRNAs can be produced to specifically inhibit translation initiation or deadenylation. Conceivably, miRNAs could use multiple combinations of additional, as yet unknown, mechanisms to regulate different targets. Though, in the end, it may be that the mechanism of every miRNA is fine-tuned for its specific target(s).

REFERENCES

- Bartel, D. P. (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function, *Cell* 116, 281–297.
- Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005) Combinatorial microRNA target predictions, *Nat. Genet.* 37, 495–500.
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003) Prediction of mammalian microRNA targets, *Cell* 115, 787–798.
- Wightman, B., Ha, I., and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*, *Cell* 75, 855–862.
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougiev, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*, *Nature* 403, 901–906.
- Olsen, P. H., and Ambros, V. (1999) The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation, *Dev. Biol.* 216, 671–680.

- Humphreys, D. T., Westman, B. J., Martin, D. I., and Preiss, T. (2005) MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function, *Proc. Natl. Acad. Sci. U.S.A.* 102, 16961–16966.
- Kim, J., Krichevsky, A., Grad, Y., Hayes, G. D., Kosik, K. S., Church, G. M., and Ruvkun, G. (2004) Identification of many microRNAs that copurify with polyribosomes in mammalian neurons, *Proc. Natl. Acad. Sci. U.S.A.* 101, 360–365.
- Nelson, P. T., Hatzigeorgiou, A. G., and Mourelatos, Z. (2004) miRNP:mRNA association in polyribosomes in a human neuronal cell line, *RNA* 10, 387–394.
- Petersen, C. P., Bordeleau, M. E., Pelletier, J., and Sharp, P. A. (2006) Short RNAs repress translation after initiation in mammalian cells, *Mol. Cell* 21, 533–542.
- Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005) Inhibition of translational initiation by *let-7* MicroRNA in human cells, *Science* 309, 1573–1576.
- Wu, L., Fan, J., and Belasco, J. G. (2006) From the cover: MicroRNAs direct rapid deadenylation of mRNA, *Proc. Natl. Acad. Sci. U.S.A.* 103, 4034–4039.
- Bagga, S., Bracht, J., Hunter, S., Massierer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005) Regulation by *let-7* and *lin-4* miRNAs results in target mRNA degradation, *Cell* 122, 553–563.
- Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Van Dongen, S., Inoue, K., Enright, A. J., and Schier, A. F. (2006) Zebrafish miR-430 promotes deadenylation and clearance of maternal mRNAs, *Science* 312, 75–79.
- Shyu, A. B., Greenberg, M. E., and Belasco, J. G. (1989) The *c-fos* transcript is targeted for rapid decay by two distinct mRNA degradation pathways, *Genes Dev.* 3, 60–72.
- Wu, L., and Belasco, J. G. (2005) Micro-RNA regulation of the mammalian *lin-28* gene during neuronal differentiation of embryonal carcinoma cells, *Mol. Cell. Biol.* 25, 9198–9208.
- Gorgoni, B., Andrews, S., Schaller, A., Schumperli, D., Gray, N. K., and Muller, B. (2005) The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway, *RNA* 11, 1030–1042.
- Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000) RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals, *Cell* 101, 25–33.
- Wilusz, C. J., Wormington, M., and Peltz, S. W. (2001) The cap-to-tail guide to mRNA turnover, *Nat. Rev. Mol. Cell Biol.* 2, 237–246.
- Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., and Parker, R. (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies, *Nat. Cell Biol.* 7, 719–723.
- Sheth, U., and Parker, R. (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies, *Science* 300, 805–808.
- Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M., and Parker, R. (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs, *RNA* 11, 371–382.

Mammalian Fatty Acid Synthase: X-ray Structure of a Molecular Assembly Line

Francisco J. Asturias*

Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Fatty acids are ubiquitous cellular components with a variety of structural and metabolic functions, from energy storage to modulation of gene expression. In plants and bacteria, *de novo* synthesis of fatty acids is carried out by individual enzymes. In contrast, in animal cells, the enzymes involved in fatty acid synthesis are organized in a single fatty acid synthase polypeptide (Figure 1). Encompassing all of the activities required to retain and extend an acyl chain from acetyl-CoA and malonyl-CoA precursors and then release a full-length, long-chain fatty acid, mammalian fatty acid synthase (FAS) constitutes a true molecular assembly line (1). Human FAS is a target for drug development against obesity and related diseases, and FAS inhibitors have been shown to have antitumor activity. Although the biochemistry of fatty acid synthesis was established over 30 years ago, structural understanding of FAS had, until recently, been limited (high-resolution structures of individual prokaryotic enzymes and a couple of FAS domains were available). The situation has changed dramatically with the recent publication of the first X-ray structure of FAS by Nenad Ban and colleagues (2).

In mammalian cells, the functional form of FAS is a homodimer (MW ~540 000 Da) with two centers for fatty acid synthesis. The model for FAS organization that prevailed for over 20 years was based on a logical but naive interpretation of the possible structure of an FAS monomer (3). It was observed that the bifunctional compound dibromopropanone could cross-link

the thiol in the active site cysteine of the N-terminal β -ketoacyl synthase (KS) domain of one monomer with the thiol in the phosphopantetheine group of the acyl carrier protein (ACP) located near the C-terminus of the other monomer. Therefore, a model with two extended FAS monomers in an antiparallel arrangement was proposed, thereby providing a simple explanation for the complementary interaction of the N-terminal condensing domains of one monomer with the C-terminal reducing and chain terminating domains of the other. However, functional complementation studies of baculovirus-expressed active site FAS mutants revealed a number of intramonomer interactions that could not be explained by the extended, antiparallel model of FAS (4). Careful re-examination of the dibromopropanone experiment (5) also revealed both intra- and intermonomer cross-links. These investigations eventually resulted in the formulation of an alternative model for FAS organization in which the two monomers were intertwined (6) (Box 1).

The paucity of structural information on FAS that prevailed until very recently illustrates some of the challenges in structural analysis of large macromolecular complexes. A relatively new technique that can often be used to determine low (10–30 Å)-resolution structures of large macromolecules is molecular electron microscopy (EM). Analysis of FAS by EM was complicated by the flexibility of the molecule, and a first low-resolution 3-D EM reconstruction of FAS was interpreted

ABSTRACT Mammalian fatty acid synthase (FAS) is a homodimeric, multifunctional polypeptide which comprises two full sets of catalytic subunits that carry out fatty acid synthesis. A recently published X-ray structure of FAS reveals, for the first time, the organization of all active sites involved in acyl chain elongation and provides a structural framework for interpretation of extensive functional studies. Further analysis with techniques capable of providing information about single molecule conformations will eventually provide a more complete understanding of FAS.

* To whom correspondence should be addressed.
E-mail: asturias@scripps.edu.

Published online April 21, 2006
10.1021/cb6001448 CCC: \$33.50
© 2006 by American Chemical Society

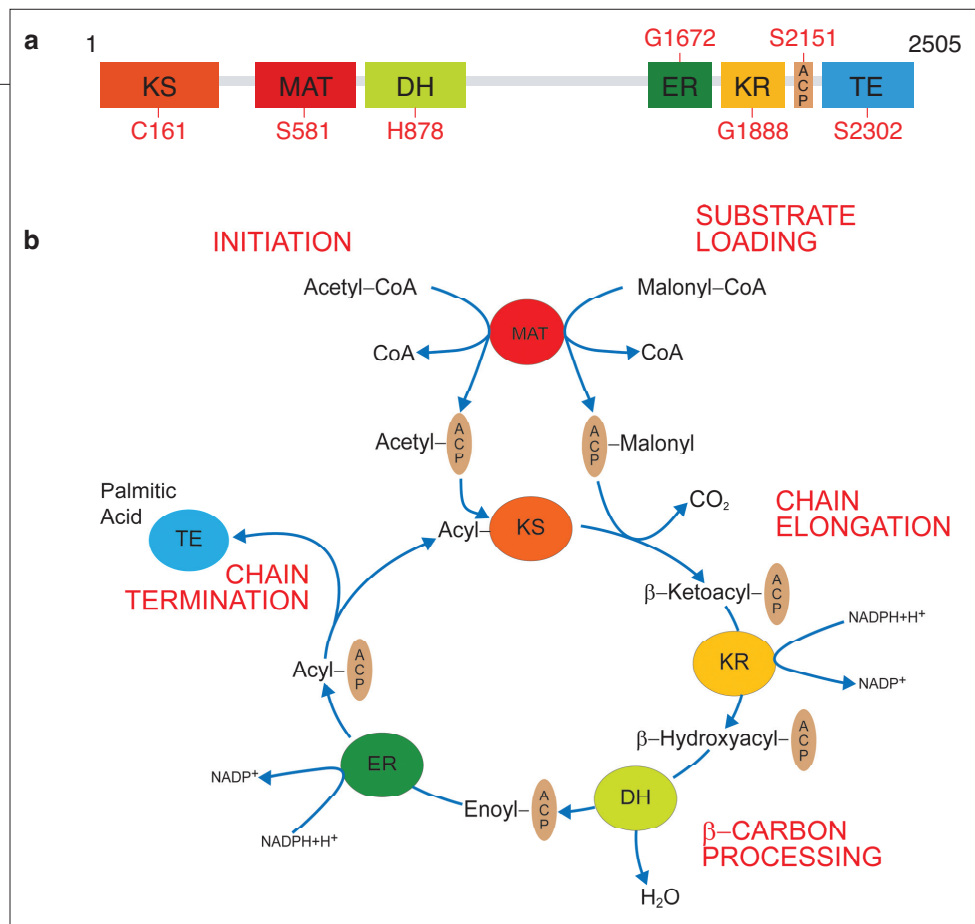


Figure 1. Active sites in the primary sequence of an FAS monomer and the fatty acid catalytic cycle. **a)** Arrangement of domains in the primary sequence of a fatty acid synthase monomer. Critical active site residues are indicated. The N-terminal β -ketoacyl synthase (KS, orange) domain is followed by malonyl/acetyl transferase (MAT, red), dehydratase (DH, light green), α,β -enoyl reductase (ER, dark green), β -ketoacyl reductase (KR, yellow), acyl carrier protein (ACP, maroon) with its prosthetic phosphopantetheine group, and finally the C-terminal thioesterase (TE, light blue) domain. The core region between the DH and ER domains has no catalytic activity. **b)** Cycle of reactions catalyzed by FAS, which culminate in synthesis of a long-chain fatty acid. The cycle is initiated when MAT catalyzes transfer of the acyl moiety of acetyl-CoA (initiation substrate) to the ACP. The acyl moiety is momentarily transferred to the KS domain, and MAT catalyzes transacylation of the malonyl group of malonyl-CoA (elongation substrate) to the ACP. KS catalyzes decarboxylative condensation to an acetoacetyl-ACP. KR catalyzes NADPH-dependent reduction of the β -carbon, and DH dehydrates the resulting β -hydroxyacyl-ACP to an α,β -enoyl intermediate. ER catalyzes the NADPH-dependent reduction of the enoyl to produce a four-carbon acyl chain to which two-carbon units derived from malonyl-CoA are attached in subsequent elongation cycles. When the acyl chain reaches a length of 16–18 carbons, it is released from the ACP by TE.

as providing evidence for the extended antiparallel model, disregarding the inconsistency with the results from FAS functional complementation assays (7). In a more recent EM study, variability in FAS conformation was reduced by the use of a KS point mutant imaged in the presence of substrates. This approach yielded a more faithful 3-D reconstruction of FAS, which

was asymmetric and clearly inconsistent with the predictions of the extended antiparallel model. Analysis of the 3-D structure of an FAS monomer and localization of the N-termini (by tagging an FAS mutant with a metal cluster directly visible in images of single FAS particles) to the center of the FAS structure were combined to propose a revised model for FAS organization that was

consistent with the biochemical and functional information (8). This represented a significant advance, but the organization of the different active sites in FAS remained obscure.

Our understanding of FAS has been taken to the next level with the publication of its X-ray structure (2). This was a long-awaited result, as needle-like crystals of FAS were reported over 30 years ago, in one of the first papers to describe a protocol for FAS purification. Because the resolution of the X-ray electron density map extended to only 4.5 Å, it was not possible to identify individual residues or trace the polypeptide backbone of the FAS monomers. However, secondary structure elements were clearly identifiable, and the electron density map was interpreted by fitting high-resolution structures of individual domain homologues (Figure 2). The overall X-shape of the FAS structure is very similar to the most recent EM reconstruction (8) and, like it, appears asymmetric. As predicted by biochemical and EM analysis (8, 9), the KS domains are dimeric and, along with the dehydratase (DH) and β -enoyl reductase (ER) domains, form the central portion of the FAS structure. The ER domains are also dimeric (contributing

significantly to the monomer–monomer interface), and the DH domains adopt a pseudo-dimeric fold within each monomer. Protruding from this central portion of the structure are pseudo-symmetrically placed malonyl/acetyl transferase (MAT) and β -ketoacyl reductase (KR) domains, both of which are monomeric. A large, catalytically inactive region that follows

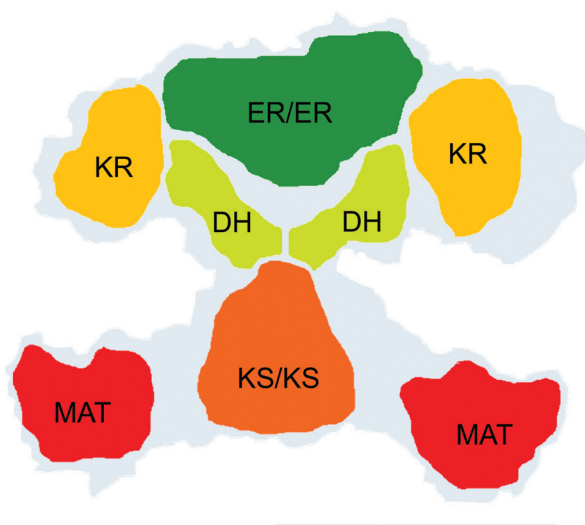


Figure 2. Organization of FAS active sites. Overlaid on an outline of the X-ray structure are the different active sites. The KS dimer forms the bottom part of the central portion of the structure. Protruding from the bottom are the (monomeric) MAT domains. The DH domains sit atop the KS dimer, and the ER domains (also dimeric) form the top central portion of the structure. Finally, extending from the top are the monomeric KR domains. The noncatalytic “core” in the FAS primary sequence appears to be distributed among the other active sites. The active sites on either side of the pseudo-symmetric structure are arranged around a reaction chamber. Extra density after the KR domain on the right likely corresponds to part of the ACP and TE domains in that monomer. The active sites are in all cases located near the center of the domains. The length of the prosthetic group in the ACP is long enough to reach each active site from the periphery of the corresponding domain, but the ACP itself must be highly mobile to reach the different domains around each reaction chamber. The scale bar corresponds to 100 Å.

the DH domain appears to be distributed throughout the central portion of the FAS structure, contributing to the interfaces between the KS, MAT, and DH domains.

The arrangement of domains around two reaction chambers formed by the coiled FAS monomers makes intuitive sense, in that it accommodates necessary functional interactions without requiring major domain rearrangements. Perhaps the most significant issue that remains unresolved is the localization and range of motion of the C-terminal ACP and thioesterase (TE) domains not included in the X-ray structure of FAS. The prosthetic phosphopantetheine group of the ACP plays a critical role by translocating substrates from one FAS active site to another, and must therefore be able to reach those sites. From the available ACP structures (10, 11), it is clear that the phosphopantetheine “arm” could only extend far enough to reach from the edge of a given domain to the domain’s internal active site. This implies that the ACP must have a significant degree of mobility. Harder to understand is the reason for the high mobility of the C-terminal TE domain, which results from more than its attachment to the mobile ACP, as the ACP and TE domains are connected by what appears

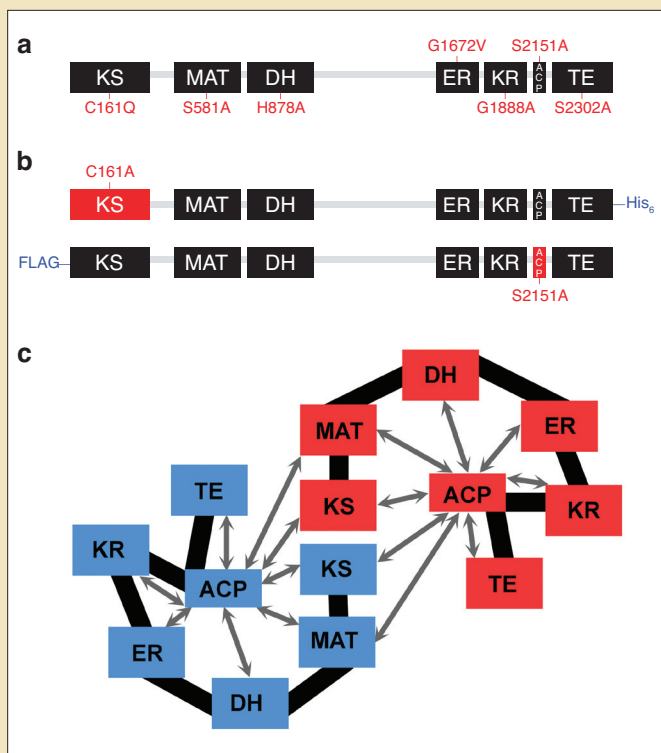
to be a highly flexible, disordered linker whose length affects TE functionality (12).

The X-ray structure (2) does not address the issue of monomer organization in FAS, but consideration of the EM results (8) suggests that the two monomers likely cross over each other in the central portion of the structure. Regardless, the static X-ray structure cannot explain both the intra- and intermonomer functional interactions revealed by functional complementation studies (4) (ACP domains can work with the KS and MAT in the same or the other monomer, see Box 1). Those interactions must be facilitated by the combined effect of large ACP mobility and conformational changes in the portions of the FAS structure included in the current X-ray map. The existence of a significant degree of flexibility in FAS has been established by alignment and multivariate statistical analysis of single particle images and by normal-mode elastic deformation analysis of the low-resolution structure of FAS (8, 13).

FAS has long been considered a paradigm for understanding the modular structure of the giant polyketide synthases (PKS) responsible for biosynthesis of some antibiotics and other important biologically active compounds. The X-ray structure of FAS

now reveals that the dimerization interface comprises domains that are often absent in PKS modules. Interestingly, along with the structure of mammalian FAS, Ban and colleagues also reported on the structure of a vastly different fungal FAS (14). Previous studies had established fundamental differences in structure between mammalian and yeast FAS (15), but the X-ray structures now illustrate how two synthases with completely different organization carry out the same biosynthetic task. Perhaps some of the strategies that made possible calculation of an FAS X-ray structure could be adapted to enable similar analysis of PKS modules.

Differences between the FAS structures calculated by X-ray crystallography and EM (particularly apparent when top or bottom views of the structures compared) suggest that the role that dimerization plays in FAS function and the reasons behind the asymmetric nature of the FAS structure are yet to be understood. The X-ray structure of FAS determined by Ban and colleagues (2) will prove very valuable in designing experiments and interpreting results from techniques such as optical, fluorescence, and electron microscopy, that can record “snapshots” of individual FAS molecules under functionally relevant conditions.



Box 1. Functional characterization of FAS using baculovirus-expressed active site point mutants.

Extensive functional studies have been carried out to characterize the interactions between FAS active sites, taking advantage of point mutations that knock-out the activity of individual domains. a) Point mutations that inactivate each of the FAS active sites have been identified, and FAS mutants with a given mutation can be expressed in *Sf9* insect cells and purified by affinity chromatography. b) Reversible dissociation of the FAS monomers induced by low temperature (4 °C) incubation makes possible assembly of heterodimers with specific mutations, which can be used to characterize inter- and intramonomer functional interactions. c) Functional complementation studies using a variety of mutants have resulted in a map of active site interactions. The acyl carrier protein (ACP) groups can interact with the β -ketoacyl synthase (KS) and malonyl/acetyl transferase (MAT) sites of either FAS monomer. In general, 65–80% of elongation cycles involve interactions between one ACP and the KS group of the other monomer. However, the other 20–35% of elongation cycles involve intramonomer ACP/KS interactions, a finding that was in clear conflict with the initial model for FAS organization depicting two extended monomers in an antiparallel orientation. In fact, the ultimate demonstration of the significance of intramonomer functional comple-

mentation came from an experiment in which a wild-type FAS monomer was dimerized with a fully inactive one. The resulting dimer showed fatty acid synthesis activity corresponding to about 30% that of a wild-type FAS dimer. These functional complementation studies were carried out in the last several years by Stuart Smith and his group at Children’s Hospital Oakland Research Institute.

REFERENCES

- Smith, S., Witkowski, A., and Joshi, A.K. (2003) Structural and functional organization of the animal fatty acid synthase, *Prog. Lipid Res.* 42, 289–317.
- Maier, T., Jenni, S., and Ban, N. (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution, *Science* 311, 1258–1262.
- Chirala, S. S., and Wakil, S. J. (2004) Structure and function of animal fatty acid synthase, *Lipids* 39, 1045–1053.
- Rangan, V. S., Joshi, A.K., and Smith, S. (2001) Mapping the functional topology of the animal fatty acid synthase by mutant complementation *in vitro*, *Biochemistry* 40, 10792–10799.
- Witkowski, A., Joshi, A. K., Rangan, V. S., Falick, A. M., Witkowska, H. E., and Smith, S. (1999) Dibromopropanone cross-linking of the phosphopantetheine and active-site cysteine thiols of the animal fatty acid synthase can occur both inter- and intrasubunit. Reevaluation of the side-by-side, antiparallel subunit model, *J. Biol. Chem.* 274, 11557–11563.
- Joshi, A. K., Rangan, V. S., Witkowski, A., and Smith, S. (2003) Engineering of an active animal fatty acid synthase dimer with only one competent subunit, *Chem. Biol.* 10, 169–173.
- Brink, J., Ludtke, S. J., Yang, C. Y., Gu, Z. W., Wakil, S. J., and Chiu, W. (2002) Quaternary structure of human fatty acid synthase by electron cryomicroscopy, *Proc. Natl. Acad. Sci. U.S.A.* 99, 138–143.
- Asturias, F. J., Chadick, J. Z., Cheung, I. K., Stark, H., Witkowski, A., Joshi, A. K., and Smith, S. (2005) Structure and molecular organization of mammalian fatty acid synthase, *Nat. Struct. Mol. Biol.* 12, 225–232.
- Witkowski, A., Ghosal, A., Joshi, A. K., Witkowska, H. E., Asturias, F. J., and Smith, S. (2004) Head-to-head coiled arrangement of the subunits of the animal fatty acid synthase, *Chem. Biol.* 11, 1667–1676.
- Roujeinikova, A., Baldock, C., Simon, W. J., Gilroy, J., Baker, P. J., Stuijter, A. R., Rice, D. W., Slabas, A. R., and Rafferty, J.B. (2002) X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site, *Structure* 10, 825–835.
- Holak, T. A., Nilges, M., Prestegard, J. H., Gronenbom, A. M., and Clore, G.M. (1988) Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry, *Eur. J. Biochem.* 175, 9–15.
- Joshi, A. K., Witkowski, A., Berman, H. A., Zhang, L., and Smith, S. (2005) Effect of modification of the length and flexibility of the acyl carrier protein-thioesterase interdomain linker on functionality of the animal fatty acid synthase, *Biochemistry* 44, 4100–4107.
- Ming, D., Kong, Y., Wakil, S. J., Brink, J., and Ma, J. (2002) Domain movements in human fatty acid synthase by quantized elastic deformational model, *Proc. Natl. Acad. Sci. U.S.A.* 99, 7895–7899.
- Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006) Architecture of a fungal fatty acid synthase at 5 Å resolution, *Science* 311, 1263–1267.
- Kolodziej, S. J., Penczek, P. A., Schroeter, J. P., and Stoops, J. K. (1996) Structure–function relationships of the *Saccharomyces cerevisiae* fatty acid synthase. Three-dimensional structure, *J. Biol. Chem.* 271, 28422–28429.

Chemical Genomics: Dialed in Transcriptional Network Control with Non-steroidal Glucocorticoid Receptor Modulators

Jasmina J. Allen and Kevan M. Shokat*

University of California-San Francisco, 600 16th Street, GH-N512D, San Francisco, California 94143 and Howard Hughes Medical Institute

Nuclear hormone receptors (NHRs) act as transmitters to communicate the hormonal status of an organism to the nucleus where they initiate and modulate transcriptional programs (1). Nuclear receptors are attractive drug targets because they have the power to control transcription at several gene loci, thereby having powerful and lasting effects on cellular processes. However, targeting the combinatorial nature of gene transcriptional regulation by NHRs also has drawbacks because undesirable transcriptional reprogramming leads to serious side effects. Therefore, a major challenge in NHR drug discovery is achievement of gene-specific regulation. For example, selective estrogen receptor modulators (SERMs) (2) are able to decouple some of the functions of the estrogen receptor. A parallel venture is development of selective glucocorticoid receptor modulators (SGRMs) (3) that retain the desired anti-inflammatory and immunosuppressive functions of glucocorticoids but do not induce harmful side effects such as osteoporosis (bone loss), and metabolic disorders. In the March issue of *Genes and Development* (4), a new class of non-steroidal arylpyrazole compounds, designed by Shah and Scanlan (5), were shown by Keith Yamamoto's laboratory to modulate glucocorticoid receptor (GR)-mediated transcription in a gene- and cell-specific manner. Although the panel of compounds all bind GR with nanomolar affinities and differ from each other by substitution at a single position, the transcrip-

tional profiles and resultant cellular effects varied dramatically. Many of the compounds were able to induce the desired anti-inflammatory effects of glucocorticoids without affecting osteoblast (bone cell) differentiation, potentially decoupling some of the beneficial effects of GR activation from detrimental effects such as osteoporosis. Additionally, such compounds are valuable tools to study how subtle changes in ligand structure affect GR conformation and the resulting biological output.

GR is a ligand-inducible transcriptional regulator, comprising DNA, ligand, and protein-protein interaction domains. Upon ligand binding, inhibitory chaperones are shed and the receptor enters the nucleus, where transcriptional regulation occurs by three mechanisms: (i) directly binding to simple glucocorticoid response elements (GREs), (ii) cooperatively binding at promoters that contain GREs and additional transcription factor binding sites (composite GREs), and (iii) allosteric tethering through nonreceptor transcription factors. Transcriptional regulation by GR reflects information integrated from promoter architecture, ligand structure, and cofactor composition. It is believed that differential utilization of GR protein surfaces dictates GRE and cofactor binding (6). Small molecule control of regulatory surface availability is better understood in the case of the estrogen receptor, where structural studies revealed that different SERMs induce different transactivation domain conformations (7), affecting corepressor recruitment. The

ABSTRACT A recent study analyzed the transcriptional effects induced by a panel of non-steroidal glucocorticoid receptor modulators. The authors discover patterns of cell-, gene-, and mechanism-specific regulation, with implications for development of improved anti-inflammatory agents.

*To whom correspondence should be addressed.
E-mail: shokat@cmp.ucsf.edu.

Published online April 21, 2006
10.1021/cb600137r CCC: \$33.50
© 2006 by American Chemical Society

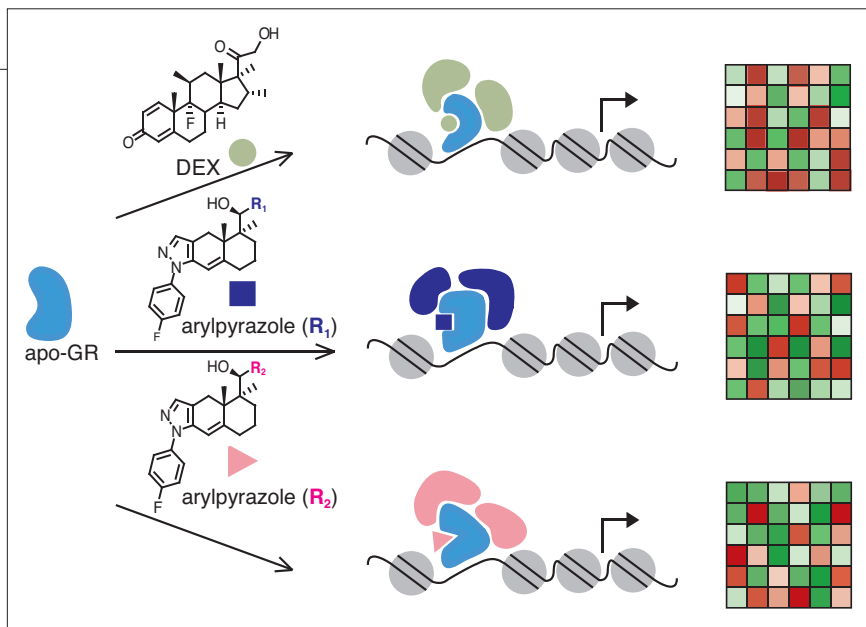


Figure 1. Model of the effects of ligand structure on GR-mediated transcription. Upon binding different ligands different GR regulatory surfaces are exposed, resulting in distinct cofactor recruitment. Extrapolating these effects to a set of GR-regulated genes yields compound-specific transcriptional signatures.

ability of ligands to modulate protein–protein and protein–DNA interactions is manifested in the activation or repression of genes that require these interactions.

What does the ideal therapeutic GR transcriptional profile look like? Unfortunately these details are not known, but it is anticipated that an ideal SGRM would influence transcription at a subset of glucocorticoid-regulated genes while leaving others unaffected. The experimental approach taken by the Yamamoto and Scanlan laboratories addressed the relationship between chemical structure and the resulting transcriptional output, by systematically correlating the arylpyrazole scaffold substituents with expression profiles. First, they established the effects of the compounds in cell-based assays. Like dexamethasone (DEX), a classical synthetic glucocorticoid, many of the compounds were able to suppress growth of glucocorticoid target cells that respond to pro-inflammatory signals. A different set of the compounds induced differentiation of pre-adipocytes, suggesting that subsets of compounds affect transcription by distinct mechanisms. Most provocatively, none of the compounds inhibited

differentiation of pre-osteoblasts, which is strongly inhibited by DEX.

All but one of the compounds activated a GRE containing reporter plasmid, but when the authors profiled the expression of 17 endogenous GR-regulated genes, distinct patterns of activation and repression were observed with each compound, highlighting the importance of chromosomal context. Differences in expression were then correlated with GR occupancy; some of the compounds inhibited or enhanced GR binding to GREs, but others had modest effects. One compound in particular allowed GR to bind the epithelial sodium channel GRE, but transcription was not activated. This was correlated with a lack of histone acetylation, an activating chromatin modification. Presumably, this compound rendered GR unable to recruit a histone acetylase. At another promoter this same compound inhibited the GR:GRE interaction. Revealing promoter-specific requirements and exemplifying how mechanistic questions can be addressed with this panel of chemical tools.

Chemical genomics, the study of small molecule regulation of gene expression, is

being advanced by both the development of techniques to rapidly analyze transcription networks and the availability of appropriate chemical tools. Through fruitful collaborations between synthetic chemists and biologists who study the fundamentals of transcription, we are beginning to understand how subtle chemical changes to GR ligands lead to biological effects mediated through genetic reprogramming. Whether these arylpyrazoles or similar compounds will become pharmaceutical agents remains to be seen, but their utility as chemical tools for systematically studying mechanisms of GR-mediated transcription is already clear.

REFERENCES

1. Edwards, D. P. (2000) The role of coactivators and corepressors in the biology and mechanism of action of steroid hormone receptors, *J Mammary Gland Biol. Neoplasia* 5, 307–324.
2. Jordan, V. C. (2004) Selective estrogen receptor modulation: Concept and consequences in cancer, *Cancer Cell* 5, 207–213.
3. Rosen, J., and Miner, J. N. (2005) The search for safer glucocorticoid receptor ligands, *Endocr. Rev.* 26, 452–464.
4. Wang, J. C., Shah, N., Pantoja, C., Meijnsing, S. H., Ho, J. D., Scanlan, T. S., and Yamamoto, K. R. (2006) Novel arylpyrazole compounds selectively modulate glucocorticoid receptor regulatory activity, *Genes Dev.* 20, 689–699.
5. Shah, N., and Scanlan, T. S. (2004) Design and evaluation of novel nonsteroidal dissociating glucocorticoid receptor ligands, *Bioorg. Med. Chem. Lett.* 14, 5199–5203.
6. Rogatsky, I., Wang, J. C., Derynck, M. K., Nonaka, D. F., Khodabakhsh, D. B., Haqq, C. M., Darimont, B. D., Garabedian, M. J., and Yamamoto, K. R. (2003) Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor, *Proc. Natl. Acad. Sci. U.S.A.* 100, 13845–13850.
7. Brzozowski, A. M., Pike, A. C., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G. L., Gustafsson, J. A., and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 389, 753–758.

Cholesterol, It's Not Just For Heart Disease Anymore

Amy Kerzmann and Andrew L. Feig*

Department of Chemistry, Indiana University, 800 East Kirkwood Avenue, Bloomington, Indiana 47405

An insidious threat lurks in the dark corners of hospital wards. Because of the widespread use of broad-spectrum antibiotics, *C. difficile*, an opportunistic pathogen, has become one of the most common hospital-acquired infections in the United States and Canada (1–3). It is also a growing threat to patients in nursing homes and extended care facilities (2). *C. difficile* colonizes the underpopulated anaerobic niches in the GI tracts of patients after their normal microflora has been killed. This organism causes pseudomembranous colitis and severe antibiotic-associated diarrhea, also called *C. difficile*-associated diarrhea (CDAD) (1). The organism secretes two toxins, Toxin A (TcdA) and Toxin B (TcdB), that are the virulence factors responsible for the cellular damage (4). It was estimated in 2002 that the U.S. medical community spends more than \$1.1 billion annually combating these infections (5). Recent work by Giesemann and colleagues shows that membrane cholesterol levels play a significant role in the transport of the *C. difficile* toxins into eukaryotic cells (6). The data suggest the intriguing possibility that cholesterol may act as a small molecule chaperone to facilitate the insertion of the protein into the membrane, thus, generating the pore necessary to translocate the catalytic component of the toxin into the cytoplasm.

TcdA and TcdB are very large proteins (on the order of 300 kDa). They share 48% identity and *in vitro* behave in a very similar fashion. Each toxin is comprised of three functional domains (Figure 1,

panel a) (7). A C-terminal domain binds cell surface receptors for target recognition, and an N-terminal domain carries a glucosyltransferase functionality that targets a family of Ras-like G-proteins, disrupting their function upon modification. The large central domain participates in the translocation process. It creates the pore necessary to transport the catalytic domain across the endosomal membrane after endocytosis. Once internalized, it is believed that the catalytic domain is proteolytically processed by a cellular protease around residue 543, freeing it from the remainder of the toxin which, having completed its function, remains in the endosomal membrane (8). Interrupting any of these three major steps (cellular recognition, internalization, or catalysis) could in theory disrupt intoxication.

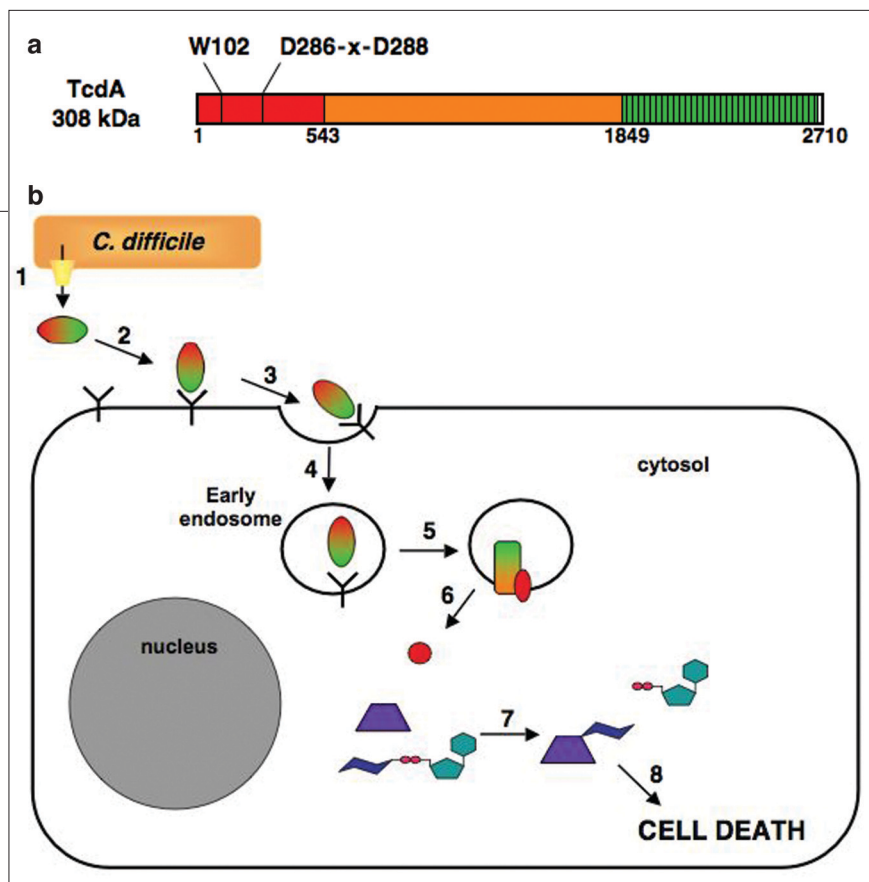
The Internalization Process. One of the critical events during intoxication by the bacterial toxin is a protein transduction step, the movement of the catalytic domain from outside the cell into the cytoplasm (Figure 1, panel b, steps 3–6) (9, 10). In this process, the toxin must trigger endocytosis and then escape from the endosome into the cytoplasm. The toxins that use endosome-mediated uptake can be divided into two classes. Cholera and Shiga toxins exemplify one class. They travel from the endosome into the Golgi body and eventually to the endoplasmic reticulum before leaving the vacuole and entering the cytoplasm. This mode of uptake is called the long-trip model. TcdA and TcdB are members of the other class of toxins that

ABSTRACT Recent studies have shown that cholesterol plays a significant role in the ability of Toxin A from *Clostridium difficile* to enter eukaryotic cells. The translocation process is one of three major steps during intoxication that could be targeted for intervention against the severe antibiotic-associated diarrhea caused by *C. difficile*.

*To whom correspondence should be addressed.
E-mail: afeig@indiana.edu.

Published online April 21, 2006
10.1021/cb600133b CCC: \$33.50
© 2006 by American Chemical Society

Figure 1. a) The three functional regions of TcdA: the N-terminal enzymatic domain (red), the central translocation region (orange), and the C-terminal repetitive oligopeptide (CROP) domain (green). There is also a small C-terminal hydrophobic region present after the CROP region in TcdA (white), not observed in TcdB. The highlighted residues are essential for UDP-Glc binding (W102) and/or catalysis (DxD motif at 286–288). b) Cellular intoxication by *C. difficile* TcdA and TcdB. (Step 1) TcdA and TcdB are exported from the bacterium; (step 2) the C-terminal CROP motif binds to cell-surface carbohydrates; (steps 3 and 4) the toxin–receptor complex is internalized through receptor-mediated endocytosis; (step 5) acidification of the maturing endosome by V-type ATPases drives a pH-dependent conformational change of the central translocation domain, resulting in insertion into cholesterol-containing endosomal membranes; (step 6) the N-terminal catalytic fragment is released into the cytosol; (step 7) once in the cytosol, the toxin fragment catalyzes the transfer of glucose from UDP-glucose to a conserved threonine residue of specific Ras-like GTPases; and (step 8) monoglucosylation of the G-proteins blocks the conformational changes that normally occur as the protein switches between GDP- and GTP-bound states. By preventing these conformational changes, these proteins are effectively “turned off” and are unable to interact with their effector proteins leading to depolymerization of the actin cytoskeleton, cellular rounding, and ultimately cell death.



also includes the diphtheria and anthrax toxins. These toxins escape directly from the early endosomes (Figure 1b, step 6). The endosomal compartment is acidified by vacuolar ATPases as it begins to travel toward the Golgi body. The toxin uses the acidification process as its cue to insert into the membrane, forming a pore and extruding its catalytic domain into the cytoplasm.

Giesemann and colleagues have explored the translocation phase of TcdA intoxication and found it to be highly dependent upon the presence of membrane cholesterol (6). The authors preloaded cells with $^{86}\text{Rb}^+$. They then used $^{86}\text{Rb}^+$ efflux measurements and single channel conductance to probe whether the toxin successfully inserted into the membrane and formed a channel. By analyzing intoxication of cell lines that differed in membrane cholesterol content, they could compare the relative susceptibility to the toxin. They also treated cells with methyl- β -cyclodextrin

(M β CD), a reagent that binds to and depletes cholesterol from the plasma membrane. Cells treated with M β CD showed marked reductions in their mortality after exposure to the toxin. The toxin still bound to the surface of the cholesterol-depleted membranes, but failed to form pores, and the catalytic domain was incapable of being transported into the cytosol.

One possible reason for cholesterol-dependent toxin uptake would be involvement of lipid raft structures or the receptors associated with them. Lipid rafts are specialized microdomains in the membranes that have a high concentration of both cholesterol and membrane proteins (11, 12). By depleting cells of cholesterol, one might disrupt these rafts and thereby disperse the receptors contained therein. To address this possibility, Giesemann and colleagues treated cells with a phosphatidylinositol-specific phospholipase C (PI-PLC). This enzyme hydrolyzes the proteins and

polysaccharide receptors from glycosylphosphatidylinositol (GPI)-anchored structures, leaving the remainder of the lipid raft intact. HT-29 cells treated with PI-PLC were just as susceptible to intoxication as the untreated cells. This result demonstrated that GPI-anchored receptors are not required for uptake, but the remainder of the lipid raft might still play a role.

So what is the function of cholesterol in toxin translocation? Are there direct interactions between cholesterol and the toxin or does cholesterol simply affect the physical properties of the membrane making it more susceptible to protein insertion? A clue may come from examining the role of cholesterol in cytolysins, another class of bacterial toxins exemplified by perfringolysin O (PFO) (13). Cytolysins kill eukaryotic cells by forming oligomeric structures that breach the plasma membrane with large pores up to 300 Å across (13). It has been hypothesized that stable folding of

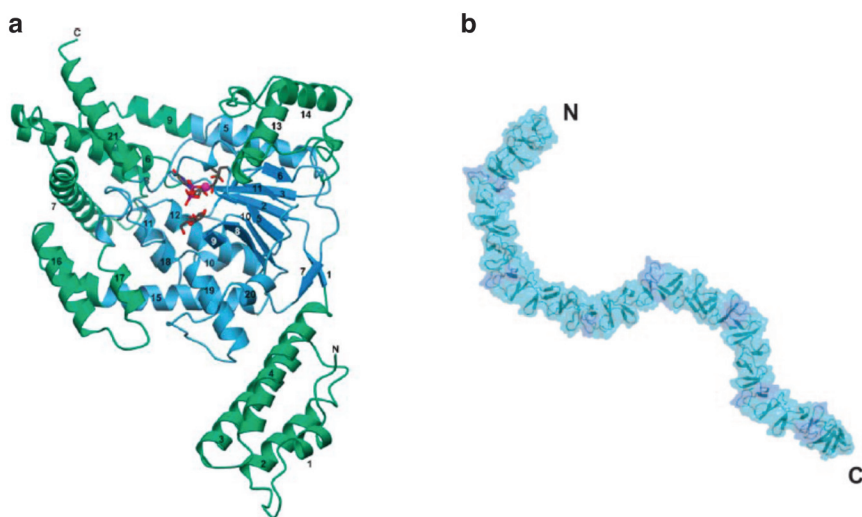


Figure 2. Structural biology of *C. difficile* toxin fragments. a) Model resulting from the X-ray crystal structure of a catalytically active N-terminal fragment of TcdB consisting of residues 1–543 (reprinted with permission from ref 16, Copyright 2005 Elsevier B.V.). The conserved GT-A fold is shown in blue. The catalytic DxD motif is shown in ball-and-stick mode, as well as UDP, glucose, and the catalytic Mn^{2+} ion. b) A structural model of the CROP domain from TcdA. The model is based on a crystal structure of a 127 amino acid fragment (residues 2573–2709) revealing stacked pairs of β -hairpins in a β -solenoid fold (reprinted with permission from ref 17, Copyright 2005 National Academy of Sciences, U.S.A.).

the monomeric PFO prevents premature aggregation of the cytolysins in solution. In cases where cholesterol has been depleted from target membranes, PFO oligomerizes on the membrane surface, but fails to insert into the membrane (14). Thus, cholesterol then may help to unfold the preinsertion structure of the toxin during the initial stages of membrane interaction (15).

C. difficile toxins do not need to aggregate the way PFO and related cytolysins do. However, the translocation domains of TcdA and TcdB do need to refold from their initial solution conformation to their membrane-inserted pore conformation at the proper time during entry into the cytosol (Figure 1, panel b, step 5). If the toxin refolds too early, it may be subject to aggregation and precipitation. If it occurs too late, it will have lost its chance to escape from the endosome. It is possible that acidification alone is insufficient to destabilize the preinsertion structures of TcdA and TcdB. Giesemann's results

are consistent with a mechanism whereby cholesterol may chaperone the membrane insertion process. The search for the mechanism by which cholesterol facilitates intoxication may provide a detailed window into the nature of protein insertion into membranes in addition to providing a potential mode for therapeutic intervention for CDAD.

Receptor Binding and Catalysis. Recent work has also expanded our understanding of the two other steps in the intoxication process. A high-resolution crystal structure of the catalytically active N-terminal domain consisting of residues 1–543 of TcdB was recently reported (Figure 2, panel a) (16). The model shows an extensive network of β structure at its core surrounded by a cluster of a helices. The core of the fold is homologous to the GT-A family of glycosyltransferases that also includes glycogenin, α 3-GalT, and LgtC. Whereas the basic glycosyltransferase behavior is probably quite similar to that of other members

of this family of enzymes, the toxins are unique in their exquisite selectivity for their protein acceptors. Besides TcdA and TcdB, there are several other large clostridial toxins, and each has a unique set of cellular targets (4). The potential to use this selectivity in specifically targeting the toxins for inhibition is so far underutilized.

A 127 amino acid fragment from the C-terminus of TcdA has also been structurally characterized (Figure 2, panel b) (17). This fragment derives from the CROP region of the protein responsible for binding receptors on the surface of colonic epithelial cells (Figure 1, panel b, step 2). This repetitive sequence folds into a series of β hairpins that stack on top of one another to form an extended filamentous assembly with a significant helical twist. This domain most likely protrudes from the body of the toxin in search of an appropriate cell-surface receptor, believed to be a short oligosaccharide motif. Several candidate trisaccharides have been reported as potential targets for TcdA binding, including Gal- α 1,3-Gal β -1,4-GlcNAc (18) and GalNAc- β 1,3-Gal- β 1,4-GlcNAc (19), but it remains unclear whether these are in fact the biologically relevant motifs. Much work remains to be done on this aspect of the cellular recognition problem.

Together, these most recent studies on the role of cholesterol and the mechanism by which the *C. difficile* toxins recognize, penetrate, and kill host cells will facilitate a host of additional experiments on these systems. Scientists have been working tirelessly to develop novel antibiotics and immunization against *C. difficile*, and this route still holds significant promise for long-term efficacy. The deeper understanding of the molecular details of intoxication, however, may allow direct targeting of the toxins in the ongoing battle against CDAD.

Acknowledgment: The authors acknowledge support from NIH Training Grant GM-007757 (to A.K.) and a Women in Science Fellowship (to A.K.). A.L.F. is a Cottrell Scholar of Research Corporation.

REFERENCES

- Kelly, C. P., Pothoulakis, C., and LaMont, J. T. (1994) *Clostridium difficile* colitis, *N. Engl. J. Med.* **330**, 257–262.
- Lamont, J. T. (2002) Theodore E. Woodward Award. How bacterial enterotoxins work: Insights from *in vivo* studies, *Trans. Am. Clin. Climatol. Assoc.* **113**, 167–180.
- Valiquette, L., Low, D. E., Pepin, J., and McGeer, A. (2004) *Clostridium difficile* infection in hospitals: A brewing storm, *Can. Med. Assoc. J.* **171**, 27–29.
- Just, I., and Gerhard, R. (2004) Large clostridial cytotoxins, *Rev. Physiol. Biochem. Pharmacol.* **152**, 23–47.
- Kyne, L., Hamel, M. B., Polavaram, R., and Kelly, C. P. (2002) Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*, *Clin. Infect. Dis.* **34**, 346–353.
- Geisemann, T., Jank, T., Gerhard, R., Maier, E., Just, I., Benz, R., and Aktories, K. (2006) Cholesterol-dependent pore formation of *Clostridium difficile* toxin A, *J. Biol. Chem.* **281**, 10808–10815.
- Moncrief, J. S., Lyerly, D. M., and Wilkins, T. D. (1997) Molecular biology of the *Clostridium difficile* toxins. In *The Clostridia: Molecular Biology and Pathogenesis* (Rood, J. I., McClane, B. A., Songer, J. G., and Titball, R. W., Eds.), pp 369–392, Academic Press, San Diego, CA.
- Rupnik, M., Pabst, S., Rupnik, M., von Eichel-Streiber, C., Urlaub, H., and Soling, H. D. (2005) Characterization of the cleavage site and function of resulting cleavage fragments after limited proteolysis of *Clostridium difficile* toxin B (TcdB) by host cells, *Microbiology* **151**, 199–208.
- Lord, J. M., and Roberts, L. M. (1998) Toxin entry: Retrograde transport through the secretory pathway, *J. Cell Biol.* **140**, 733–736.
- Lord, J. M., Smith, D. C., and Roberts, L. M. (1999) Toxin entry: How bacterial proteins get into mammalian cells, *Cell. Microbiol.* **1**, 85–91.
- Munro, S. (2003) Lipid rafts: Elusive or illusive? *Cell* **115**, 377–388.
- Pike, L. J. (2004) Lipid rafts: Heterogeneity on the high seas, *Biochem. J.* **378**, 281–292.
- van der Goot, F. G. (2003) Membrane-damaging toxins: Pore formation. In *Bacterial Protein Toxins* (Burns, D. L., Barbieri, J. T., Iglewski, B. H., and Rappuoli, R., Eds.), pp 189–202, ASM Press, Washington, DC.
- Giddings, K. S., Johnson, A. E., and Tweten, R. K. (2003) Redefining cholesterol's role in the mechanism of the cholesterol-dependent cytolysins, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11315–11320.
- Heuck, A. P., Hotze, E. M., Tweten, R. K., and Johnson, A. E. (2000) Mechanism of membrane insertion of a multimeric beta-barrel protein: Perfringolysin O creates a pore using ordered and coupled conformational changes, *Mol. Cell* **6**, 1233–1242.
- Reinert, D. J., Jank, T., Aktories, K., and Schulz, G. E. (2005) Structural basis for the function of *Clostridium difficile* toxin B, *J. Mol. Biol.* **351**, 973–981.
- Ho, J. G., Greco, A., Rupnik, M., and Ng, K. K. (2005) Crystal structure of receptor-binding C-terminal repeats from *Clostridium difficile* toxin A, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18373–18378.
- Krivan, H. C., Clark, G. F., Smith, D. F., and Wilkins, T. D. (1986) Cell surface binding site for *Clostridium difficile* enterotoxin evidence for glycoconjugate containing the sequence Gal α 1-3Gal β 1-4GlcNAc, *Infect. Immun.* **53**, 573–578.
- Teneberg, S., Lonnroth, I., Torres Lopez, J. F., Galili, U., Halvarsson, M. O., Angstrom, J., and Karlsson, K. A. (1996) Molecular mimicry in the recognition of glycosphingolipids by Gal α 3Gal β 4-GlcNAc β -binding *Clostridium difficile* toxin A, human natural anti- α -galactosyl IgG and the monoclonal antibody Gal-13: Characterization of a binding-active human glycosphingolipid, non-identical with the animal receptor, *Glycobiology* **6**, 599–609.

Targeting Translation in Hypoxic Tumors

David Ron^{†,*} and Alan G. Hinnebusch^{‡,*}

[†]Skirball Institute of Biomolecular Medicine and the Departments of Medicine and Cell Biology, New York University School of Medicine, New York, New York 10016, and [‡]Laboratory of Gene Regulation and Development, National Institute of Child Health and Human Development, Bethesda, Maryland 20892

Unlike normal tissues that are nourished by robust and well-regulated capillaries, malignant tumors have fragile and inadequate blood vessels. Tumor cells must therefore adapt to transient fluctuations in the supply of oxygen and other nutrients (1). Recent work has emphasized the role of translational control in coping with this stress (2). The ability to rapidly attenuate global protein synthesis is believed to conserve energy and other resources in the acutely stressed cell. Translational control also interfaces with other aspects of regulated gene expression to alter the internal milieu for longer-term adaptation to metabolic challenges. Two recent papers address the complex signaling pathways by which hypoxia attenuates mRNA translation (3, 4). These will be considered for the insight they provide on the possibility of targeting the translational apparatus to selectively inhibit hypoxic cancer cells and possibly manipulate other physiological processes.

Global protein synthesis is regulated at multiple levels, but among the best understood are two aspects that modulate initiation of mRNA translation. In eukaryotes most mRNAs are capped at their 5' ends by m⁷Gppp, which serves as a ligand for eukaryotic translation initiation factor 4E (eIF4E), the cap-binding subunit of a complex (eIF4F) that recruits 40S ribosomes to the mRNA. Both stimulatory and inhibitory signals converge on eIF4E to regulate eIF4F activity and thereby rates of translation initiation (5). The participation of the 40S ribosome in translation initiation

is also regulated, as it must be charged with a ternary complex of eIF2, GTP, and amino-acylated initiator methionyl tRNA. Formation of this ternary complex is limited by the activity of a guanine nucleotide exchange factor for eIF2, eIF2B. The latter is inhibited, in *trans*, by phosphorylation of its substrate, eIF2 on serine 51 of its α subunit. Specific kinases have evolved to couple eIF2 α phosphorylation and attenuated translation initiation to stressful events, and eIF2(α P)-specific phosphatases have evolved to counter-regulate and fine-tune translational repression (6, 7).

It has long been known that hypoxia actively attenuates protein synthesis, but the mechanisms involved have remained obscure (8). A clue was provided by the finding that ischemia and the attendant hypoxia and nutrient deprivation activate PERK (9, 10), an eIF2 α kinase that responds specifically to the stress of unfolded and misfolded proteins in the endoplasmic reticulum (ER stress) (11). The mechanism(s) by which hypoxia elicits ER stress and PERK activation remain obscure; however, the importance of PERK activation and eIF2(α P) to translational control in hypoxic cells is well documented by analysis of *PERK*^{-/-} cells and cells bearing a Ser51 to Ala mutation in the regulatory residue of eIF2 α (12). This adaptation to hypoxia has proven to promote survival of cancer cells in ischemic animal tumor models, as cancer cells compromised in their ability to effect eIF2 α phosphorylation in response to hypoxia were compromised in their ability to proliferate and form large,

ABSTRACT Recent insight into how mammalian cells adapt their translational machinery to hypoxic conditions raises the possibility of targeting components of the regulatory networks involved to *selectively* inhibit metabolically compromised tumor cells and possibly manipulate a broad range of other physiological processes.

*To whom correspondence should be addressed.
Email: ron@saturn.med.nyu.edu.
Email: ahinnebusch@nih.gov.

Published online April 21, 2006
10.1021/cb600125y CCC: \$33.50
© 2006 by American Chemical Society

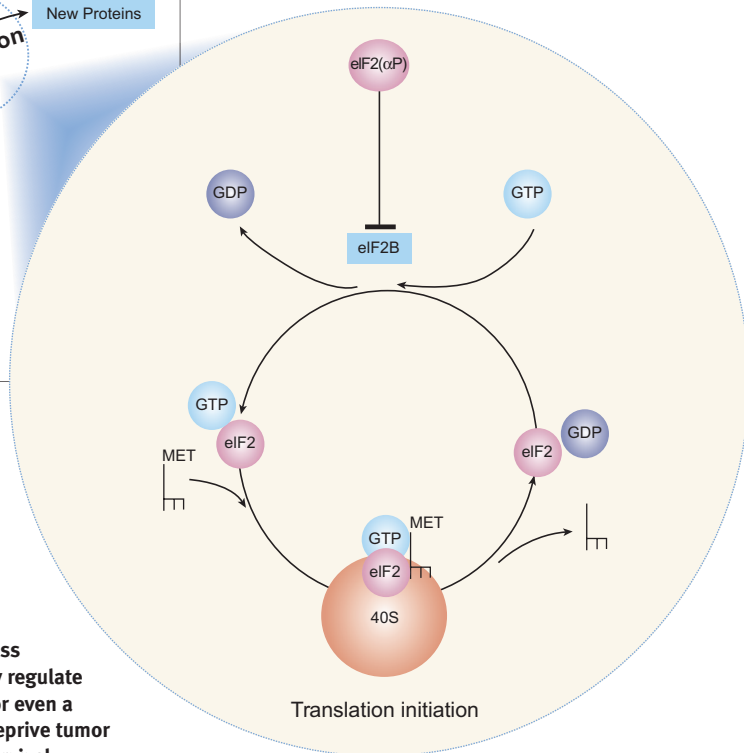
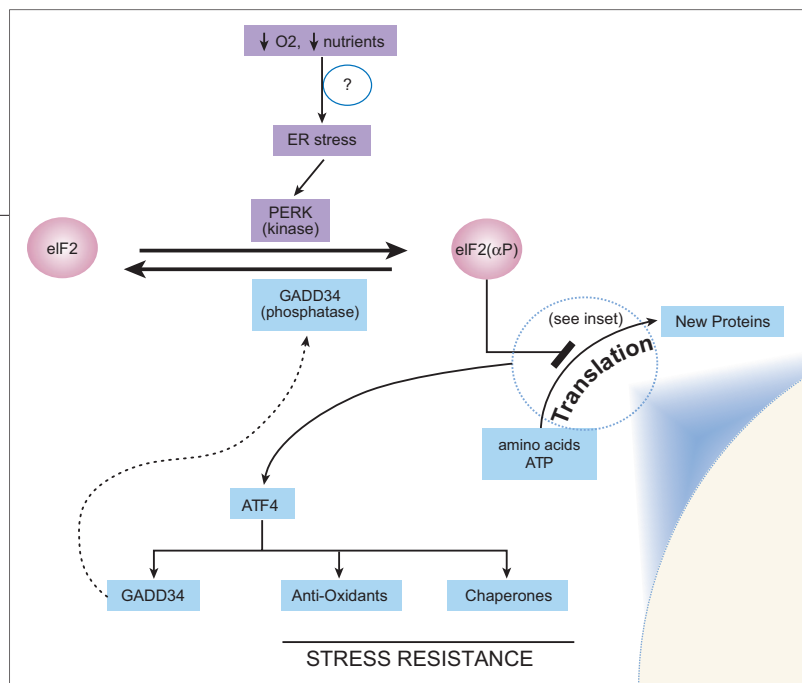


Figure 1. eIF2-Mediated translation regulation in hypoxic tumors. Hypoxia and nutrient deprivation, consequences of an unstable blood supply, promote endoplasmic reticulum stress in tumor cells. The eIF2 α kinase PERK is activated. eIF2(α P) inhibits the GTP exchange factor, eIF2B, and reduces levels of eIF2-GTP-charged initiator methionyl tRNA ternary complexes (see inset). Global protein synthesis is inhibited at the initiation step, conserving ATP, but some proteins, such as the transcription factor ATF4, are translationally induced as eIF2(α P) levels rise and these activate a downstream gene expression program that promotes stress resistance. ATF4 also upregulates a phosphatase, GADD34, to finely regulate levels of eIF2(α P) in the stressed cell. A specific inhibitor of PERK, or even a more general inhibitor of the entire class of eIF2 α kinases, would deprive tumor cells of this adaptation to hypoxia and thereby compromise their survival.

ischemic tumors (12). The adaptive role of eIF2 α phosphorylation is likely played out both at the level of its global effects on protein synthesis in hypoxic cells and through the induction of a cytoprotective gene expression program, known as the integrated stress response, which is activated by translationally controlled eIF2(α P)-dependent transcription factor(s) such as ATF4 (13) (Figure 1).

The two new papers from Liu (4), Korintzinsky (3), and colleagues confirm the contribution of eIF2 α phosphorylation to translational control in hypoxic cells but also suggest the importance of eIF4F regulation in the process. The cap-binding subunit eIF4E partitions between non-productive associations with a family of inhibitory eIF4E-binding proteins (4EBPs) and productive associations that form eIF4F. This equilibrium is regulated by the phosphorylation of the 4EBPs, which

reduces their affinity for eIF4E and thereby de-represses translation initiation. Not surprisingly, a variety of growth pathways contribute to 4EBP phosphorylation, whereas growth-factor or nutrient deprivation increases the levels of dephosphorylated, inhibitory 4EBPs (5) (Figure 2).

Hypoxia, as shown in both papers, also promotes higher levels of dephosphorylated 4EBPs and inhibits eIF4F formation. The paper by Liu and colleagues (4) examines in detail the signaling pathways involved in 4EBPs dephosphorylation. They focus attention on a recently-identified pathway that gauges energy insufficiency by monitoring the AMP/ATP ratio and in which the AMP-activated protein kinase (AMPK) serves as a major node. AMPK indirectly represses the TOR kinase and thereby leads to lower levels of phosphorylated 4EBPs and inhibits translation initiation. Liu and colleagues (4) call attention to redun-

dancy in pathways that couple hypoxia to TOR repression by demonstrating that interference with single components does not completely block 4EBPs dephosphorylation.

The paper by Korintzinsky and colleagues (3) contains an interesting twist concerning regulation of eIF4F; 4EBP dephosphorylation does not account for the entire disruptive effect of hypoxia on eIF4F formation. As expected, they report that an mRNA cap analog affinity matrix recovers eIF4E in complex with other components of eIF4F in lysates of unstressed cells. Hypoxia rapidly disrupts this complex, but not by the expected mechanism of replacing eIF4F components with dephosphorylated 4EBP as eIF4E's binding partners. Rather, early in the course of hypoxia, a fraction of eIF4E is whisked to the nucleus by a shuttling factor, eIF4T. The upstream signals remain obscure, but the nuclear sequestration of eIF4E, away

from eIF4F (and other components of the translational apparatus), correlates with the dephosphorylation of eIF4T. Hours later, they observe the expected association of dephosphorylated 4EBPs with eIF4E. By contrast eIF2 α phosphorylation is a rapid and transient response to hypoxia.

Interestingly, Koritzinsky and colleagues (3) note that abolishing eIF2 α phosphorylation eliminates most detectable changes in global translation in hypoxic cells. Thus, the physiological significance of the dramatic regulation of eIF4F in hypoxic cells, documented by both groups, remains to be established. Furthermore, abolishing eIF2 α phosphorylation also eliminates most detectable changes in mRNA-specific translation. However, it remains unclear if this reflects an ascertainment bias due to the limited number of genes examined or whether indeed eIF2(α P) also dominates as a signal to alter gene-specific translational profiles in hypoxic cells.

The existence of two discrete ways to regulate translation in hypoxic cells begs the question: How does regulation of eIF2 and eIF4F affect survival of hypoxic cells and growth of tumors? The role of eIF2(α P) in promoting the survival of hypoxic tumor cells is supported by genetic experiments (12); the contribution of hypoxia-mediated inhibition of eIF4F is less clear. Experiments in flies show that mutations compromising the ability to repress TOR and promote the regulated dephosphorylation of 4EBPs sensitize animals to hypoxia (14). Such experiments predict that stress-mediated 4EBPs dephosphorylation might contribute to survival of hypoxic tumor cells. At the same time, components of the apparatus

for linking AMPK to TOR repression are encoded by anti-oncogenes, and some measure of constitutive deregulation of mRNA translation may be a common feature of cancer cells (5). This suggests that de-repression of eIF4F, rather than its regulation by hypoxia, might be favored in cancer cells. It is notable in this regard that rapamycin, a drug that inhibits activation signals from TOR to eIF4F is being tested as an anti-tumor agent (15).

Chemical biology can help sort through the pathophysiological significance of the myriad pathways regulating translation in hypoxic tumor cells. As these rely on kinase cascades, they are potentially susceptible

to specific inhibitors. Compounds that inhibit AMPK (for which iodotubercidin is a relatively non-selective prototype) could be tested for their ability to inhibit survival of hypoxic tumor cells, and the findings of Liu (4), Koritzinsky (3), and their colleagues suggest the presence of other nodes that could be explored by more specific inhibitors, were they available. The relative simplicity of signaling through phosphorylated eIF2(α P) suggests a pathway that might be especially amenable to study with chemical inhibitors. The crystal structures of two eIF2 α kinases, PKR and GCN2 have been reported recently (16), and these contain design features unique to this

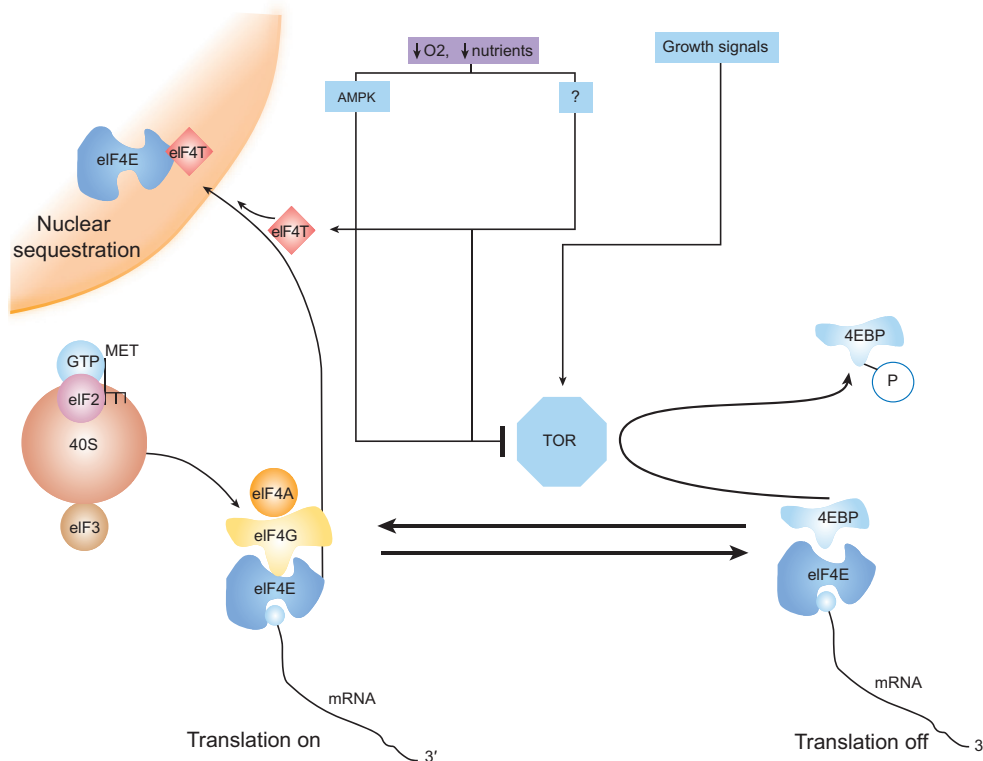


Figure 2. eIF4F-Mediated translational regulation in hypoxic tumors. Translation initiation requires the formation of an eIF4F complex (consisting of eIF4E, 4G and 4A), which binds the capped 5' end of mRNAs via eIF4E and recruits the fully-equipped 40S ribosome to initiate scanning for an AUG start codon on the mRNA. Hypoxia and nutrient deprivation activate the AMP-activated kinase and other stress-responsive pathways to inhibit the TOR kinase, which is under positive control of growth signals. Diminished TOR signaling results in hypophosphorylation of eIF4E binding proteins, the 4EBPs, and these sequester eIF4E in an inactive complex, repressing translation initiation. By poorly understood mechanisms hypoxia also promotes the eIF4T-mediated sequestration of eIF4E in the nucleus away from the translational apparatus.

entire family that could be exploited by specific inhibitors. Even the phosphatases that dephosphorylate eIF2(α P) might be accessed by small molecules, as attested to by the properties of salubrinal, an indirect inhibitor that has recently been shown to promote survival of ER stressed cells (17).

The two major pathways for general control of mRNA translation, touched on above, are very old and have arborized to influence not only survival of stressed cells but also a variety of other biological processes. These range from proliferation of immune cells required for graft rejection (pharmacological targets of the TOR-inhibitor rapamycin) to control of feeding behavior (18, 19) and memory formation (20). Small molecules that manipulate these pathways could find broad application as physiological probes and, possibly, valuable therapeutic agents.

REFERENCES

- Harris, A. L. (2002) Hypoxia—a key regulatory factor in tumour growth, *Nat. Rev. Cancer* 2, 38–47.
- Wouters, B. G., van den Beucken, T., Magagnin, M. G., Koritzinsky, M., Fels, D., and Koumenis, C. (2005) Control of the hypoxic response through regulation of mRNA translation, *Semin. Cell Dev. Biol.* 16, 487–501.
- Koritzinsky, M., Magagnin, M. G., van den Beucken, T., Seigneuric, R., Savelkoul, K., Dostie, J., Pyronnet, S., Kaufman, R. J., Wepler, S. A., Voncken, J. W., Lambin, P., Koumenis, C., Sonenberg, N., and Wouters, B. G. (2006) Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control, *EMBO J.* 25, 1114–1125.
- Liu, L., Cash, T. P., Jones, R. G., Keith, B., Thompson, C. B., and Simon, M. C. (2006) Hypoxia-Induced Energy Stress Regulates mRNA Translation and Cell Growth, *Mol. Cell* 21, 521–531.
- Richter, J. D., and Sonenberg, N. (2005) Regulation of cap-dependent translation by eIF4E inhibitory proteins, *Nature* 433, 477–480.
- Hinnebusch, A. G. (2000) Mechanism and regulation of initiator methionyl-tRNA binding to ribosomes. In *Translational Control of Gene Expression*; Sonenberg, N., Hershey, J. W. B., Mathews, M. B., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor; pp 185–243.
- Ron, D. (2002) Translational control in the endoplasmic reticulum stress response, *J. Clin. Invest.* 110, 1383–1388.
- Lefebvre, V. H., Van Steenbrugge, M., Beckers, V., Roberfroid, M., and Buc-Calderon, P. (1993) Adenine nucleotides and inhibition of protein synthesis in isolated hepatocytes incubated under different pO₂ levels, *Arch. Biochem. Biophys.* 304, 322–331.
- Kumar, R., Azam, S., Sullivan, J., Owen, C., Cavener, D., Zhang, P., Ron, D., Harding, H., Chen, J., Han, A., White, B., Krause, G., and DeGracia, D. (2001) Brain ischemia and reperfusion activates the eukaryotic initiation factor 2 α kinase, PERK, *J. Neurochem.* 77, 1418–1421.
- Koumenis, C., Naczki, C., Koritzinsky, M., Rastani, S., Diehl, A., Sonenberg, N., Koromilas, A., and Wouters, B. G. (2002) Regulation of protein synthesis by hypoxia via activation of the endoplasmic reticulum kinase PERK and phosphorylation of the translation initiation factor eIF2 α , *Mol. Cell. Biol.* 22, 7405–7416.
- Harding, H., Zhang, Y., and Ron, D. (1999) Translation and protein folding are coupled by an endoplasmic reticulum resident kinase, *Nature* 397, 271–274.
- Bi, M., Naczki, C., Koritzinsky, M., Fels, D., Hu, N., Harding, H., Novoa, I., Varia, M., Raleigh, J., Scheuner, D., Kaufman, R., Bell, J., Ron, D., Wouters, B., and Koumenis, C. (2005) ER stress-regulated translation increases tolerance to extreme hypoxia and promotes tumor growth, *EMBO J.* 24, 3470–3481.
- Harding, H., Zhang, Y., Zeng, H., Novoa, I., Lu, P., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., Stojdl, D., Bell, J., Hettmann, T., Leiden, J., and Ron, D. (2003) An integrated stress response regulates amino acid metabolism and resistance to oxidative stress, *Mol. Cell.* 11, 619–633.
- Reiling, J. H., and Hafen, E. (2004) The hypoxia-induced paralogs Scylla and Charybdis inhibit growth by down-regulating S6K activity upstream of TSC in *Drosophila*, *Genes Dev.* 18, 2879–2892.
- Sawyers, C. L. (2003) Will mTOR inhibitors make it as cancer drugs? *Cancer Cell* 4, 343–348.
- Hinnebusch, A. G. (2005) eIF2 α kinases provide a new solution to the puzzle of substrate specificity, *Nat. Struct. Mol. Biol.* 12, 835–838.
- Boyce, M., Bryant, K. F., Jousse, C., Long, K., Harding, H. P., Scheuner, D., Kaufman, R. J., Ma, D., Coen, D., Ron, D., and Yuan, J. (2005) A selective inhibitor of eIF2 α dephosphorylation protects cells from ER stress, *Science* 307, 935–939.
- Hao, S., Sharp, J. W., Ross-Inta, C. M., McDaniel, B. J., Anthony, T. G., Wek, R. C., Cavener, D. R., McGrath, B. C., Rudell, J. B., Koehnle, T. J., and Gietzen, D. W. (2005) Uncharged tRNA and sensing of amino acid deficiency in mammalian piriform cortex, *Science* 307, 1776–1778.
- Maurin, A., Jousse, C., Averous, J., Parry, L., A. B., Cherasse, Y., Zeng, H., Zhang, Y., Harding, H., Ron, D., and Fafournoux, P. (2005) The GCN2 kinase biases feeding behavior to maintain amino-acid homeostasis in omnivores, *Cell Metab.* 1, 273–277.
- Costa-Mattioli, M., Gobert, D., Harding, H. P., Herdy, B., Azzi, M., Bruno, M., Ben Mamou, C., Marcinkiewicz, E., Yoshida, M., Imataka, H., Cuello, A. C., Seidah, N., Sossin, W., Lacaille, J.-C., Ron, D., Nader, K., and Sonenberg, N. (2005) Translational control of hippocampal synaptic plasticity and memory by an eIF2 kinase, GCN2. *Nature* 436, 1166–1173.

B₁₂ Trafficking in Mammals: A Case for Coenzyme Escort Service

Ruma Banerjee*

Redox Biology Center and the Department of Biochemistry, University of Nebraska, Lincoln, Nebraska 68588-0664

Organic cofactors function as partners in critical enzyme-mediated transformations, facilitating the completion of chemically challenging reactions that would otherwise be impossible on a biological time scale. In mammals, many of these coenzymes are vitamins, *i.e.*, compounds that are not synthesized endogenously but are derived from the diet. While *de novo* pathways for the synthesis of vitamins do not exist in mammals, they do contain enzymes that catalyze the assimilation of vitamin precursors obtained from the diet into the corresponding coenzyme forms. For instance, pyridoxine, or vitamin B₆, is converted to pyridoxal-5'-phosphate for use in a plethora of reactions involved in amino acid metabolism and riboflavin, or vitamin B₂, is converted to FMN and FAD, which are employed by many redox enzymes (Table 1). The dependence of cellular metabolism on rare and often reactive coenzymes poses a challenge for their efficient delivery to target enzymes in a crowded cellular milieu. We posit that it is unlikely that cells resort to a collision-based delivery system for coenzymes to dependent enzymes but rather, use chaperones to shelter reactive coenzymes from adventitious side reactions and to escort them to dependent enzymes. A kin strategy is used in metabolic pathways for direct transfer of metabolites, *i.e.*, channeling in multifunctional enzymes and multienzyme complexes, a phenomenon that was first observed more than 50 years ago (1). The convergence of chemical, clinical, and biological approaches to the study of vitamin B₁₂ has helped unravel some of the complexities of the intracellular pathway for its assimilation and utilization and is reviewed here. A model is presented, which in its general features, could represent a paradigm for organic cofactor trafficking in mammals.

Chemical Characteristics of B₁₂. B₁₂ is a complex organometallic cofactor with a central cobalt atom tethered equatorially to four nitrogens donated by a tetrapyrrolic macrocycle, the corrin ring (2) (Figure 1).

ABSTRACT Many coenzymes are vitamins that are assimilated in mammals into their active form from precursors obtained from the diet. They are often both rare and reactive rendering the likelihood low that the cell uses a collision-based strategy for their delivery to dependent enzymes. In humans, there are only two known B₁₂ or cobalamin-dependent enzymes: methionine synthase and methylmalonyl-CoA mutase. However, the pathway for intracellular assimilation and utilization of this cofactor is complex as revealed by careful clinical analyses of fibroblasts from patients with disorders of cobalamin metabolism. In the recent past, six of the eight human genes involved in the B₁₂ pathway have been identified and these have yielded important insights into their roles. The recent literature on the encoded proteins is reviewed, and a model for intracellular B₁₂ trafficking is proposed in which B₁₂ is escorted to its target proteins in the cytoplasmic and mitochondrial compartments in complex with chaperones, thereby averting problems of dilution and adventitious side reactions.

*To whom correspondence should be addressed.

E-mail: rbanerjee1@unl.edu.

Received for review March 12, 2006
and accepted April 4, 2006

Published online April 21, 2006

10.1021/cb6001174 CCC: \$33.50

© 2006 American Chemical Society

TABLE 1. Vitamins that are coenzyme precursors

Vitamin	Coenzyme
Biotin (B ₇)	Biotin
Cobalamin (B ₁₂)	MeCbl and AdoCbl
Folic acid	H ₄ folate and its one-carbon derivatives
Nicotinamide, niacin (B ₃)	NADP ⁺ , NAD ⁺
Panθοthenate (B ₅)	Coenzyme A
Pyridoxine (B ₆)	Pyridoxal phosphate
Riboflavin (B ₂)	FAD, FMN
Thiamin (B ₁)	Thiamin pyrophosphate

The ring system has a number of peripheral ornamentations with the largest being a pendant base that serves as the α - or lower axial ligand. The identity of the base appended from the propanolamine side chain varies in Nature, and in cobalamins, it is the bulky group, dimethylbenzimidazole (DMB) (3). Diversity at the upper axial ligand site is also seen, and cyano, methyl, deoxyadenosyl, and aquo groups are found in vitamin B₁₂, methylcobalamin, 5'-deoxyadenosylcobalamin or coenzyme B₁₂, and aquocobalamin, respectively. Of these, methylcobalamin (MeCbl) and 5'-deoxyadenosylcobalamin (AdoCbl) are the active coenzyme forms utilized by the mammalian enzymes, methionine synthase and methylmalonyl-CoA mutase, respectively.

The oxidation state of the cobalt ion in cobalamin ranges from +1 to +3 and, with it, the preferred coordination state ranges from four- to six-coordinate. The active coenzyme forms are examples of alkyl cob(III)alamins and adopt a six-coordinate conformation at physiological pH, with DMB serving as the lower axial ligand (4, 5). This is referred to as the “base-on” state that can be shifted to the “base-off” state in solution by lowering the pH, which results in protonation of DMB (Figure 1, panel b). The pK_a values for this transition are ~2.13, 2.9, and 3.7 for aquocobalamin, MeCbl, and AdoCbl, respectively

(6). Interestingly, both mammalian B₁₂ enzymes bind their cofactor in yet another conformation, in which a histidine residue donated by the protein replaces DMB, which is referred to as the “base-off/His-on” conformation. Clearly, given the pK_a values for the “base-on” to “base-off” transition, the concentration of free “base-off” cobalamins at physiological pH is expected to be exceedingly low. In addition, alkylcobalamins are extremely photolabile due to the presence of a relatively weak cobalt–carbon bond with a bond dissociation energy of 30 kcal/mol in “base-on” AdoCbl and 37 kcal/mol in “base-on” MeCbl (7).

Cob(II)alamin is paramagnetic and has an unpaired electron in the d_{z²} orbital. It is an intermediate in the mutase-catalyzed reaction (Figure 2). It is also a side product of oxidative interception of cob(II)alamin in the methionine synthase-catalyzed reaction and represents an inactive species that is converted via reductive methylation to MeCbl (8). In adenosyltransferase, an enzyme that is involved in mitochondrial assimilation of B₁₂, cob(II)alamin is a substrate that is converted via reductive adenosylation to AdoCbl (9, 10). Although

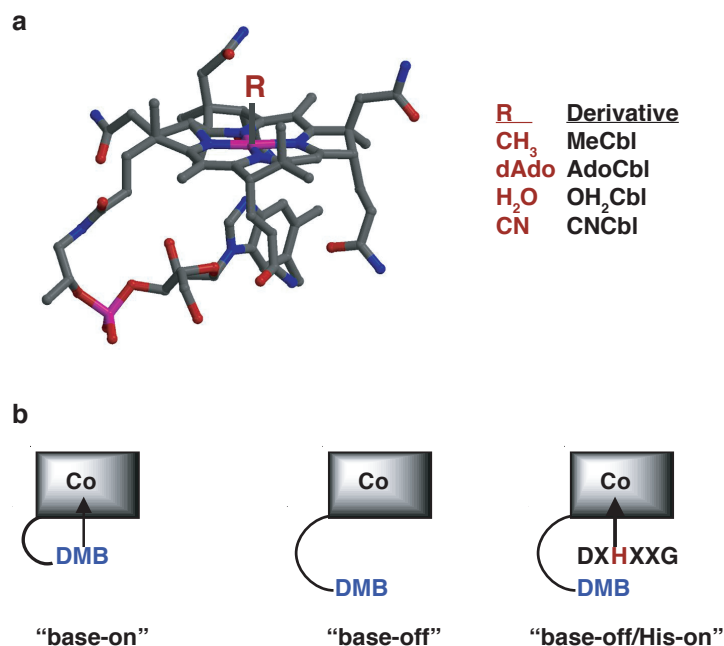


Figure 1. Structure and conformation of B₁₂. a) Crystal structure of “base-on” cobalamin in which the DMB ligand is coordinated to the cobalt at the lower axial position. The different R groups found at the upper axial position are indicated on the right. dAdo refers to 5'-deoxyadenosine. b) Alternate conformations of cobalamin seen in nature including “base-on”, “base-off”, and the “base-off/His-on” states.

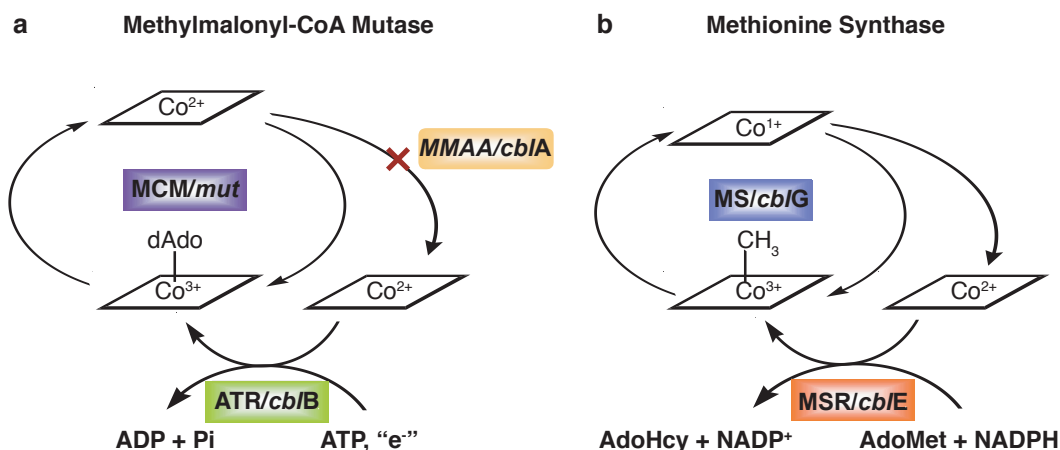


Figure 2. Oxidation states of cobalamin in reactions catalyzed by B_{12} -enzymes and auxiliary proteins. **a)** In the reaction catalyzed by methylmalonyl-CoA mutase (MCM), which is defective in the *mut* cell lines, cobalamin alternates between the cob(II)alamin and AdoCbl states. AdoCbl is synthesized from cob(II)alamin (or aquocobalmin) by adenosyltransferase (ATR), impaired in *cbB* patients. MMAA that is defective in *cbIA* cell lines protects against escape of cob(II)alamin from the catalytic turnover cycle of the mutase. **b)** In the reaction catalyzed by methionine synthase (MS), which is defective in *cbIG* patients, cobalamin cycles between the +1 and MeCbl states. Occasional oxidation leads to formation of inactive cob(II)alamin, which would accumulate were it not for a reductive methylation rescue in which methionine synthase reductase (MSR) that is defective in *cbIE* cell lines, provides electrons derived from NADPH. MSR contains FMN and FAD and accepts two electrons from NADPH while transferring one to MS.

cob(II)alamin is generally in the five-coordinate state (11), a novel four-coordinate cob(II)alamin is bound to adenosyltransferase (12). The pK_a for protonation of DMB in cob(II)alamin is estimated to be 3.1 (13).

Cob(I)alamin is an intermediate in the transmethylation reaction catalyzed by methionine synthase (Figure 2). It has been described as a “supernucleophile” and is one of the most reactive nucleophiles known in biology (14). In fact, its high reactivity underlies its sporadic oxidative interception in the methionine synthase-catalyzed reaction leading to generation of cob(II)alamin. Cob(I)alamin is also transiently generated during reductive activation of methionine synthase (15, 16) and, during conversion of cob(II)alamin to AdoCbl, catalyzed by adenosyltransferase (9) (Figure 2). It exhibits a square-planar four-coordinate geometry.

Cobalamin Uptake and Cellular Delivery. Cobalamin is a water-soluble vitamin and at physiological levels of intake (1–5 $\mu\text{g}/\text{day}$) is unable to traverse the plasma membrane unaided. Three B_{12} binding proteins, intrinsic factor, haptocorrin, and transcobalamin II, bind B_{12} with great avidity and mediate its uptake and transport (17). Haptocorrin is a salivary glycoprotein that binds B_{12} with higher affinity than intrinsic factor at the acidic pH of the stomach and liberates the cofactor

in the proximal small intestine upon being digested by pancreatic enzymes. Intrinsic factor is a secretory glycoprotein that binds B_{12} in the small intestine and in complex with a specific receptor, cubulin, located on the brush border of enterocytes, is endocytosed (17). B_{12} emerges from these cells into the portal circulation in complex with transcobalamin II (18, 19). Haptocorrin is also found in the serum where it is derived from leukocytes. It appears to function as a major storage protein for cobalamin and may play a role in removal of degraded derivatives (20, 21). All three proteins exhibit a very high affinity for B_{12} with K_d values for aquocob(III)alamin of 1 pM (intrinsic factor), 0.01 pM (haptocorrin), and 5 fM (transcobalamin II) measured at pH 7.5 and 20 $^\circ\text{C}$ (22). Intrinsic factor is very specific for cobalamin, *i.e.*, the cofactor form in which the lower axial DMB ligand is intact. This may represent an early screening mechanism for preventing degraded cobalamins from gaining cellular access.

Clinical Insights into the Complexity of the Intracellular B_{12} Trafficking. Insights into the complexities of the intracellular cobalamin assimilation pathway have emerged from careful clinical studies on patients who exhibit derangements in cobalamin metabolism, inherited as autosomal recessive

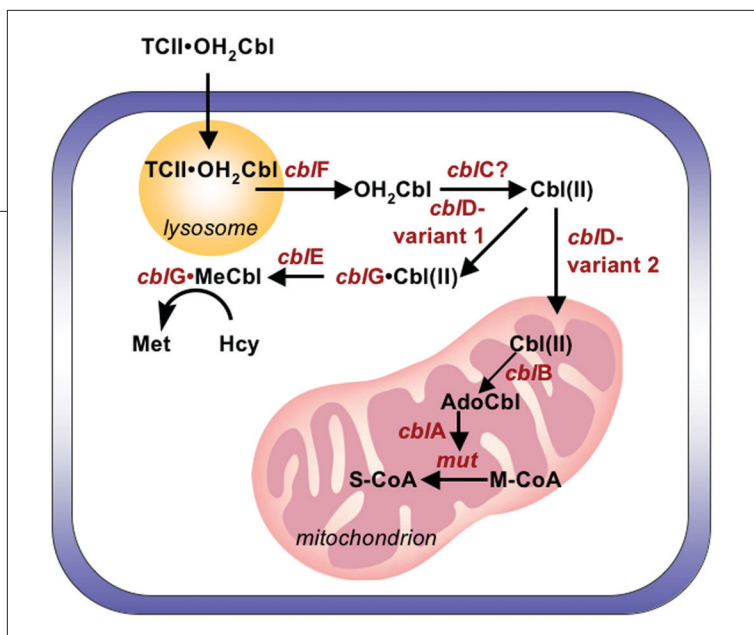


Figure 3. Disorders of intracellular cobalamin metabolism. The genetic complementation groups that have been described in patients are indicated in red (*cbIA-G* and *mut*). TCII denotes transcobalamin II and Met, Hcy, M-CoA, and S-CoA denote methionine, homocysteine, methylmalonyl-CoA, and succinyl-CoA, respectively.

disorders (23, 24) (Figure 3). There are only two known B₁₂ enzymes in mammals (25, 26): methionine synthase, a cytoplasmic enzyme that catalyzes the transfer of a methyl group from methyltetrahydrofolate to homocysteine to yield tetrahydrofolate and methionine (27), and methylmalonyl-CoA mutase, a mitochondrial enzyme that catalyzes the isomerization of methylmalonyl-CoA to succinyl-CoA (28). A functional deficiency in methionine synthase or methylmalonyl-CoA mutase leads to homocystinuria and methylmalonic aciduria, respectively (29, 30). Depending on the location of the defect in the pathway, malfunction of one or both B₁₂ enzymes ensues, resulting in either isolated or combined homocystinuria and methylmalonic aciduria. Clinical evaluations of these patients revealed heterogeneity and suggested that in many cases, loci other than the genes encoding the B₁₂-dependent enzymes were involved in the etiology of the disorder. A clever genetic approach has been used to uncover the underlying complexity and has led to delineation of at least eight distinct complementation groups (designated *cbIA-G* and *mut*) (29). Identification of the class of defect is based on the ability of patient fibroblasts to correct for the impairment in either the methionine synthase- or mutase-based incorporation of radiolabel (from [¹⁴C]-methyltetrahydrofolate and propionate, respectively) into acid-precipitable material in viral- (31) or polyethylene glycol-induced heterokaryons (32). If fusion of cells derived from two patients increases radiolabel incorporation under these conditions, it indicates that mutations are harbored at distinct genetic loci in

these cell lines. From such studies, the following three subclasses of cobalamin disorders emerged: those resulting (i) in combined defects in MeCbl and AdoCbl synthesis (*cbIF*, *cbIC*, and *cbID*), (ii) in an isolated defect in MeCbl synthesis (*cbID*-variant 1, *cbIE*, and *cbIG*), and (iii) in an isolated defect in AdoCbl synthesis (*cbID*-variant 2, *cbIA*, and *cbIB*). The *mut* locus encodes methylmalonyl-CoA mutase (Figure 3). An additional locus, *cbIH*, originally described as causing an isolated deficiency of methylmalonyl-CoA mutase (33), is now believed to be *cbID*-variant 2 (34).

Inherited Defects Affecting Both AdoCbl and MeCbl Synthesis. B₁₂ is ferried to cells with a very high affinity transporter, transcobalamin II, which binds to a specific receptor on the cell surface (35). This complex is endocytosed and delivered to the lysosome (Figure 3). The *cbIF* class of mutations impairs egress of B₁₂ from the lysosome and leads to the accumulation of the cofactor in this compartment (36). The affected gene in *cbIF* patients remains to be identified.

The function lost or impaired in *cbIC* patients was first suggested to be a cytoplasmic reductase that reduces cob(III)alamin to cob(II)alamin (37). In fact cob(III)alamin reductase activities associated with the microsomal and mitochondrial membranes that exhibited diminished activity in *cbIC* (and *cbID*) cell lines have been reported (38, 39). However, both locations seem improbable for delivery of the resulting cob(II)alamin to the target cytoplasmic and mitochondrial matrix enzymes, methionine synthase and adenosyltransferase (that synthesizes AdoCbl), respectively. It is possible that a nonspecific reductase activity was measured in these studies since the redox potential for the cob(III)alamin/cob(II)alamin redox couple is relatively high ($E^{\circ} = +240$ mV vs the standard hydrogen electrode (40)), making this a relatively facile reduction. The gene responsible for methylmalonic aciduria and homocystinuria, of the *cbIC* class has been identified recently (41) and as discussed below, based on its sequence, does not appear to be a reductase.

Alternatively, an NADPH and FAD-dependent β -ligand transferase function has been ascribed to the *cbIC* (or *cbID*) gene product that catalyzes the reductive exchange of the aquo or cyano substituent at the upper or β -axial position with glutathione to generate glutathionylcobalamin (39). However, the presence of glutathionylcobalamin in cells has not been established unequivocally since under conditions used for isolation,

i.e., deproteinized cobalamin in cell extracts containing glutathione (39, 42), detection of sulfitecobalamin (a breakdown product of glutathionyl-cobalamin), may be expected. Furthermore, the uncatalyzed conversion of aquocobalamin to glutathionylcobalamin is predicted to be rapid and essentially irreversible under physiological conditions (43), and the very relevance of cyanocobalamin in biology is uncertain. The presence of the cyano ligand found in vitamin B₁₂ that was initially crystallized was ascribed as an artifact of the isolation procedure that was employed (44). Although cyanocobalamin is the cofactor form used in many vitamin supplements today, in the absence of a known biological source of this derivative, the question as to whether mammalian cells would have evolved a system to specifically catalyze its decyanation is an open one. Thus, while *cbIC* is the most common of the inherited disorders in cobalamin utilization (45), the nature of its encoded function is presently unclear.

The *cbID* defect is enigmatic. Initially described in only two patients who were siblings, it was reported to result in a combined deficiency of methionine synthase and methylmalonyl-CoA mutase (32). As described above, limited biochemical characterization of these cell lines revealed similarities to *cbIC* cells, albeit with deficits that were less severe (37, 39). A more recent study on three additional patients revealed heterogeneity within the *cbID* mutant class, which has consequently been divided into variants 1 and 2 respectively (34). Variant 1 is associated with an isolated deficiency in MeCbl synthesis and variant 2 with an isolated defect in AdoCbl synthesis (Figure 3). On the basis of these observations, it has been suggested that a single *cbID* gene product “presents” B₁₂ to methionine synthase and to the mitochondrial membrane or that alternative splicing results in cytosolic and mitochondrial variants of the resulting protein (34). Mutations in *cbID* affecting a common domain, *viz.*, B₁₂ binding, would result in a combined deficiency of both cofactors whereas other mutations could impair interactions with one or the other protein target.

Inherited Defects Affecting Only MeCbl Synthesis.

The defects in the *cbIE* and *cbIG* cell lines have been localized to methionine synthase reductase (46–48) and methionine synthase (49–51), respectively. Methionine synthase belongs to the subfamily of B₁₂-dependent methyltransferases and the cofactor alternates between the MeCbl and cob(II)alamin

oxidation states during the reaction cycle (Figure 2). Occasional oxidation of the reactive cob(II)alamin intermediate is the source of the inactive cob(III)alamin form of the enzyme, which is repaired via a reductive methylation reaction to MeCbl (8, 52). The source of electrons for the reactivation reaction is NADPH, which is delivered via the dual flavoprotein oxidoreductase, methionine synthase reductase, to cob(II)alamin bound to methionine synthase (53). The activation reaction illustrates the sufficiency of the *cbIG* and *cbIE* loci for converting cob(III)alamin to the active MeCbl cofactor form and precludes requirement for additional cytosolic reductases/factors for this reaction.

Inherited Defects Affecting Only AdoCbl Synthesis.

The defect in the *cbIA* cell line was postulated to be a mitochondrial cob(II)alamin reductase (54, 55). However, the biochemical rationale for a stand-alone cob(II)alamin reductase is weak, since its product, cob(II)alamin, is highly reactive and would rapidly succumb to side reactions. Furthermore, cob(II)alamin bound to the protein encoded by the *cbIB* locus, adenosyltransferase, is reduced to cob(II)alamin *in situ*, in the first step of the reductive adenosylation reaction that generates AdoCbl (9). While the dual flavoprotein reductase, methionine synthase reductase, can serve as the reductase *in vitro* (10), the existence of a mitochondrial reductase dedicated for reduction of cob(II)alamin bound to adenosyltransferase appears unlikely based on the recent identification of all the mitochondrial specific *cbI* functions (10, 56, 57). Instead, a reductase that serves additional mitochondrial acceptors is likely used for the adenosyltransferase, explaining why it is not associated with a specific *cbI* defect. The recent identification of the locus harboring the *cbIA* defect (57) precludes a role for the encoded protein as an oxidoreductase as discussed below.

Proteins Involved in Intracellular B₁₂ Trafficking and Assimilation. Large gains in our understanding of the components of the intracellular B₁₂ pathway have occurred in recent years with the identification of genes associated with several of the *cbI* defects. In this section, the properties of the auxiliary *cbIE* and *cbIA-C* genes are discussed.

cbIE. Methionine synthase reductase is the locus of

KEYWORDS

Coenzyme: Organic cofactors of enzymes.

Chaperones: Protein porters for targeted delivery.

Cobalamin: Chemical name for B₁₂.

Homocysteine: A nonprotein sulfur-containing amino acid.

Since flavodoxin is absent in humans, this led to the discovery of methionine synthase reductase, a cytosolic dual flavoprotein oxidoreductase.

genetic defects in the *cbIE* complementation group (48). The susceptibility of the methionine synthase reaction to inactivation creates a dependency on a repair system that can rescue tightly bound oxidized cofactor to the turnover cycle (Figure 2). In bacteria, a two-component flavoprotein system comprising FMN-containing flavodoxin and FAD-containing NADPH-flavodoxin reductase transfer electrons to cob(II)alamin bound to methionine synthase (58, 59). Knowledge of the bacterial system guided the search for the presence of consensus FAD, FMN, and NADPH binding sites on a single polypeptide since flavodoxin is absent in humans, and this led to the discovery of methionine synthase reductase, a cytosolic dual flavoprotein oxidoreductase (48). Methionine synthase reductase has been purified to homogeneity and shown to be capable of reducing cob(II)alamin bound to methionine synthase, which is rapidly converted to MeCbl in the presence of the methyl group donor, *S*-adenosylmethionine (53, 60). Hence, the combination of methionine synthase and methionine synthase reductase is sufficient for generating the MeCbl form of the cofactor required for cytosolic B₁₂-dependent catalysis.

cbIC. The defective protein in *cbIC* patients is denoted MMACHC (since it causes methylmalonic aciduria type C and homocystinuria) and was discovered recently by a combination of linkage and haplotype analyses and homozygosity mapping (41). It encodes a soluble protein of unknown function that has 282 amino acid residues and does not appear to contain any known protein domains. Limited sequence conservation of the canonical B₁₂ binding motif found in the “base-off/His-on” family of B₁₂ enzymes is seen, suggesting that MMACHC can bind the cofactor in this conformation. The C-terminal third of MMACHC is reported to resemble the bacterial TonB protein (41) that is involved in vectorial transport of B₁₂ and other molecules across the outer and inner membranes in Gram-negative bacteria. TonB interacts with proteins that

contain a consensus Ton box motif and functions to open a hatch on the periplasmic side of the outer membrane transporter (BtuB for cobalamin) (61) thereby allowing translocation of the ligand through the membrane in to the periplasmic space.

cbIA. The defect in *cbIA* was widely expected to be in a cob(II)alamin reductase. However, identification of the genetic locus revealed a 418 residue long protein (including the mitochondrial leader sequence), designated MMAA (for causing methylmalonic aciduria type A) (57). It belongs to the G3E family of P-loop GTPases and contains the signature motifs associated with this superfamily: Walker A and Walker B motifs, an aspartate residue involved in Mg²⁺-binding, and a GTP-binding [N/T]KxD sequence (62). A bacterial ortholog of MMAA, MeaB, which is strongly conserved in operons encoding methylmalonyl-CoA mutase, is better characterized (63, and unpublished data). MeaB and methylmalonyl-CoA mutase form a stable complex that can be isolated on a native gel (63). MeaB exhibits low intrinsic GTPase activity that is activated >100-fold in complex with the mutase (unpublished data). MeaB in turn influences the mutase, protecting it from oxidative inactivation, but only in the presence of nucleotides (Figure 2) (unpublished data). In addition, MeaB increases the *k*_{cat} of the mutase 2-fold and, interestingly, the presence of nucleotides is without further effect on rate enhancement by MeaB. Although the concentration of AdoCbl is low in *cbIA* cell lines (64), AdoCbl synthesis in extracts from these cells is comparable to that in control cell lines (54). These results suggest that in addition to the mutase, MeaB interacts with adenosyltransferase and influences AdoCbl synthesis in mitochondria.

cbIB. As expected from biochemical studies on *cbIB* cell lines, the defect in them resides in the gene encoding adenosyltransferase also designated MMAB (for causing methylmalonic aciduria type B) (10, 56). Unlike methionine synthase that is the site for MeCbl synthesis in the presence of the auxiliary protein, methionine synthase reductase, inactive cofactor bound to methylmalonyl-CoA mutase cannot be converted to AdoCbl in situ (Figure 2). This activity is catalyzed by adenosyltransferase that converts cob(II)alamin or aquocobalamin in the presence of a reductant to AdoCbl. The human enzyme is trimeric and binds the various cobalamin derivatives with moderate affinity (~2–8 μM) (64). Spectroscopic analyses have revealed an unusual coordination environment for the cofactor bound to the enzyme, which is “base-off” in both the substrate cob(II)alamin and product AdoCbl oxidation states (64, 66). In the presence of ATP, four-coordinate cob(II)alamin (*i.e.*, lacking a lower water ligand) with very unusual electron paramagnetic resonance and

KEYWORDS

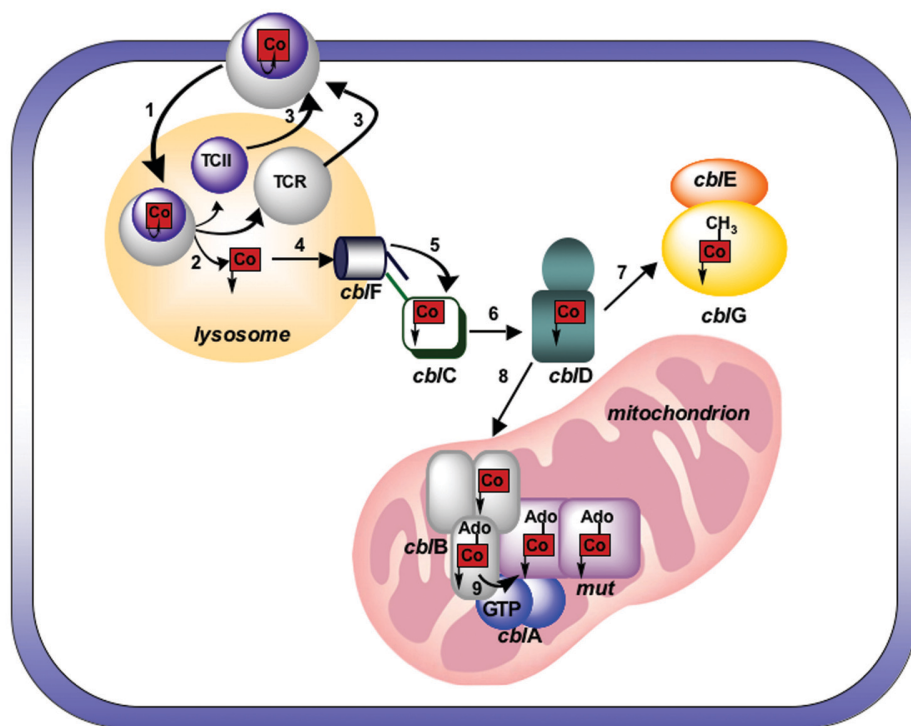
Methylmalonic aciduria: Accumulation of methylmalonic acid in urine.

cbI defects: Inherited disorders in B₁₂ metabolism belonging to distinct genetic complementation groups.

Hyperhomocysteinemia: Accumulation of homocysteine in plasma.

Inborn errors: Inherited defects leading to metabolic deficits.

Figure 4. Proposed model for intracellular cobalamin trafficking and utilization. Cobalamin bound to transcobalamin II is endocytosed (1) into the lysosome where the cofactor is released (2) while TCII and its receptor, TCR, are recycled back to the membrane (3). In the acidic environment of the lysosome, the concentration of “base-off” cobalamin is increased and this form is transported out via an unidentified transporter (4) that is defective in *cbIF* cell lines. MMAHC is shown as opening the hatch of the lysosomal transporter (5) and binding the released cobalamin, which is then transferred to the protein associated with the *cbID* defect (6), which transfers the cofactor to its cytoplasmic (7) and mitochondrial (8) targets. In the mitochondrion, adenosyltransferase (shown as a trimer) transfers its product, AdoCbl (9) directly to methylmalonyl-CoA mutase. MMAA, which has GTPase activity, is involved in this process, although the exact mechanism is not known.



magnetic circular dichroism spectroscopic properties is observed (66).

Hypothesis for Chaperoned Intracellular B₁₂ Delivery.

Three major challenges for cellular utilization of B₁₂ are its low abundance (intracellular concentrations are estimated to range from 0.03 to 0.7 μM depending on the organ (67)), its reactivity in all three oxidation states as discussed above, and its delivery to cells in the “base-on” conformation whereas the intracellular target enzymes bind the cofactor in the “base-off” conformation that is exceedingly rare at physiological pH. In the following section we propose a hypothesis for how cells might meet these challenges based on the components of the intracellular metabolic pathway for B₁₂ that have come to light in recent years although their biochemical details remain sketchy (Figure 4). The crystal structure of transcobalamin II reveals that it binds B₁₂ in the “base-on” conformation (68) and delivers it to the lysosomal compartment where the acidic pH should increase the concentration of the “base-off” species. We propose that the cofactor exits the lysosome in the “base-off” conformation via an unidentified transporter that is defective in *cbIF* patients (36) and that from this point onward remains bound to intracellular porters that maintain this

functionally important conformation that is sparsely populated in solution at physiological pH.

Reduction of aquocobalamin to cob(II)alamin in the lysosomal compartment would further enhance the concentration of the “base-off” form of the cofactor by virtue of an increase in the pK_a for protonation of DMB by ~5 orders of magnitude (from -2.1 to +3.1 (6, 13)). The redox potential for the aquocob(III)alamin/cob(II)alamin is high in both the “base-on” (+200 mV) and “base-off” (+510 mV) states (40), making reduction of either conformer within relatively easy reach of cellular reductants.

On the basis of its homology to TonB (41), it is tempting to speculate that MMACHC may function to open a hatch in the lysosomal transporter thereby promoting release of B₁₂ to the cytoplasmic compartment. The presence of the remnants of a B₁₂-binding motif in MMACHC further suggests the possibility that this protein itself is the proximal carrier of the cofactor released from the lysosome, as shown in Figure 4. Alternatively, the next protein in the pathway, which is unidentified but defective in *cbID* patients, could represent the first cytoplasmic B₁₂ binder and MMACHC may simply function as a sentry gating entry of the cofactor into the cytoplasmic compartment.

The recent description of two variants of the *cbiD* locus supports the model that the encoded protein interacts with targets in both the cytoplasmic (*variant-1*) and mitochondrial (*variant-2*) compartments (34). *Variant-1* may directly transfer bound B₁₂ directly to methionine synthase, which can catalyze the conversion of oxidized cofactor to the active MeCbl form in a reductive methylation reaction that relies upon methionine synthase reductase for the injection of an electron, and S-adenosylmethionine as a source of the methyl group. *Variant-2* could be an alternatively spliced form that is present in the cytoplasm or the mitochondrion. In the former case, it could function as a cytoplasmic porter that delivers bound cofactor to the mitochondrion. If *variant-2* is a mitochondrial protein (34), it could accept B₁₂ from MMACHC and deliver it to the mitochondrial target, adenosyltransferase. Alternatively, *variants 1* and *2* could represent mutations in two different domains in a single modular protein that interacts with the cytoplasmic and mitochondrial trafficking machinery, respectively.

The mechanism by which B₁₂ traverses the mitochondrial membranes and reaches adenosyltransferase in the matrix is not known, and a specific B₁₂ transporter may not be involved. In fact it has been suggested that mitochondrial uptake of B₁₂ may occur by passive transport (69). We propose that the three mitochondrial proteins in this pathway exist as a complex. Synthesis of AdoCbl from cob(II)alamin, the cofactor form presumed to be delivered to this organelle, by adenosyltransferase is postulated to be modulated by the G-protein, MMAA, although the mechanism by which it exerts its influence is presently not known. One possibility is that MMAA gates release of AdoCbl from adenosyltransferase and its direct transfer to methylmalonyl-CoA mutase is powered by GTP hydrolysis. MMAA also plays a role in protecting the mutase from oxidative inactivation (unpublished data). In contrast to the cytoplasmic system in which methionine synthase reductase-dependent reductive methylation intervenes to repair inactive methionine synthase, the mitochondrial system uses a preventative strategy employing MMAA to protect the radical enzyme, methylmalonyl-CoA mutase, from inactivation (Figure 2).

We have postulated that the mirror “base-off” conformation of B₁₂ that is seen in adenosyltransferase and the mutase together with the lower coordination environment (five-coordinate in the former and

six-coordinate in the latter) for bound AdoCbl is likely to be functionally significant (70). The histidine in the DXHXXG motif that coordinates B₁₂ in the mutase is predicted, based on comparisons with the related glutamate mutase, to be on an unstructured loop in the absence of cofactor and may be important in the direct transfer of AdoCbl from adenosyltransferase to the methylmalonyl-CoA mutase (71, 72). Similarities to this model in which lower (or incomplete) coordination of the metal in the donor chaperone is exploited to achieve facile transfer to the target metalloprotein are seen in the yeast copper chaperone, Atx1(73), and the human iron–sulfur cluster scaffold protein Nfu (74).

This model provides a strategy for increasing the cellular concentration of “base-off” B₁₂ that is required for its binding to target enzymes, by proposing that the cofactor exits the acidic environment of the lysosome and, thereafter, remains protein bound in this conformation. The use of porters also averts the problem of dilution that would result inevitably from release of the cofactor into solution and, furthermore, protects it from myriad potential ligands (not least of all thiols) that could trap the cofactor in unproductive complexes. Escorted delivery of B₁₂ is consistent with the observation that >95% of intracellular cobalamin is bound and that it is associated predominantly with methionine synthase and methylmalonyl-CoA mutase (29, 75). It is important to note that a subclass of patients with inborn errors of cobalamin metabolism is responsive to B₁₂ treatment. The amelioration of the clinical phenotype can be reconciled with the proposed model for B₁₂ trafficking as follows. If the defect in one or both mutant alleles results in diminished affinity for the bound cofactor, increasing the concentration of B₁₂ would circumvent this defect. Indeed, this is proposed to be the basis of the *mut*⁻ class of mutations in methylmalonyl-CoA mutase, which is supported experimentally (76). However, if the defect in both mutant alleles leads to the complete absence of the encoded protein, for instance due to alternative splicing or premature truncation, then the responsiveness to large doses of B₁₂ could result from increased intracellular concentrations bypassing the defective chaperone by direct binding to the target enzymes or to a downstream escort in the pathway.

A cytoplasmic cob(II)alamin reductase has been repeatedly invoked in the literature based on activity assays in cell extracts and its diminution in both *cbiC*

and *cbiD* cell lines (38, 39) and is a component that is not directly addressed in this model. Two redox couples are biologically relevant for cobalamin. The cob(III)alamin/cob(II)alamin redox potential is high (+240 mV (40)) and represents a reduction that can be achieved relatively easily in the intracellular milieu. In contrast the cob(II)alamin/cob(I)alamin potential is very low (−610 mV for the base-on form and −500 mV for the “base-off” form (40)) and represents a challenging reduction for biological reducing systems in aerobic organisms. The synthesis of both MeCbl and AdoCbl from cob(II)alamin proceeds via a cob(I)alamin intermediate that is generated transiently within the protective confines of the active sites and alkylated, essentially irreversibly, by coupling to exergonic methylation or adenosylation reactions, respectively. The sequence of *MMACHC* does not suggest that it is a cob(III)alamin reductase, but this issue awaits direct biochemical evaluation. Alternatively, the protein associated with the *cbiD* complementation group could function as a dedicated cob(III)alamin reductase while also transporting the cofactor to its downstream targets, or a non-specific lysosomal or cytoplasmic reductase could be employed for this step. As discussed earlier, lysosomal reduction of aquocobalamin to cob(II)alamin offers the advantage of significantly increasing the concentration of the “base-off” conformer of cob(II)alamin in this acidic compartment.

The model is parsimonious with respect to invoking additional B₁₂ derivatives and enzymes for handling

such intermediates in the pathway. Of note is glutathionylcobalamin, which would form rapidly and essentially irreversibly if the cofactor were released in solution in the presence of 1–10 mM glutathione that is found in most cell types (43). Alternatively, a specific β-ligand transferase could catalyze this transformation as suggested in the literature (39). While our model does not include these players, it does not rule out the existence of additional complexity either.

Summary. An emerging theme in metal cofactor trafficking is the deployment of a battery of chaperones that pick up and deliver metals from their sites of entry to their sites of utilization and thereby contribute to exceedingly low intracellular concentrations of free metals (77). The rationale for sequestering transition and other metals is obvious, it reduces the scope for side reactions and achieves specificity by ensuring that the correct metal is inserted. The same line of reasoning predicates intracellular systems for organic cofactor trafficking. However, unlike the expanding field of metal chaperones, porters for organic cofactors, if they exist, are largely unknown. Our model for B₁₂ trafficking is informed by recent strides in identification of the components along with the biochemical elucidation of their properties. We posit that escorts similarly target other coenzymes to dependent enzymes, although our insights into these systems are even more scant and often limited to uptake studies (78–81).

Acknowledgment: This work was supported in part by a grant from the National Institutes of Health (DK45776).

REFERENCES

1. Cori, C. F., Velick, S. F., and Cori, G. T. (1950) The combination of diphosphopyridine nucleotide with glyceraldehyde phosphate dehydrogenase, *Biochim. Biophys. Acta* 4, 160–169.
2. Hodgkin, D. C., Kamper, J., Mackay, M., Pickworth, J. W., Trueblood, K. N., and White, J. G. (1956) Structure of vitamin B₁₂, *Nature (London)* 178, 64–66.
3. Kräutler, B. (2005) Vitamin B₁₂: chemistry and biochemistry, *Biochem. Soc. Trans.* 33, 806–810.
4. Lenhert, P. G., and Hodgkin, D. C. (1961) Structure of the 5,6-dimethylbenzimidazolylcobamide coenzyme, *Nature (London)* 192, 937–938.
5. Rossi, M., Glusker, J. P., Randaccio, L., Summers, M. F., Toscano, P. J., and Marzilli, L. G. (1985) The structure of a B₁₂ coenzyme: Methylcobalamin studies by X-ray and NMR methods, *J. Am. Chem. Soc.* 107, 1729–1738.
6. Brown, K. L., and Hakimi, J. M. (1984) Heteronuclear NMR studies of cobalamins. 3. ³¹P NMR of aquocobalamin and various organocobalamins, *J. Am. Chem. Soc.* 106, 7894–7899.
7. Waddington, M. D., and Finke, R. G. (1993) Neopentylcobalamin (NeopentylB₁₂) Cobalt–Carbon Bond Thermolysis Products, Kinetics, Activation Parameters and Bond Dissociation Energy: A Chemical Model Exhibiting 10⁶ of the 10¹² Enzymic Activation of Coenzyme B₁₂'s Cobalt–Carbon Bond, *J. Am. Chem. Soc.* 115, 4629–4640.
8. Banerjee, R. V., and Matthews, R. G. (1990) Cobalamin-dependent methionine synthase, *FASEB J.* 4, 1450–1459.
9. Leal, N. A., Park, S. D., Kima, P. E., and Bobik, T. A. (2003) Identification of the human and bovine ATP:Cob(II)alamin adenosyltransferase cDNAs based on complementation of a bacterial mutant, *J. Biol. Chem.* 278, 9227–9234.
10. Leal, N. A., Olteanu, H., Banerjee, R., and Bobik, T. A. (2004) Human ATP:Cob(II)alamin adenosyltransferase and its interaction with methionine synthase reductase, *J. Biol. Chem.* 279, 47536–47542.
11. Kräutler, B., Keller, W., and Kratky, C. J. (1989) Coenzyme B₁₂ chemistry: The crystal and molecular structure of cob(II)alamin, *J. Am. Chem. Soc.* 111, 8936–8938.
12. Stich, T. A., Buan, N. R., Escalante-Semerena, J. C., and Brunold, T. C. (2005) Spectroscopic and computational studies of the ATP:corrinoid adenosyltransferase (CobA) from *Salmonella enterica*: insights into the mechanism of adenosylcobalamin biosynthesis, *J. Am. Chem. Soc.* 127, 8710–8719.
13. Brown, K. L., and Zou, X. (1991) Facile α/β diastereomerism in organocobalt corrins. Evidence for thermodynamic control in the synthesis of alkylcobamides, *J. Inorg. Chem.* 30, 4185–4192.
14. Schrauzer, G. N., and Deutsch, E. (1969) Reactions of cobalt(II) supernucleophiles. The alkylation of vitamin B₁₂'s, cobaloximes(II), and related compounds, *J. Am. Chem. Soc.* 91, 3341–3350.

15. Banerjee, R. V., Harder, S., Ragsdale, S. W., and Matthews, R. G. (1990) Mechanism of reductive activation of cobalamin-dependent methionine synthase: An electron paramagnetic resonance spectro-electrochemical study, *Biochemistry* 29, 1129–1137.
16. Jarrett, J. T., Hoover, D. M., Ludwig, M. L., and Matthews, R. G. (1998) The mechanism of adenosylmethionine-dependent activation of methionine synthase: A rapid kinetic analysis of intermediates in reductive methylation of Cob(II)alamin enzyme, *Biochemistry* 37, 12649–12658.
17. Seetharam, B. (1999) Receptor-mediated endocytosis of cobalamin (vitamin B₁₂), *Annu. Rev. Nutr.* 19, 173–195.
18. Chanarin, I., Muir, M., Hughes, A., and Hoffbrand, A. V. (1978) Evidence for intestinal origin of transcobalamin II during vitamin B₁₂ absorption, *Br. Med. J.* 1, 1453–1455.
19. Seetharam, B., Bose, S., and Li, N. (1999) Cellular import of cobalamin (Vitamin B₁₂), *J. Nutr.* 129, 1761–1764.
20. Kanazawa, S., Herbert, V., Herzlich, B., Drivas, G., and Manusselis, C. (1983) Removal of cobalamin analogue in bile by enterohepatic circulation of vitamin B₁₂, *Lancet* 1, 707–708.
21. Allen, R. H. (1975) Human vitamin B₁₂ transport proteins, *Prog. Hematol.* 9, 57–84.
22. Fedosov, S. N., Berglund, L., Fedosova, N. U., Nexø, E., and Petersen, T. E. (2002) Comparative analysis of cobalamin binding kinetics and ligand protection for intrinsic factor, transcobalamin, and haptocorrin, *J. Biol. Chem.* 277, 9989–9996.
23. Rosenberg, L. E. (1983) *The Metabolic Basis of Inherited Disease*, 5th ed. pp 474–497, McGraw-Hill, New York.
24. Rosenblatt, D. S., and Fenton, W. A. (2001) in *The Metabolic & Molecular Bases of Inherited Diseases* (Scriver, C. R., and Sly, W. S., Eds.) pp 3897–3933, McGraw-Hill, New York.
25. Kolhouse, J. F., and Allen, R. H. (1977) Recognition of two intracellular cobalamin binding proteins and their recognition as methylmalonyl-CoA mutase and methionine synthetase, *Proc. Natl. Acad. Sci. U.S.A.* 74, 921–925.
26. Mellman, I. S., Youngdahl-Turner, P., Huntington, F. W., and Rosenberg, L. E. (1977) Intracellular binding of radioactive hydroxocobalamin to cobalamin-dependent apoenzyme in rat liver, *Proc. Natl. Acad. Sci. U.S.A.* 74, 916–920.
27. Ludwig, M. L., and Matthews, R. G. (1997) Structure-based perspectives on B₁₂-dependent enzymes, *Annu. Rev. Biochem.* 66, 269–313.
28. Banerjee, R., and Ragsdale, S. W. (2003) The many faces of vitamin B₁₂: Catalysis by cobalamin-dependent enzymes, *Ann. Rev. Biochem.* 72, 209–247.
29. Shevell, M. I., and Rosenblatt, D. S. (1992) The neurology of cobalamin, *Can. J. Neurol. Sci.* 19, 472–486.
30. Chandler, R. J., and Venditti, C. P. (2005) Genetic and genomic systems to study methylmalonic acidemia, *Mol. Genet. Metab.* 86, 34–43.
31. Gravel, R. A., Mahoney, M. J., Ruddle, F. H., and Rosenberg, L. E. (1975) Genetic complementation in heterokaryons of human fibroblasts defective in cobalamin metabolism, *Proc. Natl. Acad. Sci. U.S.A.* 72, 3181–3185.
32. Willard, H. F., Mellman, I. S., and Rosenberg, L. E. (1978) Genetic complementation among inherited deficiencies of methylmalonyl-CoA mutase activity: Evidence for a new class of human cobalamin mutant, *Am. J. Hum. Genet.* 30, 1–13.
33. Watkins, D., Matiaszuk, N., and Rosenblatt, D. S. (2000) Complementation studies in the *cbIA* class of inborn error of cobalamin metabolism: Evidence for interallelic complementation and for a new complementation class (*cbIH*), *J. Med. Genet.* 37, 510–513.
34. Suormala, T., Baumgartner, M. R., Coelho, D., Zavadakova, P., Koich, V., Koch, H. G., Berghauer, M., Wraith, J. E., Burlina, A., Sewell, A., Herwig, J., and Fowler, B. (2004) The *cbID* defect causes either isolated or combined deficiency of methylcobalamin and adenosylcobalamin synthesis, *J. Biol. Chem.* 279, 42742–42749.
35. Youngdahl-Turner, P., Rosenberg, L. E., and Allen, R. H. (1978) Binding and uptake of transcobalamin II by human fibroblasts, *J. Clin. Invest.* 61, 133–141.
36. Rosenblatt, D. S., Hosack, A., Matiaszuk, N. V., Cooper, B. A., and Laframboise, R. (1985) Defect in vitamin B₁₂ release from lysosomes: Newly described inborn error of vitamin B₁₂ metabolism, *Science* 228, 1319–1321.
37. Mellman, I. S., Huntington, F. W., Youngdahl-Turner, P., and Rosenberg, L. E. (1979) Cobalamin coenzyme synthesis in normal and mutant human fibroblast. Evidence for a processing enzyme activity deficient in *cbIC* cells, *J. Biol. Chem.* 254, 11847–11853.
38. Watanabe, F., Saïdo, H., Yamaji, R., Miyatake, K., Isegawa, Y., Ito, A., Yubisui, T., Rosenblatt, D. S., and Nakano, Y. (1996) Mitochondrial NADH- or NADPH-linked aquacobalamin reductase activity is low in human skin fibroblasts with defects in synthesis of cobalamin coenzymes, *J. Nutr.* 126, 2947–2951.
39. Pezacka, E. H. (1993) Identification and Characterization of two enzymes involved in the intracellular metabolism of cobalamin. Cyanocobalamin β -ligand transferase and microsomal cob(II)alamin reductase, *Biochim. Biophys. Acta* 1157, 167–177.
40. Lexa, D., and Saveant, J.-M. (1983) The electrochemistry of vitamin B₁₂, *Acc. Chem. Res.* 16, 235–243.
41. Lerner-Ellis, J. P., Tirone, J. C., Pawelek, P. D., Dore, C., Atkinson, J. L., Watkins, D., Morel, C. F., Fujiwara, T. M., Moras, E., Hosack, A. R., Dunbar, G. V., Antonicka, H., Forgetta, V., Dobson, C. M., Leclerc, D., Gravel, R. A., Shoubridge, E. A., Coulton, J. W., Lepage, P., Rommens, J. M., Morgan, K., and Rosenblatt, D. S. (2006) Identification of the gene responsible for methylmalonic aciduria and homocystinuria, *cbIC* type, *Nat. Genet.* 38, 93–100.
42. Pezacka, E. H., Green, R., and Jacobsen, D. W. (1990) Glutathionylcobalamin as an intermediate in the formation of cobalamin coenzymes, *Biochem. Biophys. Res. Commun.* 169, 443–450.
43. Xia, L., Cregan, A. G., Berben, L. A., and Brasch, N. E. (2004) Studies on the formation of glutathionylcobalamin: Any free intracellular aquacobalamin is likely to be rapidly and irreversibly converted to glutathionylcobalamin, *Inorg. Chem.* 43, 6848–6857.
44. Abeles, R. H., and Dolphin, D. (1976) The vitamin B₁₂ coenzyme, *Acc. Chem. Res.* 9, 114–120.
45. Fenton, W. A., Gravel, R. A., and Rosenblatt, D. S. (2001) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.) pp 2165–2193, McGraw-Hill, Inc., New York.
46. Watkins, D., and Rosenblatt, D. S. (1989) Functional Methionine Synthase Deficiency (*cbIE* and *cbIG*): Clinical and Biochemical Heterogeneity, *Am. J. Med. Genet.* 34, 427–434.
47. Gulati, S., Brody, L. C., Rosenblatt, D. S., and Banerjee, R. (1997) Defects in auxiliary redox proteins lead to functional methionine synthase deficiency, *J. Biol. Chem.* 272, 19171–19175.
48. Leclerc, D., Wilson, A., Dumas, R., Gafuik, C., Song, D., Watkins, D., Heng, H. H. Q., Rommens, J. M., Scherer, S. W., Rosenblatt, D. S., and Gravel, R. A. (1998) Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria, *Proc. Natl. Acad. Sci. U.S.A.* 95, 3059–3064.
49. Sillaots, S. L., Hall, C. A., Hurteloup, V., and Rosenblatt, D. S. (1992) Heterogeneity in *cbIG*: Differential Retention of Cobalamin on Methionine Synthase, *Biochem. Med. Metab. Biol.* 47, 242–249.
50. Gulati, S. G., Baker, P., Fowler, B., Li, Y., Kruger, W., Brody, L. C., and Banerjee, R. (1996) Mutations in human methionine synthase in *cbIG* patients, *Hum. Mol. Genet.* 5, 1859–1866.
51. Leclerc, D., Campeau, E., Goyette, P., Adjalla, C. E., Christensen, B., Ross, M., Eydoux, P., Rosenblatt, D. S., Rozen, R., and Gravel, R. A. (1996) Human methionine synthase: cDNA cloning and identification of mutations in patients of the *cbIG* complementation group of folate/cobalamin disorders, *Hum. Mol. Genet.* 5, 1867–1874.
52. Matthews, R. G. (2001) Cobalamin-dependent methyltransferases, *Acc. Chem. Res.* 34, 681–689.
53. Olteanu, H., and Banerjee, R. (2001) Human methionine synthase reductase, a soluble P450 reductase-like dual flavoprotein, is sufficient for methionine synthase activation, *J. Biol. Chem.* 276, 35558–35563.

54. Mahoney, M. J., Hart, A. C., Steen, V. D., and Rosenberg, L. E. (1975) Methylmalonicacidemia: biochemical heterogeneity in defects of 5'-deoxyadenosylcobalamin synthesis, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2799–2803.
55. Fenton, W. A., and Rosenberg, L. E. (1978) Genetic and biochemical analysis of human cobalamin mutants in cell culture, *Annu. Rev. Genet.* **12**, 223–248.
56. Dobson, C. M., Wai, T., Leclerc, D., Kadir, H., Narang, M., Lerner-Ellis, J. P., Hudson, T. J., Rosenblatt, D. S., and Gravel, R. A. (2002) Identification of the gene responsible for the *cbIB* complementation group of vitamin B₁₂-dependent methylmalonic aciduria, *Hum. Mol. Genet.* **11**, 3361–3369.
57. Dobson, C. M., Wai, T., Leclerc, D., Wilson, A., Wu, X., Dore, C., Hudson, T., Rosenblatt, D. S., and Gravel, R. A. (2002) Identification of the gene responsible for the *cbIA* complementation group of vitamin B₁₂-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15554–15559.
58. Fujii, K., and Huennekens, F. M. (1979) Methionine Synthase: Characterization of protein components and mechanisms for activation and catalysis, 173-183, Japan Scientific Societies, Tokyo.
59. Fujii, K., Galivan, J. H., and Huennekens, F. M. (1977) Activation of methionine synthase: Further characterization of the flavoprotein system, *Arch. Biochem. Biophys.* **178**, 662–670.
60. Olteanu, H., Wolthers, K. R., Munro, A. W., Scrutton, N. S., and Banerjee, R. (2004) Kinetic and thermodynamic characterization of the common polymorphic variants of human methionine synthase reductase, *Biochemistry* **43**, 1988–1997.
61. Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nat. Struct. Biol.* **10**, 394–401.
62. Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases, *J. Mol. Biol.* **317**, 41–72.
63. Korotkova, N., and Lidstrom, M. E. (2004) MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation, *J. Biol. Chem.* **279**, 13652–13658.
64. Lerner-Ellis, J. P., Dobson, C. M., Wai, T., Watkins, D., Tirone, J. C., Leclerc, D., Dore, C., Lepage, P., Gravel, R. A., and Rosenblatt, D. S. (2004) Mutations in the *MMAA* gene in patients with the *cbIA* disorder of vitamin B₁₂ metabolism, *Hum. Mutat.* **24**, 509–516.
65. Yamanishi, M., Labunska, T., and Banerjee, R. (2005) Mirror “base-off” conformation of coenzyme B₁₂ in human adenosyltransferase and its downstream target, methylmalonyl-CoA mutase, *J. Am. Chem. Soc.* **127**, 526–527.
66. Stich, T. A., Yamanishi, M., Banerjee, R., and Brunold, T. C. (2005) Spectroscopic evidence for the formation of a four-coordinate Co²⁺ cobalamin species upon binding to the human ATP:cobalamin adenosyltransferase, *J. Am. Chem. Soc.* **127**, 7660–7661.
67. Hsu, J. M., Kawin, B., Minor, P., and Mitchell, J. A. (1966) Vitamin B₁₂ concentrations in human tissue, *Nature* **210**, 1264–1265.
68. Wuerges, J., Garau, G., Geremia, S., Fedosov, S. N., Petersen, T. E., and Randaccio, L. (2006) Structural basis for mammalian B₁₂ transport by transcobalamin, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4386–4391.
69. Qureshi, A. A., Rosenblatt, D. S., and Cooper, B. A. (1994) Inherited disorders of cobalamin metabolism, *Crit. Rev. Oncol. Hematol.* **17**, 133–151.
70. Yamanishi, M., Vlasie, M., and Banerjee, R. (2005) Adenosyltransferase: an enzyme and an escort for coenzyme B₁₂?, *Trends Biochem. Sci.* **30**, 304–308.
71. Tollinger, M., Eichmuller, C., Konrat, R., Huhta, M. S., Marsh, E. N., and Krautler, B. (2001) The B(12)-binding subunit of glutamate mutase from *Clostridium tetanomorphum* traps the nucleotide moiety of coenzyme B(12), *J. Mol. Biol.* **309**, 777–791.
72. Tollinger, M., Konrat, R., Hilbert, B. H., Marsh, E. N. G., and Krautler, B. (1998) How a protein prepares for B₁₂ binding: structure and dynamics of the B₁₂-binding subunit of glutamate mutase from *Clostridium tetanomorphum*, *Structure* **6**, 1021–1033.
73. Pufahl, R. A., Singer, C. P., Peariso, K. L., Lin, S. J., Schmidt, P. J., Fahmi, C. J., Culotta, V. C., Penner-Hahn, J. E., and O'Halloran, T. V. (1997) Metal ion chaperone function of the soluble Cu(I) receptor Atx1, *Science* **278**, 853–856.
74. Tong, W. H., Jameson, G. N., Huynh, B. H., and Rouault, T. A. (2003) Subcellular compartmentalization of human Nfu, an iron-sulfur cluster scaffold protein, and its ability to assemble a [4Fe-4S] cluster, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9762–9767.
75. Rosenblatt, D. S., and Cooper, B. A. (1987) Inherited disorders of vitamin B₁₂ metabolism, *Blood Rev.* **1**, 177–182.
76. Crane, A. M., and Ledley, F. D. (1994) Clustering of mutations in methylmalonyl-CoA mutase Associated with *mut-* methylmalonic acidemia, *Am. J. Hum. Genet.* **55**, 42–50.
77. Kuchar, J., and Hausinger, R. P. (2004) Biosynthesis of metal sites, *Chem. Rev.* **104**, 509–525.
78. Marchant, J. S., Subramanian, V. S., Parker, I., and Said, H. M. (2002) Intracellular trafficking and membrane targeting mechanisms of the human reduced folate carrier in mammalian epithelial cells, *J. Biol. Chem.* **277**, 33325–33333.
79. Subramanian, V. S., Marchant, J. S., Parker, I., and Said, H. M. (2003) Cell biology of the human thiamine transporter-1 (hTHTR1). Intracellular trafficking and membrane targeting mechanisms, *J. Biol. Chem.* **278**, 3976–3984.
80. Foraker, A. B., Khantwal, C. M., and Swaan, P. W. (2003) Current perspectives on the cellular uptake and trafficking of riboflavin, *Adv. Drug Deliv. Rev.* **55**, 1467–1483.
81. Phelps, M. A., Foraker, A. B., Gao, W., Dalton, J. T., and Swaan, P. W. (2004) A novel rhodamine-riboflavin conjugate probe exhibits distinct fluorescence resonance energy transfer that enables riboflavin trafficking and subcellular localization studies, *Mol. Pharm.* **1**, 257–266.

Enhancement of the Immunogenicity of Synthetic Carbohydrates by Conjugation to Virosomes: A Leishmaniasis Vaccine Candidate

Xinyu Liu[†], Sibylle Siegrist[‡], Mario Amacker[§], Rinaldo Zurbriggen[§], Gerd Pluschke[‡], and Peter H. Seeberger^{†,*}

[†]Laboratorium für Organische Chemie, ETH Zürich, Wolfgang-Pauli Strasse 10, 8093 Zürich, Switzerland, [‡]Swiss Tropical Institute, Socinstrasse 57, 4002 Basel, Switzerland, and [§]Pevion Biotech Ltd., Rehhagstrasse 79, 3018 Bern, Switzerland

Vaccinations based on defined antigens aim to elicit a specific immune response against a molecule expressed by the pathogen to be eliminated by the immune system. In addition to proteins, carbohydrate antigens are becoming increasingly important as target structures, and vaccines based on capsular polysaccharides extracted from *Haemophilus influenzae* type B, *Neisseria meningitidis*, and *Streptococcus pneumoniae* are already introduced into routine immunization schedules (1–4). Since polysaccharides are often very heterogeneous and difficult to isolate from their natural sources, it is expected that recent improvements in carbohydrate synthesis technologies (5) will lead to the development of new synthetic carbohydrate vaccines (6, 7). Unconjugated polysaccharide antigens typically elicit only T-cell independent short-lived and low-affinity IgM antibody responses, and the immunogenicity of smaller synthetic oligosaccharides tends to be even weaker. Therefore synthetic carbohydrate vaccine design is critically dependent on the development of an antigen delivery platform that promotes the generation of T-cell dependent immune responses against oligosaccharides. Stimulation of B-cells should be associated with the development of long-term memory, Ig affinity maturation, and isotype class switching to IgG. The T-cell independent properties of carbohydrate antigens can be overcome by conjugation to a carrier protein. In addition, the immunogenicity has to be enhanced by delivery

of the conjugates with an immunological adjuvant. Alum has remained the dominant adjuvant for human vaccines, since many other candidate adjuvants have shown unsuitable properties including reactogenicity, toxicity, instability, or high costs. Since alum has a relatively poor adjuvant effect on polysaccharide antigens (8), the use of immunostimulating reconstituted influenza virosomes (IRIVs) may constitute an attractive alternative integrated carrier and adjuvant platform for synthetic oligosaccharides. Potential advantages of IRIVs for subunit vaccine antigen delivery include their fusogenic activity, serial display of the molecularly defined antigens on the surface of the virus-like particles, and an excellent safety record both in animals and humans (9). Two IRIV-based vaccines against hepatitis A and influenza are currently on the market (10), and it has been shown that IRIVs represent an excellent delivery system for small synthetic peptides (11). We have evaluated IRIVs for the first time as a synthetic carbohydrate antigen delivery system as part of an ongoing leishmania vaccine development program. The results suggest that IRIVs represent a general platform for oligosaccharide vaccines, and details are reported here.

Kala azar, now known as visceral leishmaniasis, threatens more than 350 million people worldwide and kills 60,000 annually. Transmitted by the bite of the female phlebotomine sandfly, protozoan parasites of the genus *Leishmania* cause the tropical disease that is still treated with old

ABSTRACT Novel virosomal formulations of a synthetic oligosaccharide were prepared and evaluated as vaccine candidates against leishmaniasis. A lipophosphoglycan-related synthetic tetrasaccharide antigen was conjugated to a phospholipid and to the influenza virus coat protein hemagglutinin. These glycan conjugates were embedded into the lipid membrane of reconstituted influenza virus virosomes. The virosomal formulations elicited both IgM and IgG anti-glycan antibodies in mice, indicating an antibody isotype class switch to IgG. The antisera cross-reacted *in vitro* with the corresponding natural carbohydrate antigens expressed by leishmania cells. These findings support the concept of using virosomes as universal antigen delivery platform for synthetic carbohydrate vaccines.

*To whom correspondence should be addressed.
Email: seeberger@org.chem.ethz.ch.

Received for review February 25, 2006
and accepted April 4, 2006

Published online April 21, 2006
10.1021/cb600086b CCC: \$33.50

© 2006 by American Chemical Society

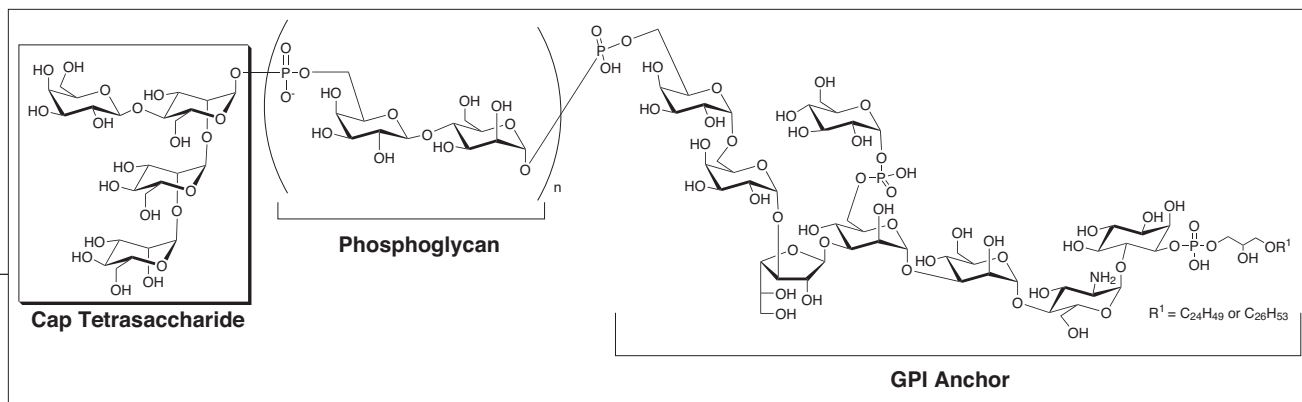


Figure 1. General structure of the leishmania lipophosphoglycan (LPG).

antimony-based drugs that are expensive and associated with significant side effects. Despite of many efforts, no effective vaccine for any form of leishmaniasis has emerged yet (World Health Organization, State of the Art of New Vaccines: Research and Development, http://www.who.int/vaccine_research/documents/new_vaccines/en/).

The lipophosphoglycans (LPGs) on the cell surface of leishmania parasites are an important virulence factor and essential for the survival and infectivity of the parasites. These cell surface glycoconjugates are therefore specific markers such as those on many cancer cells, bacteria, viruses, and parasites (12, 13). Disease-specific carbohydrates have attracted the attention of immunologists as antigens for the creation of novel vaccines in recent years as the methods for the synthesis of complex oligosaccharides have significantly improved (14, 15). Vaccines based on synthetic oligosaccharide antigens against bacterial infections, HIV, and malaria, as well as cancers including breast and prostate cancers, are at different stages of preclinical and clinical development (16–22). A unique, structurally well-defined capping tetrasaccharide that has been implicated as crucial for the invasion of the parasite into macrophages terminates the LPGs of leishmania parasites (Figure 1) (23). This tetrasaccharide was the focal point of our efforts to create a leishmaniasis vaccine candidate.

The tetrasaccharide epitope, equipped with a short PEG-linker terminated by a thiol group at the reducing end, was synthesized in a linear fashion (24–26). The thiol serves as a handle for conjugation to different carriers (27). The chemical synthesis of the tetrasaccharide epitope started with the installation of thiol-based linker **1** on the carbohydrate moiety (Figure 2). The coupling of **1** with readily available mannose building block **4** gave mannoside **5** in 92% yield (see

Supporting Information). Selective opening of 4,6-*O*-benzylidene afforded mannoside **6** in 63% yield. Union of galactosyl phosphate **2** (28) with **6** and subsequent cleavage of the acetyl ester in the presence of a pivalate using dilute HCl generated *in situ* under nonaqueous conditions gave disaccharide **7** in 76% yield. The disaccharide was further decorated with mannose by glycosylation using mannosyl trichloroacetimidate **3** to furnish trisaccharide **8** in 69% yield. Selective removal of acetate afforded **9** in 90% yield. Placement of the terminal mannose proved far more challenging as an excess (3 equiv) of **3** was needed to afford the protected tetrasaccharide **10** in good yield. The conditions of dissolving metal reductions ensured the complete removal of all protecting groups before purification by Sephadex column chromatography and dialysis furnished the pure tetrasaccharide **11** in 81% yield. Predominantly, the dimer of **11** was obtained as indicated by NMR and HR-MS analysis (see Supporting Information).

For immunological studies, the disulfide in glycan **11** was first reduced *in situ* with tricarboxyethylphosphine (TCEP) and then conjugated either to a phospholipid or to a carrier protein (Scheme 1). Maleimide-activated phospholipid **12** was prepared by combining

amine (PE) with 4-maleimidobutyric acid sulfo-*N*-succinimidyl ester (sulfo-GMBS). The *in situ* reduced glycan **11** was mixed with **12** to generate the lipid conjugate **13**. The conjugate was formulated into IRIVs as described (29), by resuspending it in octaethyleneglycol (OEG) and combining it with egg phosphatidylcholine (PC), surface glycoproteins, and phospholipids of inactivated influenza A/Singapore/6/86 (H1N1). PE-Glycan loaded IRIVs were formed by detergent removal. For the preparation of protein conjugates, keyhole limpet hemocyanin (KLH) or solubilized influenza hemagglutinin (HA) was activated with sulfo-GMBS to result in maleimide-activated carrier proteins that were mixed with the reduced glycan **11**. While the resulting glycan-HA conjugate **14** was resuspended in OEG to prepare HA-glycan loaded IRIVs, the glycan-KLH conjugate **15** was used to coat ELISA plates with the glycan **11**.

Glycan-PE **13** and glycan-HA **14** loaded IRIVs were used to immunize groups of

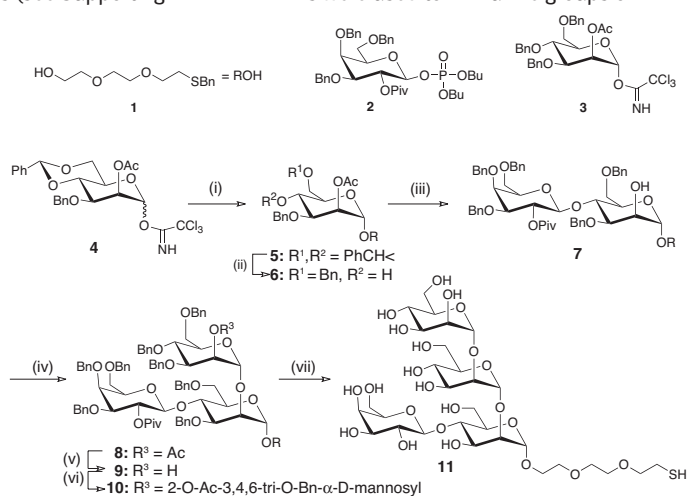
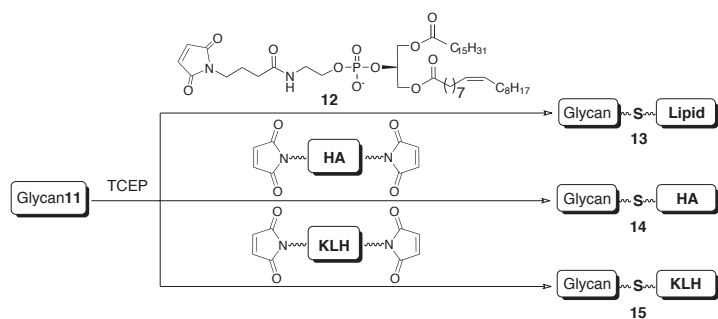


Figure 2. Synthesis of the thiol-tethered LPG tetrasaccharide epitope.

(i) **1**, TBSOTf (cat.), Et₂O/CH₂Cl₂, 0 °C, 92%; (ii) Et₃SiH, TFOH, 4 Å MS, CH₂Cl₂, –78 °C, 63%; (iii) **2**, TMSOTf, CH₂Cl₂, –40 °C; then AcCl in THF/MeOH, 0 °C to rt, 76% 2 steps; (iv) **3**, TMSOTf (cat.), Et₂O/CH₂Cl₂, 0 °C, 69%; (v) AcCl in THF/MeOH, 0 °C to RT, 90%; (vi) **3**, TMSOTf (cat.), Et₂O/CH₂Cl₂, 0 °C, 53%; (vii) Na, NH₃, THF/MeOH, –78 °C, 81%.



Scheme 1. Chemical conjugation of LPG tetrasaccharide for immunological studies

BALB/c mice intramuscularly. After pre-immunization with the IRIV-based influenza vaccine Inflexal Berna (Berna Biotech) the mice received at 3-week intervals three doses of IRIV containing 50 μ g of glycan conjugated either to PE or to HA. Priming of mice with influenza antigen enhances the antibody response against IRIV-associated antigens (30). When the generation of glycan-specific antibodies was analyzed by ELISA using KLH-glycan coated microtiter plates, both IgG and IgM response was observed. IgG isotyping indicated that IgG1 contributed dominantly to the antibody elicited, with only negligible amounts of IgG3 detected (ELISA data not shown). These data provided solid evidence that the observed IgG antibody response is T-cell dependent. Cross-reactivity with natural carbohydrate antigens produced by leishmania cells was analyzed by immunofluorescence staining of axenic amastigote *Leishmania donovani* (MHOM-ET-67/L82) parasites. The *in vitro* cultivated parasites were washed, suspended in PBS containing 1% bovine serum albumin, and fixed on diagnostic microscope slides for at least 60 min at room temperature using a 4% paraformaldehyde fixing solution containing 0.1% Triton X-100. After being washed, slides were incubated with serial dilutions of mouse sera for 2 h at room temperature in a humid chamber. After being washed, slides were incubated for 60 min with Cy3-conjugated AffiniPure F(ab')₂ fragment goat-anti-mouse IgG, Fc γ fragment specific antibodies for immunoflu-

orescence staining and 1 μ g mL⁻¹ Hoechst dye 33258 for DNA staining. After being washed, the slides were dried, mounted with mounting solution, covered with a cover slide, and assessed by fluorescence microscopy. Although LPG expression is downregulated in axenic amastocytotes (31), staining by parasite cross-reactive IgG was observed with sera of all mice immunized with glycan-loaded IRIV. Whereas glycan-HA loaded IRIVs induced an ELISA titer higher than that of glycan-PE loaded IRIVs, both formulations elicited comparable titers of parasite cross-reactive IgG in the

immunofluorescence analysis (Figure 3, panel a). No staining was observed with pre-immune sera and after three immunizations of mice with IRIV that were not loaded with glycan 8 (Figure 3, panel b). The anti-glycan antisera stained also parasites in liver sections from leishmania-infected hamsters (data not shown). These assays demonstrate that the new vaccine candidate is highly immunogenic and elicits IgG antibodies that recognize leishmania parasites. Challenge studies in animals are the next step en route to a leishmaniasis vaccine.

In conclusion, IRIVs represent a viable antigen delivery system suitable to induce T-cell dependent antibody responses against oligosaccharide antigens. Our results with two IRIV-based formulations of a leishmanial tetrasaccharide indicate that IRIVs have a great potential for the design of safe and effective synthetic carbohydrate vaccines. Thus, IRIVs present an attractive alternative to currently used carbohydrate vaccine formulations and should find application for diseases including bacterial, viral, and parasitic infections and cancer.

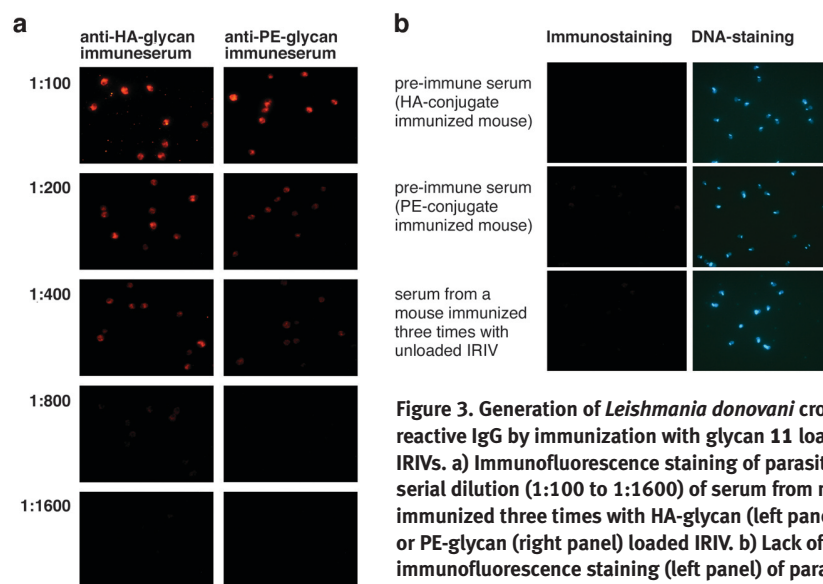


Figure 3. Generation of *Leishmania donovani* cross-reactive IgG by immunization with glycan 11 loaded IRIVs. a) Immunofluorescence staining of parasites by serial dilution (1:100 to 1:1600) of serum from mice immunized three times with HA-glycan (left panel) or PE-glycan (right panel) loaded IRIV. b) Lack of immunofluorescence staining (left panel) of parasites by pre-immune sera (dilution 1:100) and by serum (dilution 1:100) of a mouse immunized three times with unloaded IRIV. Presence of parasites on the microscopic slides was demonstrated by staining of parasite DNA with Hoechst dye 33258 (right panel). Typical results are shown.

METHODS

Preparation of Lipid–Glycan Conjugate 13.

PE (4 mg) was dissolved in a mixture of CHCl_3 /MeOH (500 μL , 9:1), and triethylamine (5 μL) was added. This solution was mixed with GMBS (3 mg) and incubated for 2 h at 30 °C using a thermomixer. Leishmania glycan 11 (2 mg) was dissolved in OEG-PBS (0.5 mL, 100 mM), and TECP solution (5 μL , 0.5 mM) added. The solution was incubated for 5 min at 30 °C and pipetted into the activated PE solution. The mixture was then further incubated for 3 h at 30 °C. The resulting solution was stored at 4 °C for immunization study.

Preparation of KLH–Glycan Conjugate 15.

Sulfo-GMBS (2 mg) was mixed with 1 mL of KLH solution and incubated for 3 h at 30 °C using a thermomixer. Leishmania glycan (2 mg) was dissolved in PBS (0.2 mL), and TECP solution (5 μL , 0.5 mM) was added. The solution was incubated for 5 min at 30 °C and pipetted into the activated KLH solution. The mixture was then further incubated for 3 h at 30 °C. The resulting solution was stored at 4 °C for immunization studies. HA–glycan conjugate was prepared in an analogous fashion.

Virosome Formulation. Glycan-IRIVs were prepared by the method described previously (29). Briefly, 32 mg of egg PC (phosphatidylcholine), 8 mg of PE, and 2 mg of the glycan-PE conjugate were dissolved in 3 mL of PBS, 100 mM OEG (OEG-PBS). Next, 4 mg HA of inactivated influenza A/Singapore virus was centrifuged at 100000g for 1 h at 4 °C, and the pellet was dissolved in 1 mL of OEG-PBS. The detergent-solubilized phospholipids and viruses were mixed and sonicated for 1 min. This mixture was centrifuged at 100000g for 1 h at 20 °C, and the supernatant was sterile filtered (0.22 μm). Virosomes were then formed by detergent removal using 1.25 g of wet SM2 Bio-Beads (BioRad, Glattbrugg, Switzerland) for 1 h at room temperature with shaking and three times for 30 min with 625 mg of SM2 Bio-Beads each.

ELISA. For enzyme-linked immunosorbent assay (ELISA) analyses, Polysorp plates were coated overnight at 4 °C with 100 μL of a 10 $\mu\text{g mL}^{-1}$ solution of KLH-glycan conjugate in PBS (pH 7.4). Wells were then blocked with 5% milk powder in PBS for 2 h at 37 °C, followed by three washes with PBS containing 0.05% Tween 20. Plates were then incubated with serial dilutions of the mouse sera in PBS containing 0.05% Tween 20 and 0.5% milk powder for 2 h at 37 °C. After washing, plates were incubated with either goat anti-mouse Ig-HRP (dilution 1:1000), goat anti-mouse IgG-HRP (dilution 1:2000), or goat anti-mouse IgM-HRP (dilution 1:2000), or for IgG isotyping, rabbit anti-mouse IgG1-HRP (dilution 1:2000) and goat anti-mouse IgG3-HRP (dilution 1:1000), respectively, for 1 h at 37 °C. After the plates were washed, O-phenylenediamine substrate was added, the plates were incubated in the dark at room temperature until the colorimetric reaction had progressed sufficiently, the reaction was stopped by addition of 100 μL of 1 M H_2SO_4 , and optical densities (OD) were read at 492 nm on a Spectra Max Plus. Endpoint titer is calculated as the reciprocal of the last serum dilution with the $\text{OD}_{\text{test sera}} \geq 2 \times \text{OD}_{\text{negative serum}}$. A strong signal is considered with endpoint titers greater than 5000.

Acknowledgment: Financial support from the Swiss National Science Foundation and ETH Zürich are gratefully acknowledged. We thank S. Rosenfellner for technical assistance and R. Brun for support with the cultivation of leishmania.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Jennings, H. J., and Pon, R. A. (1996) Polysaccharides and glycoconjugates as human vaccines, in *Polysaccharides in Medicinal Applications* (Dumitriu, S., Ed.) pp 443–479, Marcel Dekker, New York.
- Robbins, J. B., Schneerson, R., Szu, S. C., and Pozsgay, V. (1999) Bacterial polysaccharide-protein conjugate vaccines, *Pure Appl. Chem.* **71**, 745–754.
- Moreau, M., and Schulz, D. (2000) Polysaccharide based vaccines for the prevention of pneumococcal infections, *J. Carbohydr. Chem.* **19**, 419–434.
- Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens, *Carbohydr. Res.* **338**, 2539–2547.
- Ernst, B., Hart, G. W., and Sinay, P. (2000) *Carbohydrates in Chemistry and Biology, Part I: Chemistry of Saccharides, Volume 1: Chemical Synthesis of Glycosides and Glycomimetics*, Wiley-VCH, Weinheim.
- Roy, R. (2004) New trends in carbohydrate-based vaccines, *Drug Discovery Today: Technol.* **1**, 327–336.
- Borman, S. (2004) Carbohydrate vaccines, *Chem. Eng. News* **82**, 31–35.
- Lindblad, E. B. (1995) Aluminium adjuvants, in *The Theory and Practical Application of Adjuvants* (Stewart-Tull, D. E. S., Ed.) pp 21–35, John Wiley & Sons, Chichester, U.K.
- Zurbriggen, R. (2003) Immunostimulating reconstituted influenza virosomes, *Vaccine* **21**, 921–924.
- Westerfeld, N., and Zurbriggen, R. (2005) Peptides delivered by immunostimulating reconstituted influenza virosomes, *J. Pept. Sci.* **11**, 707–712.
- Pfeiffer, B., Peduzzi, E., Moehle, K., Zurbriggen, R., Gluck, R., Pluschke, G., and Robinson, J. A. (2003) A virosome-mimotope approach to synthetic vaccine design and optimization: Synthesis, conformation, and immune recognition of a potential malaria-vaccine candidate, *Angew. Chem., Int. Ed.* **42**, 2368–2371.
- Varki, A. (1993) Biological roles of oligosaccharides—All of the theories are correct, *Glycobiology* **3**, 97–130.
- Dwek, R. A. (1996) Glycobiology: Toward understanding the function of sugars, *Chem. Rev.* **96**, 683–720.
- Koeller, K. M., and Wong, C. H. (2000) Synthesis of complex carbohydrates and glycoconjugates: Enzyme-based and programmable one-pot strategies, *Chem. Rev.* **100**, 4465–4493.
- Seeberger, P. H., and Werz, D. B. (2005) Automated synthesis of oligosaccharides as a basis for drug discovery, *Nat. Rev. Drug Discovery* **4**, 751–763.
- Danishefsky, S. J., and Allen, J. R. (2000) From the laboratory to the clinic: A retrospective on fully synthetic carbohydrate-based anticancer vaccines, *Angew. Chem., Int. Ed.* **39**, 836–863.
- Dudkin, V. Y., Orlova, M., Geng, X. D., Mandal, M., Olson, W. C., and Danishefsky, S. J. (2004) Toward fully synthetic carbohydrate-based HIV antigen design: On the critical role of bivalency, *J. Am. Chem. Soc.* **126**, 9560–9562.
- Adams, E. W., Ratner, D. M., Bokesch, H. R., McMahon, J. B., O’Keefe, B. R., and Seeberger, P. H. (2004) Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology: Glycan-dependent gp120/protein interactions, *Chem. Biol.* **11**, 875–881.
- Lee, H. K., Scanlan, C. N., Huang, C. Y., Chang, A. Y., Calarese, D. A., Dwek, R. A., Rudd, P. M., Burton, D. R., Wilson, I. A., and Wong, C.-H. (2004) Reactivity-based one-pot synthesis of oligomannoses: Defining antigens recognized by 2G12, a broadly neutralizing anti-HIV-1 antibody, *Angew. Chem., Int. Ed.* **43**, 1000–1003.
- Wang, L. X., Ni, J. H., Singh, S., and Li, H. G. (2004) Binding of high-mannose-type oligosaccharides and synthetic oligomannose clusters to human antibody 2G12: Implications for HIV-1 vaccine design, *Chem. Biol.* **11**, 127–134.
- Verez-Bencomo, V., Fernandez-Santana, V., Hardy, E., Toledo, M. E., Rodriguez, M. C., Heynngnezz, L., Rodriguez, A., Baly, A., Herrera, L., Izquierdo, M., Villar, A., Valdes, Y., Cosme, K., Deler, M. L., Montane, M., Garcia, E., Ramose, A., Aguilar, A., Medina, E., Torano, G., Sosa, I., Hernandez, I., Martinez, R., Muzachio, A., Carmenates, A., Costa, L., Cardoso, F., Campa, C., Diaz, M., and Roy, R. (2004) A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b, *Science* **305**, 522–525.
- Schofield, L., Hewitt, M. C., Evans, K., Siomos, M. A., and Seeberger, P. H. (2002) Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria, *Nature* **418**, 785–789.
- Descoteaux, A., and Turco, S. J. (2002) Functional aspects of the Leishmania donovani lipophosphoglycan during macrophage infection, *Microbes Infect.* **4**, 975–981.
- Arasappan, A., and Fraser-Reid, B. (1996) *N*-Pentenyl glycoside methodology in the stereoselective construction of the tetrasaccharyl cap portion of *Leishmania lipophosphoglycan*, *J. Org. Chem.* **61**, 2401–2406.
- Upreti, M., Ruhela, D., and Vishwakarma, R. A. (2000) Synthesis of the tetrasaccharide cap domain of the antigenic lipophosphoglycan of *Leishmania donovani* parasite, *Tetrahedron* **56**, 6577–6584.
- Hewitt, M. C., and Seeberger, P. H. (2001) Solution and solid-support synthesis of a potential leishmaniasis carbohydrate vaccine, *J. Org. Chem.* **66**, 4233–4243.
- Ratner, D. M., Adams, E. W., Su, J., O’Keefe, B. R., Mrksich, M., and Seeberger, P. H. (2004) Probing protein-carbohydrate interactions with microarrays of synthetic oligosaccharides, *ChemBioChem* **5**, 379–382.
- Plante, O. J., Andrade, R. B., and Seeberger, P. H. (1999) Synthesis and use of glycosyl phosphates as glycosyl donors, *Org. Lett.* **1**, 211–214.
- Zurbriggen, R., Novak-Hofer, I., Seeliger, A., and Gluck, R. (2000) IRIV-adjuvanted hepatitis A vaccine: in vivo absorption and biophysical characterization, *Prog. Lipid Res.* **39**, 3–18.
- Poltl-Frank, F., Zurbriggen, R., Helg, A., Stuart, F., Robinson, J., Gluck, R., and Pluschke, G. (1999) Use of reconstituted influenza virus virosomes as an immunopotentiating delivery system for a peptide-based vaccine, *Clin. Exp. Immunol.* **117**, 496–503.
- Gupta, N., Goyal, N., and Rastogi, A. K. (2001) In vitro cultivation and characterization of axenic amastigotes of *Leishmania*, *Trends Parasitol.* **17**, 150–153.

The Innate Immunity of Maize and the Dynamic Chemical Strategies Regulating Two-Component Signal Transduction in *Agrobacterium tumefaciens*

Justin Maresh[†], Jin Zhang[†], and David G Lynn^{*}

Departments of Chemistry and Biology, Emerson Hall, Emory University, Atlanta, Georgia 30322

Members of the benzoxazinone family of heterocycles are major secondary metabolites found in grasses of the family Poaceae (also known as Graminae), including the cereals maize, wheat, and rye. Generally, they have been recognized as “resistance factors” that function as part of a nonspecific arsenal of metabolic defense agents, since they have been implicated in antimicrobial, antifungal, insecticidal, anti-feeding, and mutagenic activities, as well as involvement in a tritrophic interaction (1, 2). The costs associated with the maintenance and adaptive evolution of entire metabolic pathways, particularly those associated with general resistance factors, are clearly some function of the sum of their benefits. Yet, the biology of such compounds is less well-defined than that of more specific gene-for-gene relationships, ones in which complements of resistance and avirulence genes determine compatible/incompatible relationships in pathogenesis (3).

One complicating factor in understanding benzoxazinone activity is the spectrum of possible chemical mechanisms and the intrinsic hydrolytic instability of these species. For example, the primary benzoxazinone found in maize and wheat, DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one, **2**, Scheme 1), possesses many chemical reactivities with proposed biological relevance. On one hand, the ring-chain tautomeric aldehyde of DIMBOA is considered to be an electrophile implicated in reversible reaction with protein nucleophiles, particularly in the hydrophobic binding cavities accessible to the aromatic ring. In a simple model for this mode of action, DIMBOA was shown to form imine adducts with the ϵ -NH₂ group of *N*-acetyl lysine (4). DIMBOA was also shown to inhibit chymotrypsin activity by addition of the aldehyde to

ABSTRACT The naturally occurring 2-hydroxy-4,7-dimethoxybenzoxazin-3-one (HDMBOA), essentially the sole component of maize seedling organic exudate, was shown to be a potent inhibitor of the VirA–VirG two-component system which mediates host recognition and activates virulence gene transcription in the soil pathogen *Agrobacterium tumefaciens*. The hydrolytic lability of HDMBOA creates a steady-state zone of inhibition circumscribing the young maize seedling. We now show that rather than the HDMBOA natural product, an *o*-imidoquinone decomposition intermediate, (3Z)-2,2-dihydroxy-*N*-(4-methoxy-6-oxocyclohexa-2,4-dienylidene)acetamide, can function as an inhibitor of virulence gene expression in *A. tumefaciens*. Structural characterization of this *o*-imidoquinone intermediate clarifies several issues related to the decomposition pathways available to this class of antibiotics. Of direct ecological importance, this species is produced rapidly and quantitatively within the more neutral pH ranges of the *A. tumefaciens* cytoplasm, while HDMBOA is more persistent at the slightly acidic pH common to many soils. These results suggest the rather intriguing possibility that the physical instability of the benzoxazinone antibiotics may not only create a steady-state local defense, but also enable a “pro-drug” strategy directed against bacterial environmental sensing.

*To whom correspondence should be addressed.
E-mail: dlynn2@emory.edu.

[†]These authors contributed equally to this work.

Received for review February 4, 2006
and accepted April 2, 2006.

Published online April 21, 2006.

10.1021/cb600051w CCC: \$33.50

© 2006 by American Chemical Society

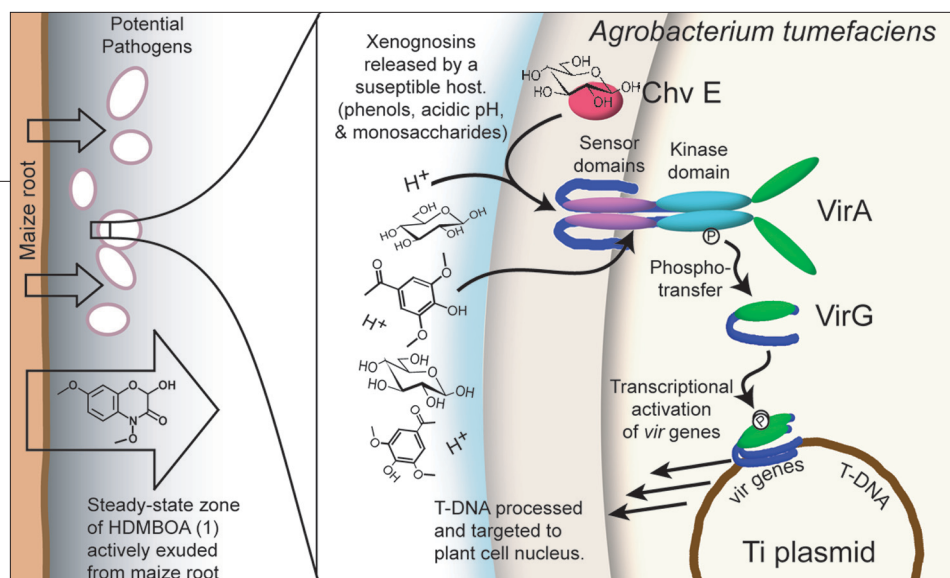


Figure 1. Overview of VirA–VirG-mediated host recognition by *Agrobacterium tumefaciens* and the chemical defensive strategy of maize seedlings. HDMBOA (1) is the primary organic component exuded by the roots of maize seedlings; its instability in this environment creates a steady-state defensive zone around the seedling. Three signals characteristic of a plant wound site, phenols, acidic pH, and monosaccharides, are perceived by the inner membrane sensor kinase VirA. Phosphotransfer to the response regulator VirG activates transcription of virulence genes. It is this signal transduction pathway that is specifically inhibited by HDMBOA.

the active site serine (5), and a quantum mechanical model suggested that DIMBOA inhibits a glutathione *S*-transferase through reaction of the aldehyde with the thiol of reduced glutathione (6). Additionally, the cyclic hydroxamic acid of DIMBOA may also function as a thiol oxidant, generating a lactam (2-hydroxy-7-methoxy-benzoxazin-3-one, HMBOA, 3) as the reduction product (7, 8). Further, metal-binding properties of the hydroxamic acid moiety have been implicated in metal-mediated reactivity and possibly ion transport (9). Finally, potent electrophilicity following metabolic activation has also been suggested. *O*-acetylated hydroxamic acids, such as ADIBOA 4 (R = OAc in the basic structure in Scheme 1), regioselectively react as strong electrophiles with C, N, and S nucleophiles (10), including DNA (11). Accordingly, DIMBOA proved to be mutagenic only after rat liver S9 activation (12). Given these various reports, it is not surprising that the benzoxazinones are regarded as rather exceptional resistance factors.

The plant cell maintains benzoxazinones in the vacuole as low toxicity, stable β -glycosides. The reactive aglycones are generated by release of glucosidases localized in the plastid upon tissue damage related to herbivore feeding or pathogen infection (13). Accordingly, DIMBOA is easily obtained from maize seedlings by homogenization of the tissue to release these enzymes before extraction (14). Although active root exudation of aglycones has been reported (15, 16), little is known about this process.

stresses by converting their considerable reserves of DIMBOA-Glc to HDMBOA-Glc, understanding the potential defensive role of the HDMBOA aglycone becomes an important question.

It was therefore surprising to discover that HDMBOA comprises ~98% of the organic extract from the surface of intact maize roots (15). Most biological studies of benzoxazinone activity focus on pathogen or herbivore viability. Therefore, it was further surprising to discover from co-cultivation of maize seedlings with the plant pathogen *Agrobacterium tumefaciens*, that HDMBOA exuded directly from the roots did not inhibit bacterial colonization. Instead, a reporter gene indicated a steady-state zone defined by the benzoxazinone was a specific natural inhibitor of the sensing system that regulates virulence gene expression (15) (Figure 1).

A. tumefaciens is a Gram-negative soil bacterium that uses a mechanism of genetic transformation of host cells to incorporate foreign oncogenetic elements into higher plants to generate tumors characteristic of Crown Gall disease (21) (Figure 1). The critical elements necessary for infection are contained within the large 210 kb tumor-inducing plasmid (Ti-plasmid). Virulence is activated upon perception of three host-derived signals, phenols, acidic pH, and sugars, all characteristic of a susceptible lesion in plant tissue. Perception and response to these host recognition factors, or xenogonins, is modulated by a single Ti-plasmid-localized, two-component signaling system, VirA–VirG (22), which relies on an

Relative to DIMBOA, fewer studies have focused on the benzoxazinone HDMBOA (1), as this compound is highly labile in aqueous media (17). Recent reports however suggest that the defensive role of HDMBOA is more significant than previously assumed. Stress-inducing defense elicitors, such as jasmonic acid and CuCl_2 , induce significant conversion of the glucoside DIMBOA-Glc to HDMBOA-Glc in both maize (18) and wheat (19). Similar induction has been observed upon colonization of maize leaves by various fungi and by herbivore feeding (20). Given that these plants respond to certain

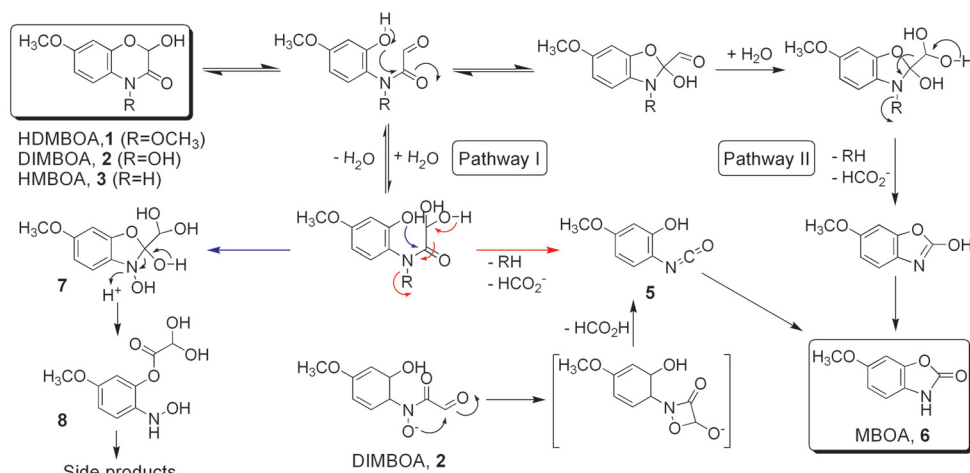
accessory chromosomal protein ChvE to mediate sugar and pH perception. Two-component systems are a common signal-transduction strategy employed by prokaryotes for response to environmental stimuli (23, 24), particularly for the regulation of virulence in pathogenic bacteria (25). VirA serves as the histidine autokinase transducer that transfers a phosphoryl group from a conserved histidine to a conserved aspartate on the response regulator VirG. Active phospho-VirG mediates the transcriptional activation of the *vir* regulon.

Therefore, this defense strategy of HDMBOA is unusual, with host perception rather than pathogen viability being targeted, changing the selection pressure for the development of resistance observed with most antibiotics. In a practical sense, effective anti-virulence agents could be useful in combination therapy to extend current antibiotics. Accordingly, bacterial two-component signal transduction systems, like the VirA–VirG system, could serve as a model for the development of new antibiotics (25, 26).

The inherent chemical reactivity of the benzoxazinones, particularly their instability in aqueous media, is critical to any understanding of the biological activity of this class of secondary products. Previous reports of an intermediate in HDMBOA decomposition (17) and reports of an orange-colored intermediate observed during decomposition of HDMBOA in methanol (27) directed our efforts to reinvestigate the decomposition. We now report kinetic analyses of HDMBOA decomposition, the structure of a decomposition intermediate, and argue that this species, rather than HDMBOA, may function as the specific inhibitor of host perception by *A. tumefaciens*.

RESULTS AND DISCUSSION

Decomposition of HDMBOA. The mechanism of benzoxazinone decomposition has been debated now for decades. Proposed mechanisms of HDMBOA decomposition assume that its reactivity is analogous to that of the well-characterized benzoxazinone,



Scheme 1. Summary of decomposition pathways previously proposed for benzoxazinones

DIMBOA, since MBOA (6-methoxybenzoxazol-2-one, 6) is the primary product of both reactions (7, 28–30), and two general models for HDMBOA decomposition are summarized in Scheme 1. The aqueous half-life of several natural benzoxazinones follow the order HMBOA 3 » DIMBOA 2 » HDMBOA 1. Both models interpret this qualitative trend in terms of the leaving group activity of the nitrogen substituent, OMe[−] > OH[−] » H[−] (17, 31). Early labeling studies indicated that the aldehyde C-2 carbon generates formic acid in benzoxazinone decomposition (32). Likewise, both models invoke a 1,5-fragmentation of HDMBOA, with methoxide and formic acid liberated simultaneously. However, the two models differ in the role assigned to the phenol hydroxyl.

Grambow and Lückge proposed Pathway I, a phenol-independent fragmentation to isocyanate 5 (17). For DIMBOA, the decomposition rate is critically dependent on the concentration of the hydroxamic acid monoanion (pK₁ of 6.9). This anion was proposed to act as an internal nucleophile in the rate-determining step of the reaction by attacking the aldehyde carbon, presumably via the four-membered ring intermediate (Scheme 1) (28, 30). This reaction sequence liberates formic acid, generating a proposed isocyanate species, which rearranges to form MBOA. However, the *N*-methoxy oxygen of HDMBOA is a poorer nucleophile than the oxoanion of DIMBOA, yet for HDMBOA, the decomposition reaction is faster. This observation argued for a mechanism distinct from DIMBOA in which a heterolytic Grob-like fragmentation generates isocyanate 5 (17) (Scheme 1, Pathway I).

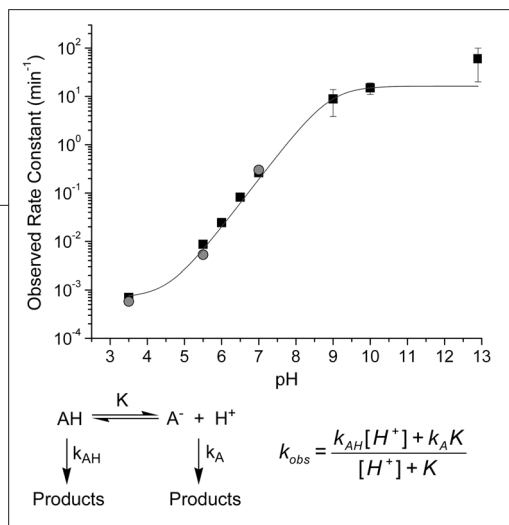


Figure 2. pH dependence for pseudo-first-order disappearance HDMBOA at 24 °C monitored by UV (squares) and NMR (circles) spectroscopy. Below pH 8.0, error bars represent one standard deviation of triplicate measurements. Above pH 8.0, the contribution of MBOA to the UV spectrum was significant after the first few seconds, therefore, error bars reflect the estimated uncertainty in the initial concentration. The solid line represent a nonlinear least-squares fit of UV and NMR data to the above equation corresponding to a monoprotic acid model: $k_{AH} = 7 \times 10^{-4} \text{ min}^{-1}$, $k_A = 16.25 \pm 0.04 \text{ min}^{-1}$, $K = (1.19 \pm 0.01) \times 10^{-9}$, $R^2 = 0.99998$.

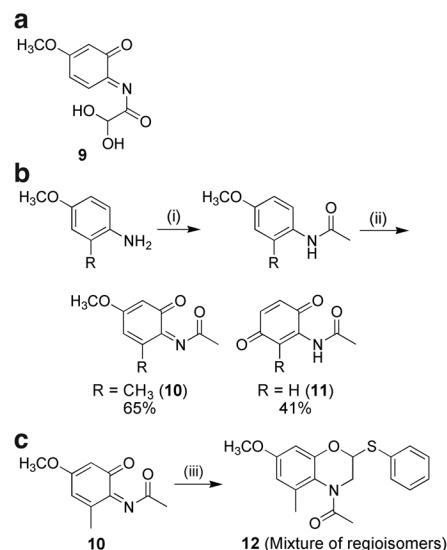
However, this proposal did not address the significant role of the phenol demonstrated by Smissman *et al.* in their study of open-chain analogue structures (31). Compounds lacking the phenol or containing an unreactive phenol ether were more stable than their benzoxazinone counterparts and did not hydrolyze to the expected isocyanate. Thus, Pathway II (Scheme 1) assigns the phenol hydroxyl the role of an active nucleophile, attacking the amide to give an intermediate benzooxazol-2-ol prior to fragmentation to MBOA. Given the report of an HDMBOA intermediate with an observable lifetime (17) and very different intermediates predicted by the two mechanisms of Scheme 1, we revisited the problem of HDMBOA decomposition.

The decomposition of HDMBOA (**1**) was followed from pH 3.5 to 13 in water and in methanol by ^1H NMR and UV spectroscopy. Under all conditions, the major decomposition product was MBOA (**6**), but the yield and the relative rate varied markedly (Figure 2). At pH 7.0, no HDMBOA was present by NMR after 15 min, being quantitatively converted to MBOA after 2.5 h (Figure 3).

For DIMBOA, the pH dependence of the pseudo-first-order rate of disappearance is biphasic, with a maximum around pH 9.0, suggesting that DIMBOA behaves like a diprotic acid in water (28). In contrast, the HDMBOA decomposition rate increases with pH over the range from pH 3.5 to 13. At basic pH, decomposition became so fast as to complicate measurements (reflected in the larger error bars of Figure 2). Despite this error, it was clear that no analogous decrease was found at basic pH, demonstrating that HDMBOA behaves either like a monoprotic acid or that removal of a second proton only increases the rate of decomposition. Assuming a simple monoprotic acid model (Figure 2), our data agrees with a mechanism in which decomposition is critically dependent on the

concentration of a species with a $\text{p}K_a$ of ~ 8.9 , a $\text{p}K_a$ most consistent with the ring-chain tautomeric phenol.

Characterization of the Major Intermediate. As previously reported (27), an orange intermediate was observed by thin layer chromatography (TLC) during HDMBOA decomposition. This color faded quickly, and direct isolation attempts were unsuccessful. However, an intermediate was also observed at early times by ^1H NMR. The rapid disappearance of HDMBOA and long relative lifetime of the intermediate at neutral pH allowed direct assignment of the ^1H NMR chemical shifts and coupling constants at 5.54 (d, $J = 2.0$, 1H), 6.33 (dd, $J = 10.2$, 2.0, 1H), and 6.40 (d, $J = 10.2$, 1H) ppm. Details may be seen



Scheme 2. a) The major intermediate of HDMBOA decomposition. b) Synthesis of **10**, an *o*-imidoquinone analogue of intermediate **9**. Reagents: (i) Ac_2O (3 equiv); (ii) DMP (4 equiv), H_2O (2 equiv), CH_2Cl_2 . c) Diels-Alder conformation of structure, (iii) vinyl phenyl sulfide (4.5 equiv), THF.

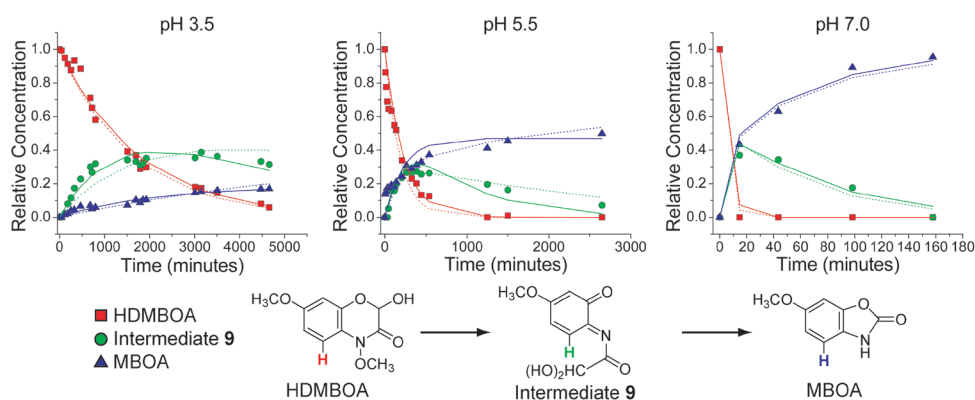


Figure 3. HDMBOA decomposition to MBOA as monitored by ^1H NMR. HDMBOA (7.5 mM) was dissolved in 100 mM phosphate buffer. The monitored peaks were HDMBOA (7.16 ppm, red squares), intermediate **9** (6.40 ppm, green circles), and MBOA (6.97 ppm, blue triangles). The lines represent theoretical data for the two models $k_4 = 0$ (solid lines) and $k_5 = 0$ (dotted lines).

in Supplementary Figure 2. These resonances were consistent with structure **9** (Scheme 2, panel a), a substituted *ortho*-imidoquinone spin system (**33**) with the proton singlet at 5.75 ppm being assigned as a hydrated aldehyde. These data were distinct from other possible aromatic intermediates including 2,4-dimethoxyphenyl-isocyanate, a stable model compound of intermediate **5** (complete assignments are found in Supporting Information).

To verify the identity of the intermediate, the analogous *o*-imidoquinone **10** was synthesized by Dess-Martin periodinane (DMP) oxidation of *N*-(4-methoxy-2-methyl-phenyl)acetamide (**34**, **35**) as shown in Scheme 2. The ^1H NMR and ^{13}C NMR spectra of the intensely yellow *o*-imidoquinone closely matched the HDMBOA decomposition intermediate in D_2O . Details may be seen in Supplementary Figure 2 and Supplementary Experimental Methods. The identity of the *o*-iminoquinone functionality of **10** was further confirmed by cycloaddition with vinyl phenyl thioether in THF at room temperature to obtain Diels-Alder adduct **12** (Scheme 2, panel c) (**34**, **36**).

The Role of Intermediate 9: A New Mechanism for HDMBOA Decomposition. Our identification of the dominant intermediate as *o*-imidoquinone **9** assigns a new role to the phenol, one not considered in the proposal of Pathway II (Scheme 1), but one that supports the critical role of the aldehyde hydrate as common to both Pathways I and II. This new mecha-

nism (Figure 4) differs considerably from decomposition mechanisms proposed for DIMBOA, as there is no requirement for nucleophilicity of either the hydroxamic acid or the phenol oxygen. Instead, the reactivity of the phenol and amide resembles the oxidation of an *ortho*-hydroquinone. Like the two models presented in Scheme 1, formation of the open-chain aldehyde hydrate (Figure 4, species **13**) is critical to this process. The transient kinetics observed by ^1H NMR in buffered D_2O solutions at pD 3.5, 5.5, and 7.0 and in methanol- d_6 (Figure 3) permitted elimination of several kinetic models for HDMBOA decomposition. For example, a simple linear mechanism of HDMBOA to **9** to MBOA did not fit our observations. Instead, a mechanism including two pathways to MBOA (Figure 4, k_1 and k_2) was required. Additionally, fitting our kinetic data to the model in Figure 4 allowed for an estimation of the half-life of the intermediate species **9**.

The semiempirical model in Figure 4 was developed on the basis of the following constraints. First, fast hydration was assumed for the HDMBOA open-chain aldehyde. Although previous models have suggested that hydration is slow, we propose that the hydrate of the aldehyde may generally dominate the open-chain structure of benzoxazinones in water as intermediate **9**, and all other open-chain HDMBOA species, were observed exclusively as aldehyde hydrates. No down-field aldehyde proton was observed during HDMBOA decomposition, and the chemical shift of 6.5 ppm

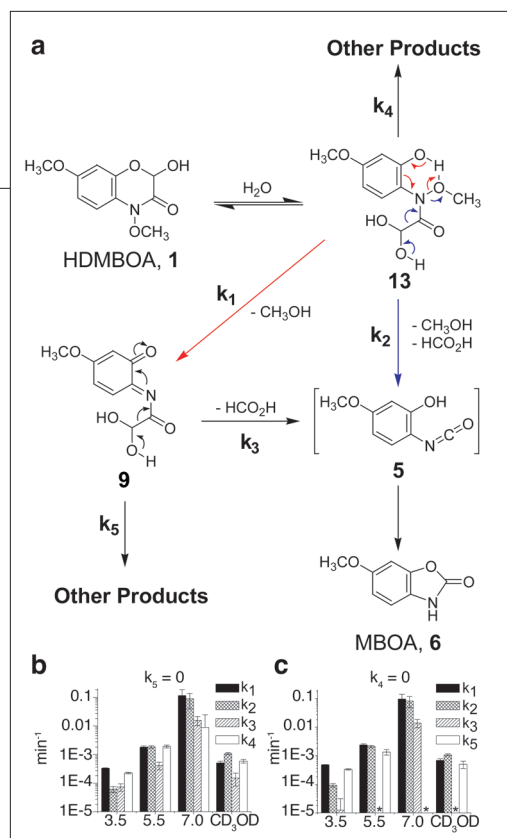


Figure 4. Proposed mechanism of HDMBOA decomposition. **a)** Inclusion of pathways k_4 and k_5 were required to obtain a reasonable fit for the experimental data. **b** and **c)** Panels represent the best fit kinetic constants to the model in panel **a** for two extremes: $k_5 = 0$ and $k_4 = 0$, respectively. An asterisk (*) indicates constants for which minimization only converged at the lower parameter bound (10^{-9}) and are treated as zero.

reversible addition of methanol to **9** was observed in pure CD_3OD , even when significant formation of **9** was observed. Further, no evidence for nucleophilic attack by water at the nitrogen center of **9** was obtained. The resulting product would be identical to the open-chain form of DIMBOA, and since the lifetime of DIMBOA is orders of magnitude longer than the total time of our experiments (37), DIMBOA should have accumulated.

Third, to account for the observation that the MBOA yield is reduced for $\text{pH} < 7$ and in methanol, it was necessary to introduce the additional decomposition pathways k_4 and k_5 . Both intermediates **9** and **13** may participate in side reactions, but in the absence of identifiable side products, it was not possible to experimentally probe the relative contributions of either pathway k_4 or k_5 . Therefore, the two models $k_4 = 0$ and $k_5 = 0$ were treated separately as limiting cases, and finally, only constants k_1 – k_5 were considered when fitting the model in Figure 4; all other steps, including hydrolytic ring opening, were assumed to be fast.

Several qualitative features are common to the limiting models $k_4 = 0$ and $k_5 = 0$. The rate constants for the

hydrate proton suggested that little or no aldehyde species was present in fast equilibrium on the NMR timescale. This effect has been previously observed in the ^1H NMR spectrum of HMBOA **3**, for which the aldehyde hydrate is apparently the exclusive open-chain species in D_2O buffer (pD 9.0) (7 and J. Zhang and J. Maresh, unpublished observations), just as these hydrates are generally observed for the simple oxo-aldehydes such as phenylglyoxal.

Second, the two pathways indicated by k_1 and k_2 (Figure 4) were assigned as irreversible. No evidence for

two pathways of HDMBOA disappearance (Figure 4, k_1 and k_2) are approximately equal at pH 5.5, pH 7.0, and in methanol. Both models also suggest a change in mechanism at acidic pH as the “direct” decomposition pathway k_2 becomes slow relative to k_1 . Despite these similarities, significant differences complicate interpretation of the rate-limiting step(s). The model $k_4 = 0$ is discontinuous, that is, rates k_3 and k_5 are effectively zero at pH 5.5 and 7.0, respectively. In contrast, for the model $k_5 = 0$, decomposition of **9** is consistently rate-limiting at every condition, yet the mechanistic reason for this dependence may in fact be complex. Principally, the equilibrium of HDMBOA ring-chain tautomerization may be pH -dependent. Although the critical pK_a of 8.9 derived from HDMBOA disappearance (Figure 2) supports this assignment, detailed equilibrium studies with stable analogues, such as HMBOA, may be necessary to establish a precise pH dependence of the ring-chain tautomer equilibrium.

Within either of the limits of $k_4 = 0$ or $k_5 = 0$, all rate constants generally increase with pH . For the $k_5 = 0$ model, in which intermediate **9** decomposes only to MBOA, the rates remain relatively congruent, suggesting that the pH -dependent ring-chain tautomerization of HDMBOA merely dictates the relative concentration of **13** (Figure 4, panel b). However, in the $k_4 = 0$ model, the pH dependence for the decomposition of intermediate **9** to either MBOA (k_3) or other products (k_5) does

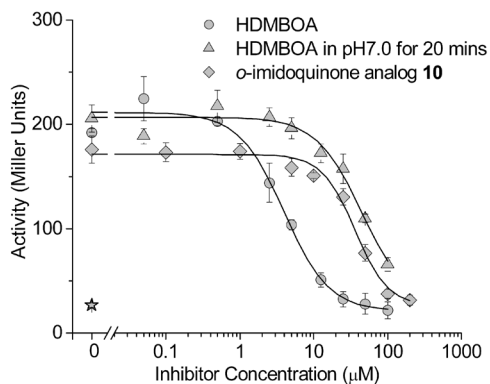


Figure 5. Activity of HDMBOA, HDMBOA decomposition intermediate(s), and analogue **10** for inhibition of *vir* gene expression in wild-type *A. tumefaciens* strain A348 pSW209 ($P_{virB}::lacZ$) in the presence of phenol inducer ($100 \mu\text{M}$ AS). Decomposition intermediates were generated by reacting HDMBOA in pH 7.0 buffer for 20 min before dilution in pH 5.5 induction medium. Hill slope was 1.4 ± 0.3 for all curve fits.

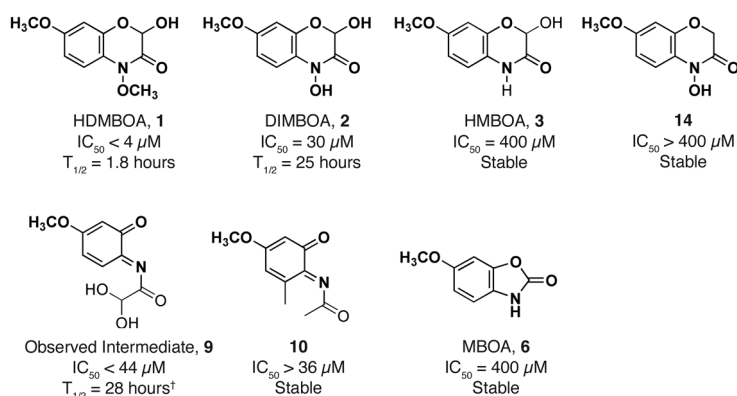


Figure 6. *Vir* inhibition activity and stability of HDMBOA and related derivatives. Activity was monitored by *virB::lacZ* expression with 100 μM AS at 28 °C. Chemical stability *in vitro* was monitored by 1H NMR at pH 5.5 and 24 °C. [†] $k_5 = 0$ assumed as in Figure 2 ($T_{1/2} = 23$ h if $k_4 = 0$ assumed).

not appear physically reasonable; k_3 is rate-limiting at pH 3.5 and 7.0, and too slow to be relevant at pH 5.5 or in MeOH (Figure 4, panel c). At these acidic conditions, the $k_4 = 0$ model assigns **9** as an off-pathway side product. However, the *o*-imidoquinone **10** was relatively stable in these conditions. No decomposition of **10** was observed in either pH 5.5 or pH 7.0 buffers after 24 h by 1H NMR and UV spectroscopy. Analogue **10** only decomposed following addition of 1 M DCl, NaOD, or isopropylamine (data not shown). Although, the $k_5 = 0$ model is physically most reasonable, both k_4 and k_5 may contribute to decomposition, and other pathways may also be accessible.

The yield of MBOA decreases at lower pH; however, specific side products were not successfully identified. Further identification of intermediates and side products should shed light on these reactions. Identification of a major side product of DIMBOA decomposition in aprotic solvents revealed that 2,3-dihydro-benzoxazole intermediate **7** and its acid-catalyzed rearrangement to **8** as likely side reactions (38). For DIMBOA, decreased MBOA yield is also correlated with decreased solvent donor number and decreased pH in protic solvents. For HDMBOA, an analogous side reaction is anticipated to contribute to k_4 , which may further react to form a variety of condensation and oxidation products.

Despite these limitations of the model, the rates of the two pathways k_1 and k_2 are not apparently dependent on these unidentified side reactions, as similar values were found in both limits $k_4 = 0$ and $k_5 = 0$ (Figure 4, panels b and c). Therefore, the value of k_3 in the $k_5 = 0$ limiting model establishes an upper

limit for the decomposition rate of intermediate **9** to MBOA at 24 °C (Figure 5). Details may also be seen in Supplementary Table 1.

Although this report is the first direct observation of intermediate **9**, a similar *o*-imidoquinone decomposition intermediate has been proposed as the chemically active species for alkylation activity of ADIBOA (**4**, R = OAc in the basic structure in Scheme 1), a synthetic DIMBOA derivative (10, 39). That species was inferred from the identity of the addition products from the reaction of ADIBOA with various biologically relevant nucleophiles. As a similar *O*-alkyl benzoxazinone derivative, the natural aglycone HDMBOA would be expected to react similarly with biologically relevant nucleophiles, yet analogous reactivity studies have not been reported. The stability of **9**, and the synthetic analogue **10**, at moderately acidic pH and in methanol will facilitate the further characterization of this reactivity. Preliminary data (not shown) confirm that intermediate **9** is very electrophilic, and more detailed reactivity studies with various nucleophiles are underway.

Activity of HDMBOA and Its Analogues on *A. tumefaciens* Xenognosis. To evaluate the critical structural features of HDMBOA responsible for inhibiting xenognosis (host recognition) by *Agrobacterium*, the benzoxazinone derivatives **3** and **14** were prepared and evaluated for inhibition of *vir* gene induction (Figure 6). Hydroxamic acids such as **14** are known to possess moderate redox activity (7) and chelate iron(III), two chemical activities often associated with benzoxazinone biological activity. Although **14** is stable under the assay conditions, its

presence did not alter *vir* gene expression even at very high concentrations (Figure 6). Likewise lactam **3** (HMBOA), which shares ring-chain tautomer aldehyde functionality with HDMBOA and DIMBOA (**7**) but lacks the labile N–O bond, is also stable in the assay media and shows no inhibitory activity (Figure 6). The analysis of these analogues suggested that hydrolytic instability correlated with *vir* gene inhibition by the benzoxazinones.

Since the accumulating stable decomposition product MBOA does not inhibit *vir* gene expression in *A. tumefaciens*, it was possible to test the decomposition intermediates directly by allowing HDMBOA to decompose. UV spectroscopy and HPLC were used to monitor the decomposition of HDMBOA at pH 7.0, and after 20 min when no HDMBOA remained, the resulting solution was immediately added to pH 5.5 induction media containing *A. tumefaciens* strain A348 carrying a $P_{virB}::lacZ$ reporter gene construct. This sample possessed an apparent IC_{50} of $\leq 44 \pm 5 \mu\text{M}$. Under these conditions, the IC_{50} of HDMBOA is $4 \mu\text{M}$, 10-fold lower than the upper limit estimate of **9** (Figure 6). The ^1H NMR decomposition data in Figure 3 suggests that the maximal concentration of **9** is only 30–40% of the initial HDMBOA concentration after 20 min at pH 7.0, but even with this correction, there appears to be a difference. Three possibilities are apparent at this point: (i) HDMBOA is more active than **9**. This proposal seems unlikely as the removal of one oxygen atom from either of two different sites on the skeleton, analogues **3** and **14**, both stabilize the molecule against rearrangement into **9** and eliminate all biological activity. Alternatively, (ii) other decomposition products, in addition to **9**, may contribute to the observed activity. Since we detect no other intermediate at pH 7 where the above decomposition experiment was conducted, and the kinetic analyses implicate no other intermediates of any significant concentration, the compound would have to be active at very low concentrations. At this point, this possibility also seems unlikely. Finally, (iii) HDMBOA has no direct activity against VirA. It may well be that HDMBOA facilitates the delivery of **9** across cell membranes, past the degradative machinery, and/or specifically to the site of action on the VirA kinase to enhance the activity, but this third possibility implicates **9** as the active inhibitor.

To further evaluate the role of the *o*-imidoquinone functionality of **9**, the *vir* inhibitory activity of the

synthetic *o*-imidoquinone analogue **10** was investigated. This compound lacks the oxo-aldehyde, which eliminates *o*-imidoquinone decomposition by path k_3 (Figure 4) and renders the compound stable to the assay conditions. Moreover, the additional ring methyl group was required by the *o*-imidoquinone synthetic procedure, but even with these modifications, **10** displayed significant *vir* inhibitory activity and an IC_{50} of $36 \pm 3 \mu\text{M}$. Moreover, the Hill factors for HDMBOA, DIMBOA, and **10** were close to 1.0, implying that all three compounds act on the same single step of the VirA–VirG signal transduction pathway. Therefore, at this level of analysis, intermediate **9** appears significantly more accepting of structural perturbation than HDMBOA and further implicates it as the actual inhibitor. Further structure/activity relationships and direct evaluation with the VirA/VirG receptor will now allow this hypothesis to be critically evaluated.

HDMBOA as a Natural Pro-Anti-Xenogonin. The exuded natural product HDMBOA has been implicated in a novel mechanism of regulating host recognition, or xenogonin, in *A. tumefaciens*. The novelty is associated with both the environmental sensing target for HDMBOA and the delivery of a physically unstable compound to mediate transient local exposure to the pathogen. Our detailed analysis of the decomposition mechanism however revealed the major decomposition intermediate, *o*-imidoquinone **9**, a compound now expected to accumulate rapidly within the neutral pH environment of the cytoplasm of proximal bacterial cells. Initial structure–activity relationships suggest that the activity of the HDMBOA analogues correlate with their physical instability. Since the *o*-imidoquinone skeleton constitutes an active inhibitor, these data raise the very interesting possibility that HDMBOA functions as a “pro-drug” to deliver the *o*-imidoquinone to the bacterial cytoplasm.

Two issues are critical to this model. First, while *Agrobacterium* is exploited biotechnologically to transform monocots, it is not a pathogen of maize. Not surprisingly, innate immunity is not the only line of defense, and while washing HDMBOA from the seedlings improves transformation efficiency, there may well be other factors limiting host/pathogen interaction. Our initial data also suggest that HDMBOA and the *o*-imidoquinone target the conserved kinase domain of the two-component signaling elements (data not shown), suggesting that the site of action may be

the more general environmental-sensing modules of bacteria. Further investigations of other two-component signaling systems will be necessary to understand host/pathogen specificity. Second, these data raise questions about the mechanism of DIMBOA, the most well-known of the benzoxazinones, in *vir* inhibition, as there is no direct evidence for DIMBOA decomposition to an *o*-imidoquinone. However, *in vivo* metabolic activation of the hydroxamic acid as a leaving group is expected to trigger the *o*-imidoquinone decomposition pathway in this compound (10). The weaker activity of DIMBOA in *A. tumefaciens* may be the result of a slower activation event or poorer cellular uptake of the charged DIMBOA. Metabolic activation of DIMBOA will need to be further explored. Despite this uncertainty, HDMBOA, not DIMBOA, is the primary benzoxazinone likely to be encountered by bacterial pathogens as the exclusive component of maize root organic exudates (15). Moreover, biotic stress has been demonstrated to induce accumulation of HDMBOA-Glc in maize leaves, even near lesions caused by herbivore feeding (20), which serve as critical sites of opportunity for bacterial pathogenesis.

Taken together, these data suggest a complete revision of our chemical and ecological understanding of the benzoxazinone class of antibiotics. A unique

model for HDMBOA-mediated inhibition of the host recognition machinery in *A. tumefaciens* is proposed. HDMBOA is released into the acidic soil media around the plant root tissues where it remains within a steady-state zone for extended periods. As an uncharged aromatic compound, it is expected to cross the bacterial membrane into the cytoplasm easily. The neutral pH of the cytoplasm induces clean rearrangement to *o*-imidoquinone **9**. Our current evidence suggests that **9** inhibits VirA–VirG two-component sensing, in this case making the maize seedling invisible to *A. tumefaciens* colonies. In this model, HDMBOA functions as a “pro-drug” or “pro-anti-xenogonin” uniquely targeting bacterial environmental sensing rather than cellular viability. While it is possible that undetected decomposition intermediates, or side reactions of **9** with cellular metabolites, may play a role in the precise mechanism of inhibition, it is equally clear that *vir* inhibitory activity is tied to the unique instability of HDMBOA. Although this study has focused on VirA–VirG, this activity may indeed be general to two-component kinase sensing in other bacterial pathogens. Therefore, this discovery opens the door for developing stable structures or adapting this decomposition mechanism for the design of related pro-drug strategies for antibiotic development.

METHODS

Chemical Synthesis. Detailed synthesis and characterization of compounds used in this study is supplied in Supporting Information.

Measurements of Decomposition Kinetics by NMR. All NMR kinetic analyses were performed at 500 MHz for protons and 24 °C, employing a 50 mM HDMBOA solution in DMSO-*d*₆ diluted 3:20 (v/v) with the corresponding 0.1 M (citrate, phosphate, or carbonate) buffer in D₂O adjusted to pH 3.5, 5.5, 7.0, 8.5, 10.0, or 11.5. Samples in CD₃OD were prepared similarly.

Consistent with previous reports, the reaction developed an orange color and a major intermediate with the following spectral characteristics accumulated transiently: ¹H NMR (D₂O and DMSO-*d*₆): δ 3.82 (s, 3H, OCH₃), 5.54 (d, *J* = 2.0, 1H), 5.75 (s, 1H, CH(OH)₂), 6.33 (dd, *J* = 10.2, 2.0, 1H), 6.40 (d, *J* = 10.2, 1H); ¹³C NMR (D₂O and DMSO-*d*₆): δ 140.1 (C6), 129.3 (C⁵), 98.7 (CH(OH)₂), 95.1 (C³), 59.0; UV (H₂O) λ_{max} = 334 nm, ε = 2200 ± 200 cm⁻¹ M⁻¹.

The primary final product, 6-methoxy-3*H*-benzoxazol-2-one (MBOA, **6**), was assigned by comparison to published NMR spectra (40, 41): ¹H NMR (D₂O and DMSO-*d*₆) δ 3.95 (3H), 6.71 (dd, *J* = 2.4, 8.6, 1H), 6.88 (d, *J* = 2.4, 1H), 6.97 (d, *J* = 8.6, 1H); ¹³C NMR (D₂O and DMSO-*d*₆) δ 88.3, 112.0, 112.6. These data were identical to an authentic standard (Sigma): ¹H NMR (D₂O and DMSO-*d*₆) δ 3.66 (3H), 6.71 (dd, *J* = 2.4, 8.6, 1H), 6.88 (d, *J* = 2.4, 1H), 6.97 (d, *J* = 8.6, 1H), consistent with previous characterization of HDMBOA decomposition (17).

¹H NMR (D₂O and DMSO-*d*₆) signals corresponding to HDMBOA (7.16 ppm), MBOA (6.97 ppm), and the intermediate

(6.40 ppm) were followed over time with the earliest time being 5 min. The relative ratio of each species was normalized to the intensity of the internal 2.50 ppm DMSO signal. For the rapid decomposition at pH 7.0, a pH 3.5 standard prepared in the identical manner was used to verify the initial concentration.

Kinetic Measurements of Decomposition by UV Spectroscopy. Ultraviolet (UV) spectra were recorded with a Jasco UV/VIS spectrophotometer (Tokyo, Japan) in quartz cuvettes with an optical pathlength of 1 cm. Data processing was accomplished with the Jasco software and Excel (Microsoft Corp.) spreadsheet for calculation of extinction coefficients and normalization. Least-squares fitting of the data was performed with Origin 7.0 (OriginLab Corp.) using the exponential function $y = A \exp(-kx)$. An authentic sample of MBOA was characterized with the following characteristics by UV: (methanol) λ_{max} = 232 nm (10 000 ± 50 M⁻¹ cm⁻¹), λ_{max} = 290 nm (6000 ± 50 M⁻¹ cm⁻¹); UV (H₂O pH 5.5–7.0) λ_{max} = 230 nm, λ_{max} = 286.5 nm (5650 ± 60 M⁻¹ cm⁻¹); UV (H₂O pH 9.0) λ_{max} = 235 nm, λ_{max} = 290 nm (6000 ± 90 M⁻¹ cm⁻¹); UV (H₂O pH 12.9) λ_{max} = 240.5 nm, λ_{max} = 294 nm (6000 ± 60 M⁻¹ cm⁻¹).

For kinetics measurements, a standard solution of HDMBOA in THF was rapidly diluted 1000-fold in 0.1 M buffer at pH 3.5 (citrate), 5.5 (MES), 6.0, 6.5, 7.0 (phosphate), 9.0, 10.0 (carbonate), or 12.9 (NaOH). The earliest measurable time point was 6.0 s. The absorbance at 262 nm (HDMBOA) and 286 nm (MBOA) was monitored over several hours at ambient temperature (24 °C). Full spectra of identically prepared samples were obtained before and after measurement. Assuming that

only these two species contributed significantly at an observed wavelength λ , the concentrations of HDMBOA and MBOA were derived from the experimental data by simultaneous solution of the relation: $\text{Abs}_\lambda = (\epsilon_{\lambda, \text{HDMBOA}})(c_{\text{HDMBOA}}) + (\epsilon_{\lambda, \text{MBOA}})(c_{\text{MBOA}})$ for both wavelengths. Extinction coefficients were calculated from pure samples at each pH. Between pH 3.5 and 7.0, this correction did not significantly alter the calculated rate constant. However, above pH 7.0, MBOA was the dominant species before a single spectrum could be measured; thus, the contribution of MBOA at 262 nm was significant. As with the NMR kinetics, initial HDMBOA concentration was determined from preparation of an identical sample at pH 3.5. At a pH ≥ 7.0 , the final spectra matched the spectra of pure MBOA at the expected concentration for clean conversion, indicating high yield. Although side products distorted the final spectra of HDMBOA at pH 3.5, the kinetics for disappearance of HDMBOA closely matched the NMR data. To estimate the critical pK_a for decomposition, pseudo-first-order rate constants were fit to a monoprotic acid model by nonlinear least-squared fitting to the equation in Figure 2 using Origin 7.0 (OriginLab, Inc.) using the UV and NMR rate data without weighting. The rate data at pH 13.0 were excluded from this analysis due to larger error.

Kinetic Modeling of Decomposition. The kinetic model was resolved using the Gepasi kinetic simulation package (42–44), a program that employs input of a kinetic model to solve the corresponding differential equations. This program additionally fits the transient concentrations of observed species at each time point in the decomposition pathway by nonlinear least-squares fitting. The built-in Levenberg–Marquardt algorithm (set at 8000 iterations) was used for optimization of rate constants describing the kinetic profiles. To ensure convergence to a global minimum in parameter space, initial guesses were defined by the “multistart” option, which automatically restarts the minimization (option set to 300 starts) using randomly generated initial values for the rate constants. Initial concentrations of the intermediate and MBOA were set to zero at time zero. This software permitted rapid evaluation of many possible kinetic models. The model presented in Figure 4 represents a best fit to the kinetic data under all observed conditions.

Activity Assays. Activity data and IC_{50} values were measured from whole cell gene expression assays as described previously (15) using strain A348(pSW209) (carrying reporter gene fusion $P_{\text{virB}}::\text{lacZ}$) in a minimal induction media (IM): 10 g L^{-1} fructose, 1.0 g L^{-1} NH_4Cl , 150 mg L^{-1} MgSO_4 , 10 mg L^{-1} CaCl_2 , 1.5 mg L^{-1} FeSO_4 , 60 mg L^{-1} K_2PO_4 , 20 mg L^{-1} NaH_2PO_4 , and 100 mg L^{-1} CasAmino acids in 50 mM MES buffer at pH = 5.5. Acetosyringone (AS) was purchased from Sigma-Aldrich Co.; all other chemicals were purchased from Fisher Scientific, Inc., and all were used without further purification. The phenol inducer AS and all inhibitory compounds were dissolved in a minimal amount of DMSO and diluted to 100–500 mM concentrations (final DMSO concentration was identical for all samples in a given experiment and $<0.1\%$). Media was sterilized by autoclave treatment in a Market Forge Sterilmatic at 15 lb in^{-2} and $121 \text{ }^\circ\text{C}$ for 25 min.

Cells were grown overnight to mid-log phase at $28 \text{ }^\circ\text{C}$ in Luria–Bertani (LB) media supplemented with appropriate antibiotics. Cells were then pelleted by centrifugation and washed by suspension in IM. After a second centrifugation, the pellet was diluted in IM and inoculated into sterile culture tubes with the desired concentrations of inducer and inhibitor such that the final volume of each sample was 1.0 mL with a cell density of ca. $\text{OD}_{600} = 0.15$. Each sample was incubated at $28 \text{ }^\circ\text{C}$ for 8 h with shaking at 250 rpm at which time the assay was stopped by freezing at $-80 \text{ }^\circ\text{C}$ for 2–16 h. After slow thawing at $4 \text{ }^\circ\text{C}$, development of β -galactosidase activity followed from the procedure of Miller (45) using SDS and chloroform for cell permeabilization. Miller units represent gene expression activity normalized to cell growth. Error bars represent triplicate

samples. Data are representative of at least three similar independent experiments.

To obtain IC_{50} values, activity data were fit to a four-parameter logistic equation using Origin 7.0 (OriginLab Corp.), applying sufficient Simplex iterations for convergence (at least 1000). All curve-fits were verified with multiple starting parameters. Reported IC_{50} values represent the average of three independent experiments. For compounds reported as “ $\text{IC}_{50} > 400 \text{ mM}$,” inhibition was insignificant up to 400 mM. For these compounds, the additional order of magnitude in concentration required to define a sigmoidal curve (of Hill slope = 1) generally resulted in insolubility and/or growth inhibition; thus, it was concluded that activity was not biologically relevant.

Acknowledgment: We gratefully acknowledge the generous support of NIH GM 47369 and the invaluable collaboration with the Andrew Binns’ lab at the University of Pennsylvania. We dedicate this work to the lasting memory of W. Scott Chilton, North Carolina State University, and to his many fundamental contributions to the chemical biology of *Agrobacterium tumefaciens*.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Sicker, D., and Schulz, M. (2002) Benzoxazinones in plants: Occurrence, synthetic access, and biological activity, *Stud. Nat. Prod. Chem.* 27, 185–232.
- Sicker, D., Frey, M., Schulz, M., and Gierl, A. (2000) Role of natural benzoxazinones in the survival strategy of plants, *Int. Rev. Cytol.* 198, 319–346.
- Ham, J. H., and Bent, A. (2002) Recognition and defense signalling in plant/bacterial and fungal interactions, *Front. Mol. Biol.* 38, 198–225.
- Perez, F. J., and Niemeyer, H. M. (1989) Reaction of DIMBOA with amines, *Phytochemistry* 28, 1831–1834.
- Cuevas, L., Niemeyer, H. M., and Perez, F. J. (1990) Reaction of DIMBOA, a resistance factor from cereals, with α -chymotrypsin, *Phytochemistry* 29, 1429–1432.
- Sant’anna, C. M. R., Passos De Souza, V., and Santos De Andrade, D. (2002) Semiempirical simulation of a theta-class glutathione S-transferase-catalyzed glutathione attack to the allelochemical DIMBOA, *Int. J. Quantum Chem.* 87, 311–321.
- Atkinson, J., Morand, P., Arnason, J. T., Niemeyer, H. M., and Bravo, H. R. (1991) Analogues of the cyclic hydroxamic acid 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3-one: decomposition to benzoxazinolones and reaction with β -mercaptoethanol, *J. Org. Chem.* 56, 1788–1800.
- Perez, F. J., and Niemeyer, H. M. (1985) The reduction of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one by thiols, *Phytochemistry* 24, 2963–2966.
- Petho, M., Levai, L., and Romheld, V. (1997) The possible role of cyclic hydroxamic acids in the iron uptake of maize, *Novenytermeles* 46, 139–144.
- Hashimoto, Y., and Shudo, K. (1996) Chemistry of biologically active benzoxazinoids, *Phytochemistry* 43, 551–559.
- Ishizaki, T., Hashimoto, Y., Shudo, K., and Okamoto, T. (1982) Reaction of 4-acetoxy-1,4-benzoxazin-3-one with DNA—a possible chemical mechanism for the antifungal and mutagenic activities. *Tetrahedron Lett.* 23, 4055–4056.
- Hashimoto, Y., Shudo, K., Okamoto, T., Nagao, M., Takahashi, Y., and Sugimura, T. (1979) Mutagenicities of 4-hydroxy-1,4-benzoxazinones naturally occurring in maize plants and of related compounds, *Mutat. Res.* 66, 191–194.
- Sicker, D., Schneider, B., Hennig, L., Knop, M., and Schulz, M. (2002) Glycoside carbamates from benzoxazolin-2(3H)-one detoxification in extracts and exudates of corn roots, *Phytochemistry* 58, 819–825. [Erratum in *Phytochemistry* 60, 203.]

14. Larsen, E., and Christensen, L. P. (2000) Simple method for large scale isolation of the cyclic arylhydroxamic acid DIMBOA from maize (*Zea mays* L.), *J. Agric. Food Chem.* **48**, 2556–2558.
15. Zhang, J., Boone, L., Kocz, R., Zhang, C., Binns, A. N., and Lynn, D. G. (2000) At the maize/Agrobacterium interface: natural factors limiting host transformation, *Chem. Biol.* **7**, 611–621.
16. Schulz, M., Friebe, A., Kueck, P., Seipel, M., and Schnabl, H. (1994) Allelopathic effects of living quack grass (*Agropyron repens* L.). Identification of inhibitory allelochemicals exuded from rhizome borne roots, *Angew. Bot.* **68**, 195–200.
17. Grambow, H. J., Lueckge, J., Klausener, A., and Mueller, E. (1986) Occurrence of 2-(2-hydroxy-4,7-dimethoxy-2H-1,4-benzoxazin-3-one)- β -d-glucopyranoside in *Triticum aestivum* leaves and its conversion into 6-methoxy-benzoxazolinone, *Z. Naturforsch., C: Biosci.* **41**, 684–690.
18. Oikawa, A., Ishihara, A., Hasegawa, M., Kodama, O., and Iwamura, H. (2001) Induced accumulation of 2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one glucoside (HDMBOA-Glc) in maize leaves, *Phytochemistry* **56**, 669–675.
19. Oikawa, A., Ishihara, A., and Iwamura, H. (2002) Induction of HDMBOA-Glc accumulation and DIMBOA-Glc 4-O-methyltransferase by jasmonic acid in poaceous plants, *Phytochemistry* **61**, 331–337.
20. Oikawa, A., Ishihara, A., Tanaka, C., Mori, N., Tsuda, M., and Iwamura, H. (2004) Accumulation of HDMBOA-Glc is induced by biotic stresses prior to the release of MBOA in maize leaves, *Phytochemistry* **65**, 2995–3001.
21. Hansen, G., and Chilton, M. D. (1999) Lessons in gene transfer to plants by a gifted microbe, *Curr. Top. Microbiol. Immunol.* **240**, 21–57.
22. Heath, J. D., Charles, T. C., and Nester, E. W. (1995) Ti plasmid and chromosomally encoded two-component systems important in plant cell transformation by *Agrobacterium* species, in *Two-Component Signal Transduction* (Hoch, J. A., and Silhavy, T. J., Eds.) pp 367–386, ASM Press, Washington, DC.
23. Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction, *Annu. Rev. Biochem.* **69**, 183–215.
24. Inouye, M., and Dutta, R., (2003) *Histidine Kinases in Signal Transduction*, p xviii, p 520, Elsevier Science, San Diego, CA.
25. Hubbard, J., Burnham, M. K. R., and Throup, J. P. (2003) Pathogenicity and histidine kinases: approaches toward the development of a new generation of antibiotics, in *Histidine Kinases in Signal Transduction* (Inouye, M., and Dutta, R., Eds.) pp 459–481, Elsevier Science, San Diego, CA.
26. Macielag, M. J., and Goldschmidt, R. (2000) Inhibitors of bacterial two-component signalling systems, *Expert Opin. Invest. Drugs* **9**, 2351–2369.
27. Escobar, C. A., Kluge, M., and Sicker, D. (1997) Syntheses of 2-hydroxy-4,7-dimethoxy-2H-1,4-benzoxazin-3(4H)-one: a precursor of a bioactive electrophile from Gramineae, *Tetrahedron Lett.* **38**, 1017–1020.
28. Niemeyer, H. M., Bravo, H. R., Pena, G. F., and Corcuera, L. J. (1982) Decomposition of 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3-(4H)-one, a hydroxamic acid from Gramineae, in *Chemistry and Biology of Hydroxamic Acids* (Kehl, H., Ed.), pp 22–28, Karger, Basel and New York.
29. Atkinson, J., Arnason, J., Campos, F., Niemeyer, H. M., and Bravo, H. R. (1992) Synthesis and reactivity of cyclic hydroxamic acids. Resistance factors in the Gramineae, *ACS Symp. Ser.* **504**, 349–360.
30. Bredenberg, J. B. S., Hokanen, E., and Virtanen, A. I. (1962) The kinetics and mechanism of the decomposition of 2,4-dihydroxy-1,4-benzoxazin-3-one, *Acta Chem. Scand.* **2**, 135–141.
31. Smismann, E. E., Corbett, M. D., Jenny, N. A., and Kristiansen, O. (1972) Mechanism of the transformation of 2,4-dihydroxy-1,4-benzoxazin-3-ones and 2-hydroxy-2-methyl-4-methoxy-1,4-benzoxazin-3-one to 2-benzoxazolinone, *J. Org. Chem.* **37**, 1700–1704.
32. Honkanen, E., and Virtanen, A. I. (1961) Formation of 2(3)-benzoxazolinone from 2,4-dihydroxy-1,4-benzoxazin-3-one, *Acta Chem. Scand.* **15**, 221–222.
33. Berger, S., and Rieker, A. (1974) Identification and determination of quinones, *Chem. Quinonoid Compd.* Pt. 1, 163–229.
34. Nicolaou, K. C., Sugita, K., Baran, P. S., and Zhong, Y. L. (2002) Iodine(V) reagents in organic synthesis. Part 2. Access to complex molecular architectures via Dess–Martin periodinane-generated o-imidoquinones, *J. Am. Chem. Soc.* **124**, 2221–2232.
35. Nicolaou, K. C., Sugita, K., Baran, P. S., and Zhong, Y. L. (2001) New synthetic technology for the construction of N-containing quinones and derivatives thereof: Total synthesis of epoxyquinomycin B, *Angew. Chem., Int. Ed.* **40**, 207–210.
36. Nicolaou, K. C., Baran, P. S., Zhong, Y. L., and Sugita, K. (2002) Iodine(V) reagents in organic synthesis. Part 1. Synthesis of polycyclic heterocycles via Dess–Martin periodinane-mediated cascade cyclization: generality, scope, and mechanism of the reaction, *J. Am. Chem. Soc.* **124**, 2212–2220.
37. Woodward, M. D., Corcuera, L. J., Helgeson, J. P., and Upper, C. D. (1978) Decomposition of 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one in aqueous solutions, *Plant Physiol.* **61**, 796–802.
38. Bravo, H. R., and Niemeyer, H. M. (1986) A new product from the decomposition of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA), a hydroxamic acid from cereals, *Heterocycles* **24**, 335–337.
39. Ishizaki, T., Hashimoto, Y., and Shudo, K. (1992) Importance of the 2-hydroxy group for the reactions of an acetate of a naturally occurring prohibition, 4-acetoxy-2-hydroxy-2H-1,4-benzoxazin-3(4H)-one, with nucleophiles, *Heterocycles* **34**, 651–656.
40. Nagao, T., Otsuka, H., Kohda, H., Sato, T., and Yamasaki, K. (1985) Benzoxazinones from *Coix lachryma-jobi* var. *ma-yuen*, *Phytochemistry* **24**, 2959–2962.
41. Nachman, R. J. (1982) Convenient preparation of 2-benzoxazolinones with 1,1-carbonyldiimidazole, *J. Heterocycl. Chem.* **19**, 1545–1547.
42. Mendes, P. (1993) Gepasi—a software package for modeling the dynamics, steady-states and control of biochemical and other systems, *Comput. Appl. Biosci.* **9**, 563–571.
43. Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3, *Trends Biochem. Sci.* **22**, 361–363.
44. Mendes, P., and Kell, D. B. (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation, *Bioinformatics* **14**, 869–883.
45. Miller, J. H. (1972) Assay of beta-galactosidase, in *Experiments in Molecular Genetics*, pp 352–355, Cold Spring Harbor Lab. Press, Plainview, N.Y.

Gene Silencing Activity of siRNAs with a Ribo-difluorotoluy Nucleotide

Jie Xia[†], Anne Noronha[†], Ivanka Toudjarska[†], Feng Li[‡], Akin Akinc[†], Ravi Braich[†], Maria Frank-Kamenetsky[†], Kallanthottathil G. Rajeev[†], Martin Egli[‡], and Muthiah Manoharan^{†,*}

[†]Anylam Pharmaceuticals, Inc., 300 Third Street, Cambridge, Massachusetts 02142, and [‡]Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville, Tennessee 37232

ABSTRACT Recently, chemically synthesized short interfering RNA (siRNA) duplexes have been used with success for gene silencing. Chemical modification is desired for therapeutic applications to improve biostability and pharmacokinetic properties; chemical modification may also provide insight into the mechanism of silencing. siRNA duplexes containing the 2,4-difluorotoluy ribonucleoside (rF) were synthesized to evaluate the effect of noncanonical nucleoside mimetics on RNA interference. 5'-Modification of the guide strand with rF did not alter silencing relative to unmodified control. Internal uridine to rF substitutions were well-tolerated. Thermal melting analysis showed that the base pair between rF and adenosine (A) was destabilizing relative to a uridine-adenosine pair, although it was slightly less destabilizing than other mismatches. The crystal structure of a duplex containing rF-A pairs showed local structural variations relative to a canonical RNA helix. As the fluorine atoms cannot act as hydrogen bond acceptors and are more hydrophobic than uridine, there was an absence of a well-ordered water structure around the rF residues in both grooves. siRNAs with the rF modification effectively silenced gene expression and offered improved nuclease resistance in serum; therefore, evaluation of this modification in therapeutic siRNAs is warranted.

Short interfering RNAs (siRNAs) are 21–23 nucleotides long RNA duplexes. One of the strands of the siRNA duplex, the guide (antisense) strand, is incorporated into a complex of enzymes called the RNA-induced silencing complex (RISC). The small RNA component serves as a guide to identify a complementary sequence in a messenger RNA (mRNA) and causes the mRNA to be cleaved by RISC or prevents it from being translated into protein (1–4). The result is that the targeted gene is silenced.

Although unmodified siRNA duplexes have been used with success for gene silencing, chemical modification of one or both of the strands is desired for therapeutic applications to enhance stability and to improve pharmacokinetic properties. Numerous chemical modifications have been evaluated, but only a few nucleobase modifications have been tested for effects on siRNA activity (5–7). It is not clear which of these nucleobase modifications are recognized and accepted by the RNA-induced silencing complex, RISC, during the course of RNA interference.

The 2'-deoxydifluorotoluy nucleoside (dF) is a nonpolar nucleoside analogue invented by Kool and co-workers that is an isostere of thymidine (T) (8). In a DNA double helix, dF is destabilizing due to lack of hydrogen bonds. It is, however, more hydrophobic and stacks more favorably than thymidine (9). An NMR solution structure of a DNA dodecamer containing dF-A pairs revealed no significant deviations from B-form geometry (10). A canonical T-A base pair and the corresponding F-A pair are shown in Figure 1. dFTP serves as a substrate for *Escherichia coli* DNA polymerase and is inserted into a template opposite adenosine (A) with an efficiency 40-fold lower than that of dTTP, but with selectivity nearly as high as that observed for dTTP (8, 11). Replication of dF-containing templates occurs in bacterial cells with high fidelity (12).

*To whom correspondence should be addressed.
E-mail: mmanoharan@anlylam.com.

Received for review February 4, 2006
and accepted April 10, 2006.

Published online April 17, 2006.

10.1021/cb600063p CCC: \$33.50

© 2006 American Chemical Society

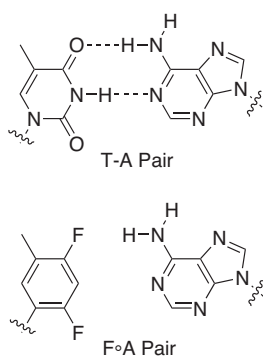


Figure 1. A canonical T-A base pair and the corresponding F-A pair.

is an evaluation of the thermal stability of duplexes containing the rF modification and a high-resolution crystal structure of a self-complementary 12-mer RNA duplex containing two rF-A pairs. A guide strand with an rF opposite to the cleavage site on the mRNA target directed cleavage of the mRNA at the predicted site. This experiment demonstrated the importance of stacking interactions rather than hydrogen bonding at this critical functional site. The biostability of siRNA duplex in serum is critical to therapeutic utility. The rF modification provided resistance to endonuclease cleavage, the first example of an induced stability that may be due to the hydration around the modification revealed by the crystal structure.

RESULTS AND DISCUSSION

The rF monomer was synthesized as described in Methods (Figure 2). In general, the thermal stability of a duplex is increased when C-5 of a pyrimidine has a methyl group as compared to a hydrogen (13); thus, we chose to evaluate the fluorine-substituted analogue of thymidine, rather than the uridine (U). Parsch and Engels reported a similar synthesis of a phosphoramidite of 2,4-difluorobenzene riboside and its incorporation to a 12-mer RNA (14). The siRNA duplexes (Table 1) were synthesized using standard RNA oligonucleotide synthesis and purification protocols (15, 16). Thermal melting profiles of the RNA duplexes were used to evaluate the effect of the rF modification on RNA duplex stability (Table 1). The rF-A pair lowered the melting temperature (T_m), but was slightly less destabilizing than the mismatches tested. For example, the fully complementary duplex had a T_m of 73 °C. Substitution

Kool rationalized that hydrogen bonds are less important, and nucleotide shape is more important, in the fidelity of DNA replication than originally believed. (12)

We have evaluated the effects of the ribonucleotide analogue of 2,4-difluorotoluy (rF) on *in vitro* RISC-mediated gene silencing activity of small interfering RNAs (siRNAs). Also reported

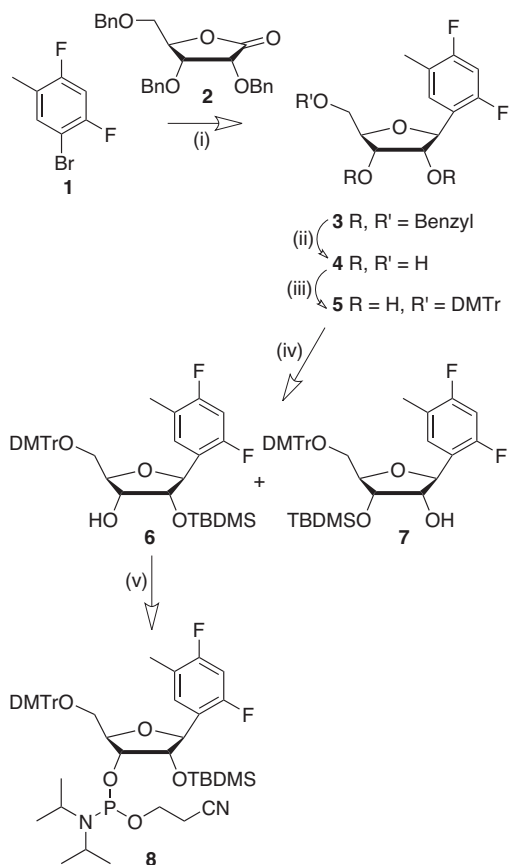


Figure 2. Scheme for the synthesis of ribo-difluorotoluy nucleoside phosphoramidite (8). (i) *n*-BuLi/THF, -78 °C, 3 h, Ar atm, then 0 °C, 3 h; Et₃SiH-BF₃-Et₂O/CH₂Cl₂, -78 °C to RT, 16 h, Ar atm., 81%; (ii) BCl₃/CH₂Cl₂, -78 to -40 °C, 4 h, Ar atm., 74%; (iii) DMTr-Cl, DMAP/Py, RT, 20 h, 71%; (iv) AgNO₃, Py, TBDMSCl/THF, RT, 12 h, 85% (combined, 6:7 2:8); (v) *i*-Pr₂NP(Cl)O(CH₂)₂CN, DMAP, *i*-Pr₂NEt/CH₂Cl₂, RT, Ar atm, 91%.

of rF for U at position 10 on the guide strand of duplex I (Table 1, duplex VII) decreased the T_m by 5.5 °C relative to the unmodified duplex, whereas substitution of cytidine (C) for U (Table 1, duplex X) lowered the T_m by 6.5 °C. Substitution with either purine lowered the T_m by 7.5 °C. Substitution of U with ribo 2,4-difluorobenzene nucleoside in a 12-mer RNA duplex decreases the thermal stability of the duplex by 10 °C (14). Though a direct comparison of thermal stability of a 19-mer duplex with two T overhangs on either end (Table 1, duplex VII) with a blunt-ended 12-mer duplex is difficult, the rF modification appears to be more stabilizing than the 2,4-difluorobenzene modification.

TABLE 1. siRNA duplex sequences, thermal stability, and biological activity

ID	Duplex sequence ^a	T_m (°C) ^b	IC ₅₀ (nM) ^c
I	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACUCAUGAAGCU	73	0.21
II	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACUCAUGAAGCF	nd	0.27
III	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACUCAFGAAGCU	nd	0.27
IV	CUUACGCUGAGUACUUCGATT TTGAAFGCGACUCAUGAAGCU	nd	0.58
V	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACUFAUGAAGCU	nd	>30
VI	CUUACGCUGAGUACUUCGATT TTGAAUGCAGAFUCAUGAAGCU	nd	>30
VII	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACFCAUGAAGCU	67.5	1.2
VIII	CUUACGCUGAGUACUUCGATT TTGAAUGCAGACAUGAAGCU	65.5	>30
IX	CUUACGCUGAGUACUUCGATT TTGAAUGCAGCGCAUGAAGCU	65.5	>30
X	CUUACGCUGAGUACUUCGATT TTGAAUGCAGCCCAUGAAGCU	66.5	0.92

^asiRNA duplexes are shown with the sense strand above (written 5' to 3') and the guide strand below (written 3' to 5'). All nucleotides were ribo with the exception of deoxythymidines (indicated by T). The rF modification is indicated by an F. ^b T_m is the midpoint of the thermal melting transition ± 0.5 °C; nd indicates not determined. ^cThe IC₅₀ value is for luciferase silencing in HeLa cells.

Modified siRNAs II–VII (Table 1) have guide strands complementary to firefly luciferase. These modified siRNAs were compared with unmodified control siRNA (Table 1, duplex I) for the ability to silence luciferase activity *in vitro*. The 50% inhibitory concentrations (IC₅₀) for duplexes tested are listed in Table 1, and selected data is shown in Figure 3. When the rF was positioned at the 5'-end of the guide strand (Table 1, duplex II), silencing efficiency was similar to that observed with an unmodified duplex. When rF was substituted for U at position 7 (Table 1, duplex III), silencing was essentially the same as that of unmodified control (Table 1, duplex I). Activity of the duplex with rF at position 16 (Table 1, duplex IV) was reduced by approximately 2-fold compared with control. Substitution of rF

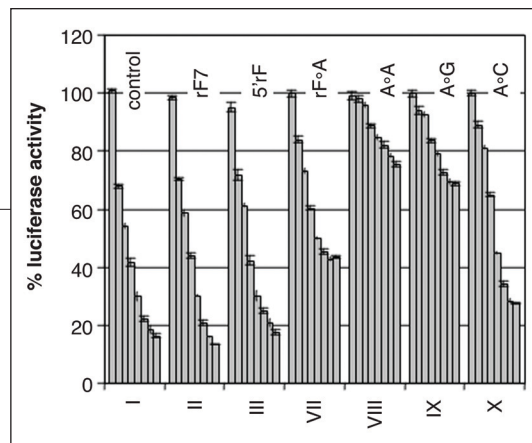


Figure 3. Effect of the rF modification vs mismatches on siRNA activity. siRNA duplexes targeted luciferase mRNA. siRNA activity was observed as a decrease in percent luciferase luminescence over concentrations ranging from 0.4 to 30 nM. Firefly/*Renilla* luciferase expression ratios were used to determine percent gene silencing relative to untreated controls

for C near the center of the duplex was not tolerated (Table 1, duplexes V and VI). Interestingly, rF or C could be substituted for uridine at position 10 (Table 1, duplexes VII and X) without much loss of activity, but purine substitution at this position was not tolerated (Table 1, duplexes VIII and IX). In agreement with previous reports (6, 7), the effect of mismatches on silencing depends on the mismatch position, the sequence surrounding and opposite the mismatch, and the geometry of the mismatch, independent of whether the nucleobase is canonical or noncanonical.

The RISC mechanism requires 5'-phosphorylation for recruitment of the guide strand and the guide strand-targeted endonuclease (*argonaute2*) cleavage of the target mRNA (17). Previous work has shown that chemically modified siRNAs that cannot be 5'-phosphorylated at the guide strand are inactive (18, 19). The guide strand containing a 5'-rF modification was efficiently labeled with 5-[γ -³²P]ATP by T4 polynucleotide kinase (Figure 4) and was effective in the *in vitro* silencing assay (Figure 3).

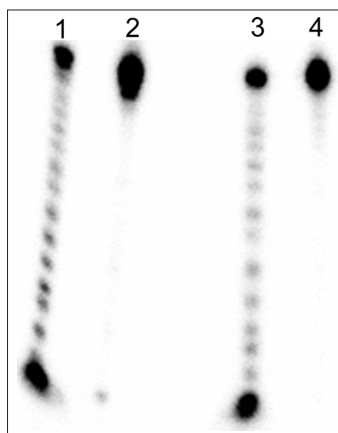


Figure 4. Phosphorylation of an oligonucleotide with a 5' rF modification: lane 1, alkaline hydrolysis of 5'-³²P-end labeled unmodified guide strand of duplex I (Table 1); lane 2, reaction of unmodified guide strand of duplex I with 5-[γ -³²P]ATP and T4 polynucleotide kinase; lane 3, alkaline hydrolysis of 5'-³²P-end labeled 5'-rF-guide strand of duplex II (Table 1); lane 4, reaction of 5'-rF-guide strand with 5-[γ -³²P]ATP and T4 polynucleotide kinase. The guide strand of duplex II was labeled with ³²P by T4 polynucleotide kinase as efficiently as an unmodified oligonucleotide.

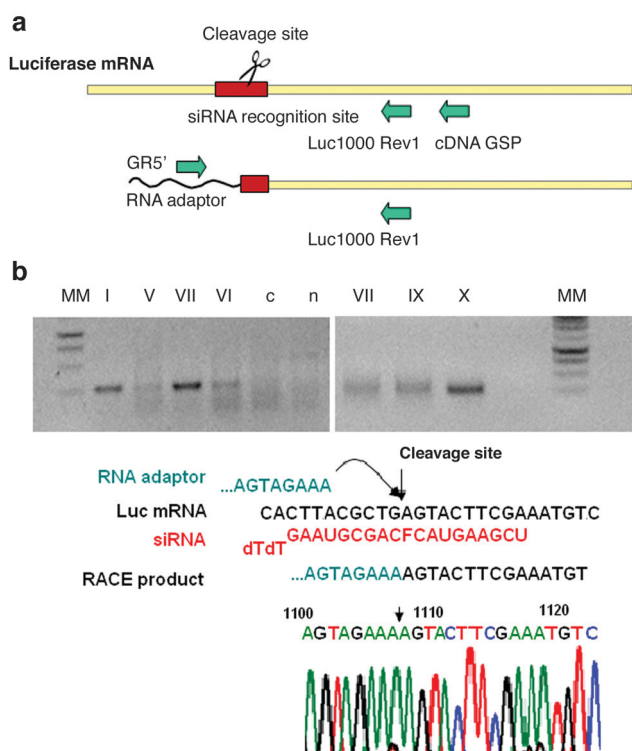


Figure 5. RACE analysis of cleavage products. **a)** Schematic representation of modified RACE analysis of siRNA-induced cleavage of luciferase mRNA. Primers used for amplification of cleavage products are indicated by the green arrows. **b)** Cells were treated with luciferase-targeting siRNA duplexes and RACE-PCR amplification products were analyzed using agarose gel electrophoresis. The lanes labeled MM are molecular weight markers, the lane labeled c indicates cells transfected with unrelated siRNA, n indicates cells alone. The cleavage site was verified by sequencing. The arrow on the chromatogram indicates the precise cleavage site where the RNA adaptor was ligated.

In a 21-mer siRNA duplex with mRNA, the cleavage site on the mRNA is located between nucleotides 10 and 11 from the 5'-end of the guide strand (20, 21). To insure that inhibition of luciferase expression observed with rF-modified siRNAs was due to siRNA-directed cleavage, specific mRNA cleavage products were characterized using a modified 5'-RACE (rapid amplification of cDNA ends) technique as previously described (23–25) (Figure 5). Sequencing of the PCR products demonstrated that siRNA-mediated cleavage occurred at the predicted site after treatment with the unmodified duplex as well as duplex VII (rF at position 10). No cleavage product was detected after treatment with duplex V (rF at position 9), in agreement

with the results of the luciferase silencing assay (Table 1, Figure 3). Some product was observed after treatment with duplex VI, with rF substituted for C at position 11. No luciferase activity was detected in the expression assay with duplex VI; the discrepancy was most likely due to the higher sensitivity of the RACE assay. No cleavage product was observed in cells treated with unrelated siRNA or cells alone.

The biostability of siRNA duplex in serum is critical to therapeutic utility. The rF modification provides resistance to endonuclease cleavage (Figure 6). In this experiment, duplexes labeled at the 5'-end of one strand were incubated in human serum and the endonuclease cleavages were visualized by PAGE. The cleavage observed in the control unmodified duplex between U₁₆ and A₁₇ near the 3'-end was not observed in the rF-modified strand. Our data suggest that the rF modification increases the resistance to endonuclease activity of an siRNA without compromising silencing activity. Two other sites, U7 and C11, were cleaved by endonuclease and serve as internal controls for cleavage at the modified residue. The modified strand also undergoes degradation by exonucleases. In summary, whereas endonuclease cleavage occurred at U7 and C11, the modified site 16 was resistant.

To visualize the geometry of the rF^oA pair in an RNA duplex, we determined the crystal structure of the self-complementary duplex with two rF^oA pairs, 5'-CGCFAAUUAGCG-3', to 1.61 Å resolution (residues of

with the results of the luciferase silencing assay (Table 1, Figure 3). Some product was observed after treatment with duplex VI, with rF substituted for C at position 11. No luciferase activity was detected in the expression assay with duplex VI; the discrepancy was most likely due to the higher sensitivity of the RACE assay. No cleavage product was observed in cells treated with unrelated siRNA or cells alone.

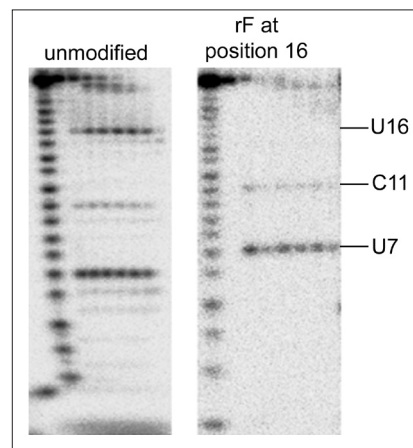


Figure 6. The rF modification imparted resistance to endonuclease cleavage to an siRNA duplex. The guide strand (5'-UCGAAGUACUCAGCGFAAGdTd-3') contained rF at position 16. The first panel shows the unmodified control siRNA. The second panel shows the rF-modified duplex IV. The positions of the observed endonuclease cleavages after U7, C11, and U16 (or rF16) are labeled. The first lane in each panel is alkaline hydrolysis of the labeled guide strand. The second lane was the duplex incubated in phosphate-buffered saline for 240 min. The subsequent lanes show cleavage observed after incubation for 0, 15, 30, 60, 120, and 240 min in 90% human serum.

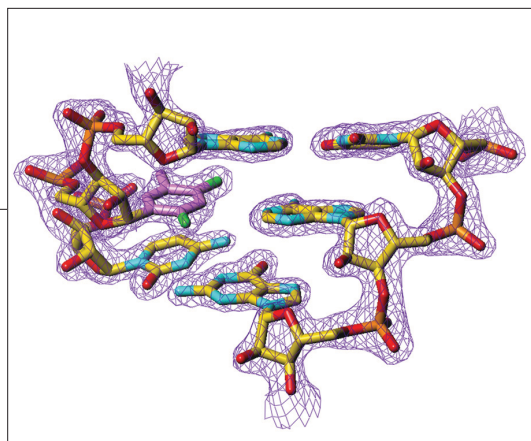


Figure 7. Example of the electron density in the region of the rF modification. Fourier ($2F_o - F_c$) sum electron density around the final structure in the region of the rF4•A21 pair. Carbon atoms of the rF moiety are highlighted in magenta and fluorine atoms in green.

strands 1 and 2 are numbered 1–12 and 13–24, respectively). The structure of an RNA duplex of similar sequence, $(5'-CGCgAAUuA\underline{GCG}-3')_2$, with two guanosine•adenosine (G•A) mismatch pairs (underlined) was determined a decade ago (26). Recently, we collected diffraction data to better than 1.1 Å resolution for crystals of the dodecamer containing G•A mismatches, and it was anticipated that replacement of G4 by rF and the formation of two rF•A pairs would not hamper crystallization (Figure 7). In the duplex, the rF and A moieties are more than 0.8 Å apart, on average, along the long dimension of the pair (Figure 8, panels a and b) relative to the positions of bases in the central U-A pairs (Figure 8, panel c) as a consequence of the loss of hydrogen bonds and the

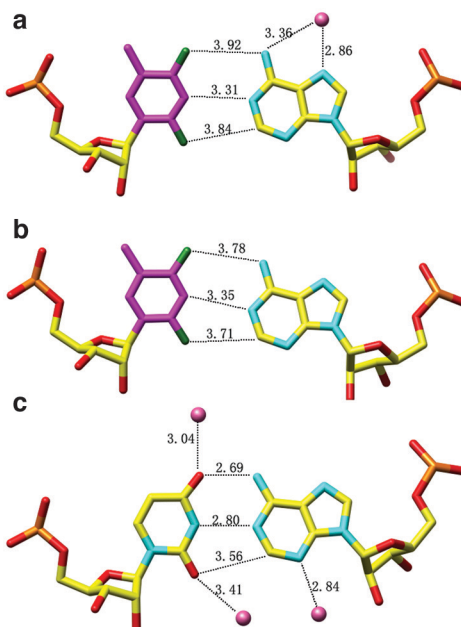
C–H...N interaction in the rF•A pair. The F4 of rF and the exocyclic N6 of A are at approximately the sum of their van der Waals radii. This leads to an opening of rF•A pairs (of 10°) compared to a U-A pair and a local widening of the major groove. Other structural changes include a marked roll between rF16 and A17 (–20°) and strong shearing in the rF16•A9 pair, with the rF moiety shifted 0.75 Å into the major groove relative to other base pairs (Figure 8, panel b). The same pair also exhibited a very strong propeller twist (–30°; –13° in the rF4•A21 pair; average value in the duplex, –17°). Unlike O2 and O4 of U, fluorine atoms in rF residues do not engage in hydrogen bonds to water molecules, and there was an absence of a well-ordered water structure around difluorotoluy moieties in both grooves.

Chemical modification of siRNA is required for therapeutic applications due to insufficient chemical stability and resistance to nuclease degradation of unmodified RNA and inefficient uptake of the siRNA into mammalian cells. Although there are virtually no technical limitations regarding chemical modification of the nucleobase, sugar, and phosphate, extensive interactions of siRNA with protein factors put constraints on the types of modifications and the number of residues that can be modified without severely impairing the efficacy of the siRNA. We have synthesized and

evaluated siRNA duplexes modified with the ribonucleotide analogue of 2,4-difluorotoluy (rF). Difluorotoluy was first synthesized and evaluated in the context of DNA duplexes by Kool and co-workers (8). The dF modification destabilizes DNA duplexes but is accepted by DNA polymerase with high fidelity (12).

In the fidelity of DNA polymerase, nucleotide shape appears to be more important than hydrogen bonding, as selectivity for the dF analogue is roughly equivalent to that of T (12). Our results suggest that hydrogen bonds are not as critical as shape in selectivity of siRNA-mediated RISC cleavage. Whereas substitution of either A or G was not tolerated in the guide strand opposite the cleavage site, neither rF nor C abolished silencing of luciferase in an expression assay (Figure 3, Table 1). A RACE assay was used to ensure that the mechanism of silencing was not altered by rF substitution (Figure 5, panel b). When rF replaced C at position 9 (Table 1,

Figure 8. Geometry and hydration of rF•A and U-A base pairs in the RNA dodecamer duplex. Base pairs (a) rF4•A21, (b) rF16•A9, and (c) A6-U19 are shown. Atoms are colored yellow, red, cyan, and orange for carbon, oxygen, nitrogen, and phosphorus, respectively. Carbon atoms of the difluorotoluene moiety are highlighted in magenta, fluorine atoms are green and water molecules are pink. Selected hydrogen bonds in the U-A base pair and the corresponding interactions in the rF•A pair are shown as dashed lines with distances in angstroms. The lengths of Watson–Crick hydrogen bonds and the O2(U) to C2(A) distance shown in panel c are average values based on the four central U-A pairs. In panel c, the water molecules hydrogen bonded to O2(U) and N3(A) are separated by 3.6 Å. The figures were generated with the program UCSF Chimera (26).



duplex V), no cleavage product was observed. An rF substitution at position 10 reduced silencing by about 6-fold compared to the unmodified siRNA, but a RACE product was observed and, as shown by sequencing of the product (data not shown), had the sequence expected for cleavage opposite positions 10 and 11 of the guide strand. To our knowledge, this is the first demonstration of RNA-mediated gene silencing using an siRNA chemically modified at the cleavage site with a nucleobase mimetic that does not form a Watson-Crick base pair.

The crystal structure of a duplex containing two rF^oA pairs indicates that the presence of an isolated rF^oA pair leads to a host of local changes that subtly

affect the geometry of the duplex both geometrically and thermodynamically compared to a duplex with canonical U-A pairs. In particular, local stacking interactions and water structure undergo changes relative to a Watson-Crick-paired duplex. On the basis of the structural and thermodynamic data, one can conclude that the duplex is not as rigid at the sites of rF incorporation as it would be if a Watson-Crick pair was present, consistent with the observations from a recent molecular dynamics simulation (27). In spite of these structural variations, the retention of silencing activity, recognition by the RISC enzymes and kinase, and improved endonuclease resistance in serum all warrant further evaluation of rF in therapeutic siRNAs.

METHODS

Synthesis of Ribo-difluorotoluyl Nucleoside

Phosphoramidite. The difluorotoluyl (rF) monomer **8** was synthesized as shown in Figure 2.

1. 2,3,5-Tri-*O*-benzyl-1-C-(2,4-difluorotoluene)- β -ribose (3**).** Compound **1** was prepared according to Kool's procedure (28). *n*-Butyl lithium (4.25 mL, 2.5 M in hexanes) was added to a cold solution of 5-bromo-2,4-difluorotoluene (2.19 g, 10.63 mmol) in anhydrous THF (50 mL) at -78°C and stirred at the same temperature for 3 h under an argon atmosphere. 2,3,5-Tri-*O*-benzyl-D-riboactone (**2**, 4.45 g, 10.63 mmol) in dry THF (17 mL) was added dropwise to the above solution and stirred at -78°C for 2 h and then at 0°C for 3 h. The reaction mixture was quenched with saturated NaHCO_3 solution and extracted with dichloromethane (3×120 mL). The organic phase was washed with saturated aqueous NaHCO_3 solution and brine, dried (Na_2SO_4), and concentrated to a crude residue that was dried under high vacuum for 1.5 h. $\text{BF}_3 \cdot \text{OEt}_2$ (4 mL) and Et_3SiH (5.1 mL) were added to a cold solution of the crude residue in anhydrous dichloromethane (80 mL) at -78°C and stirred at -78°C to ambient temperature under argon overnight. The reaction was quenched by addition of 1 N HCl, stirred at ambient temperature for 1 h, and subsequently neutralized with 1 N aqueous NaOH solution. The product was extracted into dichloromethane (3×100 mL), washed with saturated aqueous NaHCO_3 solution and then brine, dried over Na_2SO_4 , and concentrated to a crude residue that was applied to a column of silica gel and eluted with hexanes-ethyl acetate (4:1) to give pure compound **3** (4.57 g, 81%) as a syrup. Characterization of the product by NMR is described in Supporting Information.

2. 1-C-(2,4-Difluorotoluene)- β -ribose (4**).** BCl_3 (31 mL, 1 M in dichloromethane) was added to a cold solution of 2,3,5-tri-*O*-benzyl-1-C-(2,4-difluorotoluene)- β -ribose **3** (1.1 g, 2.08 mmol) in anhydrous dichloromethane (100 mL) at -78°C under an argon atmosphere. The reaction mixture was stirred at -78°C for 2.5 h and at -45°C for 1 h. The reaction was quenched with dichloromethane-methanol (50 mL, 1:1) and saturated ammonia-methanol solution. After concentration to a crude residue, the solution was applied to a column of silica gel and eluted with dichloromethane-methanol (5:1) to give compound **4** (400 mg, 74%) as a white solid. Characterization of the product by NMR is described in Supporting Information.

3. 5'-*O*-(4,4'-Dimethoxytrityl)-1-C-(2,4-difluorotoluene)- β -ribose (5**).** 4,4'-Dimethoxytrityl chloride (535 mg, 1.58 mmol) was added to a solution of 1-C-(2,4-difluorotoluene)- β -

ribose **4** (370 mg, 1.42 mmol) in anhydrous pyridine (3 mL) in the presence of 4-*N,N*-dimethylaminopyridine (DMAP, 40 mg) and stirred at ambient temperature under argon overnight. The reaction mixture was concentrated to a crude residue and co-evaporated with dry toluene (3×10 mL). The crude residue was applied to a column of silica gel saturated with 2% triethylamine in hexanes and eluted with hexanes-ethyl acetate (1.5:1) to give pure compound **5** (570 mg, 71%) as an amorphous white solid. Characterization of the product by NMR is described in Supporting Information.

4. 5'-*O*-(4,4'-Dimethoxytrityl)-2'-*O*-(*tert*-butyldimethylsilyl)-1-C-(2,4-difluorotoluene)- β -ribose (6**).** Anhydrous pyridine (907 μL) and AgNO_3 (235 mg, 1.35 mmol) were added to a solution of 5'-*O*-(4,4'-dimethoxytrityl)-1-C-(2,4-difluorotoluene)- β -ribose **5** (640 mg, 1.14 mmol) in anhydrous THF (8 mL), and the solution was stirred at ambient temperature for 20 min under argon atmosphere. *tert*-Butyldimethylsilyl chloride (235 mg, 1.48 mmol) in dry THF (3 mL) was added to the stirring mixture, and stirring was continued at ambient temperature for 3 h. Solids were filtered off, and the filtrate was concentrated to a crude residue that was applied to a column of silica gel and then eluted with hexane- Et_2O (4:1) to give compound **6** (360 mg, 46%), compound **7** (5'-*O*-(4,4'-dimethoxytrityl)-3'-*O*-(*tert*-butyldimethylsilyl)-1-C-(2,4-difluorotoluene)- β -ribose, 40 mg, 5%), and a mixture of compounds **6** and **7** (650 mg) as amorphous solid. Characterization of the product by NMR is described in Supporting Information.

5. 5'-*O*-(4,4'-Dimethoxytrityl)-2'-*O*-(*tert*-butyldimethylsilyl)-1-C-(2,4-difluorotoluene)- β -ribose-3'-*O*-caynoethyl-*N,N*-diisopropylphosphoramidate (8**).** 2-Cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (252 mg, 1.07 mmol) was added to a solution of 5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-(*tert*-butyldimethylsilyl)-1-C-(2,4-difluorotoluene)- β -ribose **6** (360 mg, 0.53 mmol), diisopropylethylamine (504 μL , 2.93 mmol), and DMAP (19 mg) in anhydrous dichloromethane (6 mL) and stirred at ambient temperature for 6 h under argon. The reaction mixture was concentrated to a crude residue, applied to a column of silica gel that was saturated with 2% triethylamine in hexanes, and eluted with hexanes-ethyl acetate (2:1) to give pure compound **8** (420 mg, 91%) as an amorphous white solid. Characterization of the product by NMR is described in Supporting Information.

Oligonucleotide Synthesis. The RNA molecules were synthesized on a 394 ABI synthesizer using the manufacturer's standard 93 step cycle with modifications described below. The dT-controlled pore glass (CPG) was prepacked, 1 μmol , 500 \AA from Prologo Biochemie GmbH. RNA phosphoramidites with fast deprotecting groups were obtained from Pierce Nucleic Acid Technologies or were synthesized as described above and were used at concentrations of 0.15 M in CH_2CN unless otherwise stated. Specifically, in addition to the rF amidite **8**, 3'-O-(2-cyanoethyl)-N,N'-diisopropyl-phosphoramidites of 5'-O-(4,4'-dimethoxytrityl)-2'-O-*t*-butyldimethylsilyl-N⁶-phenoxyacetyladenosine, 5'-O-(4,4'-dimethoxytrityl)-2'-O-*t*-butyldimethylsilyl-N²-*p*-isopropylphenoxyacetylguanosine, 5'-O-(4,4'-dimethoxytrityl)-2'-O-*t*-butyldimethylsilyl-N⁴-acetylcytidine, 5'-O-(4,4'-dimethoxytrityl)-2'-O-*t*-butyldimethylsilyluridine, and 5'-O-(4,4'-dimethoxytrityl)thymidine were used for the oligonucleotide synthesis.

The coupling times were 7.5 min for the RNA monomers and 10 min for the rF monomer. Details of other reagents were as follows: activator, 5-(ethylthio)-1*H*-tetrazole (0.25 M); Cap A, 5% phenoxyacetic anhydride/THF/pyridine; Cap B, 10% *N*-methylimidazole/THF; PO-oxidation, 0.02 M iodine in THF/water/pyridine. Detritylation was achieved with 3% TCA/dichloromethane.

The oligonucleotide was cleaved from the CPG, with simultaneous deprotection of base and phosphate groups, with 1.0 mL of a mixture of ethanolic ammonia (ammonia-ethanol, 3:1) for 16 h at 55 °C. The solution was decanted, lyophilized, and resuspended in 0.2 mL of triethylamine trihydrofluoride (TEA.3HF, Aldrich) and was incubated at 65 °C for 90 min to remove the *t*-butyldimethylsilyl (TBDMS) groups. The completely deprotected oligonucleotides were then precipitated from anhydrous methanol (MeOH, 0.4 mL).

The oligonucleotides were purified by vertical slab polyacrylamide gel electrophoresis (PAGE), generally 20% (w/v) acrylamide and 5% BIS (with respect to the mass of acrylamide). About 50 ODs were typically loaded onto preparative gels in the loading buffer of 80% formamide in 10 \times Tris-borate-EDTA buffer (TBE). The desired bands were excised and shaken overnight in 5 mL of 100 mM sodium acetate. The extracted oligonucleotides were desalted using C-18 Sep-Pak cartridges (Waters). The oligonucleotides were characterized by electrospray mass spectroscopy (ES/MS) and analytical anion-exchange HPLC and/or capillary gel electrophoresis (CGE). The LC-MS data for single strands are listed in Supplementary Table 1 in Supporting Information.

Thermal Denaturation (T_m) Studies. Molar extinction coefficients for the oligonucleotides were calculated according to nearest-neighbor approximations (units = $10^4 \text{ M}^{-1} \text{ cm}^{-1}$). Duplexes were prepared by dissolving lyophilized equimolar amounts of the complementary strands in phosphate-buffered saline (pH 7.0) to give a final concentration of 2.4 μM of each strand. The solutions were heated to 90 °C for 10 min and cooled slowly to room temperature before measurements. Prior to analysis, samples were degassed by placing them in a speed-vac concentrator for 2 min. Denaturation curves were acquired at 260 nm at a rate of heating of 0.5 °C min^{-1} using a Varian CARY spectrophotometer fitted with a 12-sample thermostated cell block and a temperature controller.

In Vitro Analysis of Luciferase Expression. HeLa cells that stably express both firefly and *Renilla* luciferase (HeLa Dual-luc cells) were grown at 37 °C in Eagle medium supplemented with 10% fetal bovine serum (FBS), 100 units mL^{-1} penicillin, 100 $\mu\text{g mL}^{-1}$ streptomycin, 0.5 $\mu\text{g mL}^{-1}$ puromycin, and 500 $\mu\text{g mL}^{-1}$ zeocin (Invitrogen). Cells were passaged regularly to maintain exponential growth. Twenty-four hours prior to siRNA transfection, cells were seeded on opaque, white 96-well plates

(Costar) at a concentration of 15000 cells per well in antibiotic-free, phenol red-free DMEM (Invitrogen).

In vitro activity of siRNAs was determined using a high-throughput 96-well plate format assay for luciferase activity. HeLa Dual-luc cells were transfected with firefly luciferase targeting siRNAs at concentrations ranging from 0.40–30 nM using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. After 24 h, cells were analyzed for both firefly and *Renilla* luciferase expression using a plate luminometer (VICTOR², PerkinElmer) and the Dual-Glo Luciferase Assay kit (Promega). Firefly/*Renilla* luciferase expression ratios were used to determine the percentage of gene silencing relative to untreated controls.

In Vitro 5'-Phosphorylation. Oligonucleotides with (Table 1, guide strand of duplex II) and without (Table 1, guide strand of duplex I) the rF modification at the 5'-end were incubated with [³²P]ATP and T4 polynucleotide kinase using standard conditions and analyzed by denaturing 20% polyacrylamide gel electrophoresis.

RACE Analysis. HeLa-Dual Luc cells were transfected with 30 nM of the indicated siRNA at a density of 500 000 cells per well in a six-well plate using Lipofectamine 2000 transfection reagent (Invitrogen). Twenty-four hours post-transfection, cells were harvested, and total RNA was isolated using Qiagen's RNeasy kit (Qiagen). One microgram of total RNA was subjected to RACE assay for detection of 3' cleavage fragments (GeneRacer, Invitrogen). The GeneRacer RNA adaptor was ligated directly to untreated total RNA following the manufacturer's protocol. cDNA was generated using a gene-specific primer (cDNA GSP, 5'-TGACATCGACTGAAATCCCTGGT-3'). To detect cleavage products, PCR was performed using primers complementary to the RNA adaptor (GR5', 5'-CTCTAGAGCGACTGGAGCAGGACACTA-3', and Luc 1000 Rev1, 5'-GGAAACGAA-CACCACGGTAGGCT-3'). Amplification products were resolved by agarose gel electrophoresis and visualized by ethidium bromide staining. The identity of specific PCR products was confirmed by sequencing.

Crystallization and Structure Determination. Crystals of the rF-modified RNA dodecamer were grown at 18 °C by the hanging-drop vapor diffusion method, using the Nucleic Acid Miniscreen (Hampton Research) (29). Droplets containing 0.5 mM oligonucleotide, 5% 2'-methyl-2,4-pentanediol (MPD), 20 mM sodium cacodylate, pH 5.5, 10 mM cobalt hexamine, 20 mM LiCl, and 10 mM MgCl_2 were equilibrated against a reservoir of 35% MPD. Crystals appeared after 1 week. Crystals of the rF-modified RNA dodecamer grew readily but exhibited a different space group and packing than the duplex without the rF substitution. Diffraction data were collected at 120 K on the insertion device beamline (5-ID) of the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) at the Advanced Photon Source (APS), Argonne, IL. Data were processed with the program X-GEN (30), and the structure was determined by the Molecular Replacement method using the program EPMR (31). Refinement of the structure was carried out with the programs CNS and REFMAC (32, 33). The geometry of the RNA duplex was analyzed with the program CURVES (Version 5.3) (34). Selected crystal data and refinement parameters are listed in Supplementary Table 2 in Supporting Information.

Accession Codes. Structure factors and final coordinates have been deposited in the RCSB Protein Data Bank with ID code 2G92.

Acknowledgment: We thank C. Wilds, Z. Wawrzak, and G. Lavine for their valuable contributions. Support from the U.S. National Institutes of Health is gratefully acknowledged (GM55237 to M.E.). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under

Contract No. W-31-109-Eng-38. The DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center at the Advanced Photon Source (Sector 5) is supported by E. I. DuPont de Nemours & Co., The Dow Chemical Company, the National Science Foundation, and the State of Illinois.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Novina, C. D., and Sharp, P. A. (2004) The RNAi revolution, *Nature* **430**, 161–164.
2. Mello, C. C., and Conte, D. (2004) Revealing the world of RNA interference, *Nature* **431**, 338–342.
3. Hannon, G. J., and Rossi, J. J. (2004) Unlocking the potential of the human genome with RNA interference, *Nature* **431**, 371–378.
4. Tuschl, T., and Borkhardt, A. (2002) Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy, *Mol. Interventions* **2**, 158–167.
5. Manoharan, M. (2004) RNA interference and chemically modified small interfering RNAs, *Curr. Opin. Chem. Biol.* **8**, 570–579.
6. Dorsett, Y., and Tuschl, T. (2004) siRNAs: applications in functional genomics and potential as therapeutics, *Nat. Rev. Drug Discovery* **3**, 318–329.
7. Kurreck, J. (2003) Antisense technologies. Improvement through novel chemical modifications, *Eur. J. Biochem.* **270**, 1628–1644.
8. Moran, S., Ren, R. X., and Kool, E. T. (1997) A thymidine triphosphate shape analog lacking Watson–Crick pairing ability is replicated with high sequence selectivity, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10506–10511.
9. Guckian, K. M., Schweitzer, B. A., Ren, R. X., Sheils, C., Paris, P. L., Tahmassebi, D. C., and Kool, E. T. (1996) Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA, *J. Am. Chem. Soc.* **118**, 8182–8183.
10. Guckian, K. M., Krugh, T. R., and Kool, E. T. (1998) Solution structure of a DNA duplex containing a replicable difluorotoluene–adenine pair, *Nat. Struct. Biol.* **5**, 954–959.
11. Morales, J. C., and Kool, E. T. (1998) Efficient replication between non-hydrogen bonded nucleoside shape analogs, *Nat. Struct. Biol.* **5**, 950–954.
12. Delaney, J. C., Henderson, P. T., Helquist, S. A., Morales, J. C., Essigmann, J. M., and Kool, E. T. (2003) High-fidelity *in vivo* replication of DNA base shape mimics without Watson–Crick hydrogen bonds. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4469–4473.
13. Freier, S. M., and Altmann, K.-H. (1997) The ups and downs of nucleic acid duplex stability: structure–stability studies on chemically-modified DNA:RNA duplexes, *Nucleic Acids Res.* **25**, 4429–4443.
14. Parsch, J., and Engels, J. W. (2000) Synthesis of fluorobenzene and benzimidazole nucleic-acid analogues and their influence on stability of RNA duplexes, *Helv. Chim. Acta* **83**, 1791–1808.
15. Manoharan, M., Xia, J., and Rajeev, K. G. (2006) Single- and double-stranded oligonucleotides containing unnatural nucleobases for inhibition of gene expression and treatment of disease. U.S. Pat. Appl. Publ. 2006035254 A1 20060216 105 pp.
16. Rajeev, K. G., Xia, J., Noronha, A., Toudjarska, I., Li, F., Akinc, A., Frank-Kamenetsky, M., Braich, R., Egli, M., and Manoharan, M. (2006) RNA interference: RISC-mediated recognition of non-canonical ribo-difluorotolyl nucleotide: Synthesis, structure and activity. Abstracts of Papers, 231st ACS National Meeting, Atlanta, GA, United States, March 26–30, 2006, CARB-066.
17. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs, *Mol. Cell* **15**, 185–197.
18. Nykanen, A., Haley, B., and Zamore, P. D. (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway, *Cell* **107**, 309–321.
19. Schwarz, D. S., Hutvagner, G., Haley, B., and Zamore, P. D. (2002) Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. *Mol. Cell* **10**, 537–548.
20. Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001) Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate, *EMBO J.* **20**, 6877–6888.
21. Harborth, J., Elbashir, S. M., Vandeburgh, K., Manninga, H., Scaringe, S. A., Weber, K., and Tuschl, T. (2003) Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing, *Antisense Nucleic Acid Drug Dev.* **13**, 83–105.
22. Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA, *Science* **20**, 2053–2056.
23. Yekta, S., Shih, I. H., and Bartel, D. P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA, *Science* **304**, 594–596.
24. Soutschek, J.; Akinc, A.; Bramlage, B.; Charisse, K.; Constien, R.; Donoghue, M.; Elbashir, S.; Geick, A.; Hadwiger, P.; Harborth, J.; John, M.; Kesavan, V.; Lavine, G.; Pandey, R. K.; Racie, T.; Rajeev, K. G.; Roehl, I.; Toudjarska, I.; Wang, G.; Wuschko, S.; Bumcrot, D.; Kotliansky, V.; Limmer, S.; Manoharan, M.; Vornlocher, H.-P. (2004) Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs, *Nature* **432**, 173–178.
25. Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T., and Hunter, W. N. (1994) Crystal and molecular structure of r(CGCGAAUUAGCG): an RNA duplex containing two G(anti)A(anti) base pairs, *Structures* **2**, 483–494.
26. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—A visualization system for exploratory research and analysis, *J. Comput. Chem.* **25**, 1605–1612.
27. Zacharias, M., and Engels, J. W. (2004) Influence of a fluorobenzene nucleobase analogue on the conformational flexibility of RNA studied by molecular dynamics simulations, *Nucleic Acids Res.* **32**, 6304–6311.
28. Schweitzer, B. A., and Kool, E. T. (1994) Nonpolar Aromatic Nucleosides as Hydrophobic Isosteres of DNA Nucleosides, *J. Org. Chem.* **59**, 7238–7242.
29. Berger, I., Kang, C. H., Sinha, N., Wolters, M., and Rich, A. A. (1996) Highly Efficient 24-Condition Matrix for the Crystallization of Nucleic Acid Fragments, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **52**, 465–468.
30. Howard, A. J. (2000) Data processing in macromolecular crystallography, in *Crystallographic Computing 7. Proceedings from the Macromolecular Crystallographic Computing School* (Bourne, P. E., Watenpaugh, K. D., Eds.), pp 150–165, Oxford University Press, Oxford, U.K.
31. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 484–491.
32. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
33. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
34. Lavery, R., and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: conventions and principles, *J. Biomol. Struct. Dyn.* **6**, 655–667.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California-Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hirolyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

To Be or Not To Be an Author

Scientists love to publish papers. We editors enjoy helping them in this process. One question invariably arises when authors prepare a manuscript: Who should be included on the author list? Papers are the measure of the worth of a scientist. The number of papers published and in which journals they appear are both scrutinized at promotion time and during the grant review process. In short, much rides on publishing papers. Multidisciplinary science projects, such as those we see in chemical biology, often involve various research groups and many scientists collaborating on a common scientific goal. Who should get credit (and subsequent fame and possibly fortune) for the work? When should someone be included as an author, and when should they be acknowledged?

These are tough questions to answer, but we at *ACS Chemical Biology* would like to offer some advice. First and foremost, we recommend that our writers discuss authorship when collaborations begin. Material transfer agreements (MTAs) should be used when unique material is provided to another laboratory. The MTA is a legal contract that governs the transfer of one or more materials (usually biological, but increasingly chemical as well) from the owner (provider) to a recipient. Different forms (Forms and Model Agreements http://ott.od.nih.gov/forms_model_agreements/forms_model_agreements.html) are used depending on who is transferring material to whom (academic to industry, for profit to nonprofit, etc.), but the process of securing these agreements is the same. Before material is sent, typically the provider writes an MTA and sends it to the investigator requesting the material. The recipient then obtains the required signatures and returns it to the provider, who also signs it and then delivers the materials. The form is eventually passed on to the sponsored research or technology transfer office at the university (for example, see A Quick Guide to Material Transfer Agreements at UC Berkeley, <http://www.spo.berkeley.edu/guide/mtaquick.html>). While the MTA form is being exchanged, investigators should discuss and ideally put in writing the possible authorship should a manuscript be forthcoming from the use of the material transferred.

The one issue that is unique to chemistry is that synthetic compounds are nonrenewable reagents. Unlike plasmids and cell lines, some compounds cannot easily be replicated because their production can require complex and sometimes expensive synthetic processes. If the molecule is easily and quickly synthesized by a knowledgeable practitioner in the field, the transfer may or may not warrant coauthorship versus an acknowledgment. The distinction is up to the provider, who must determine what level of effort was required. Again, it is best to resolve these issues before the paper is written.

Once a paper has been completed, spelling out what each potential author contributed to the paper may be helpful. Some journals, such as the *Journal of the American Medical Association* (JAMA, Editorial Policies for Authors, http://jama.ama-assn.org/ifora_current.dtl), require this, and we at *ACS Chemical Biology* support the concept underlying this request. *ACS Chemical Biology* is also considering asking corresponding authors to obtain letters from scientists listed in the acknowledgment section of a paper indicating that they agree that they should be thanked and not given authorship. Let us know if you believe requesting this information will help protect authors and collaborators.

The final order of authorship is always up to the investigators. Typically, those who conducted most of the experiments appear first in the list of authors, with principal investi-

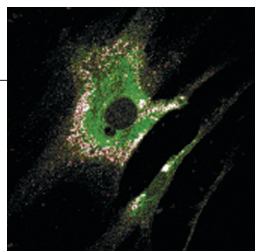
gators listed last. At *ACS Chemical Biology*, we do allow two authors to be designated as contributing equally to the work. Furthermore, we allow changes in the author list anytime before we accept the paper, as long as the corresponding author provides a reason why the author was added or removed. For example, the request from a reviewer to include additional data may require the work from a new coauthor, and in these cases, it is appropriate to include this person in the author list. The order of authors may need to be revised again, but it should be discussed before the revised paper is submitted to the journal—it is practically impossible to change the author list after it is published on our website.

These suggestions may seem very formal to some, but they are the simplest way to avoid disagreements after the paper is published. If someone disputes authorship after a paper is printed, what do we at the journal do? We will request the list of correspondence between the disputing parties. This may include MTAs as well as e-mail correspondence spelling out what material was transferred and the formal or informal agreements made at the time of transfer. On the basis of this paper trail, we will provide a recommendation to the corresponding author of the paper in question. This author will make the final decision about whether she/he should publish an erratum revising the author list. Note, however, that we cannot change the published pdf file nor can we change the author listing in PubMed or similar abstracting services. The only way to alert someone to the authorship change is through an erratum published in our journal that will be linked to the original paper. This is why we strongly recommend that authorship issues be discussed before the paper is submitted to the journal.

We hope this editorial provides some useful information to potential authors about what we at *ACS Chemical Biology* expect with respect to authorship. In future editorials, we will continue to discuss our journal policies and relevant ethical issues. We look forward to your feedback.



Evelyn Jabri
Executive Editor



Chaperoning the ER

Mutations that impair proper protein folding can cause a variety of pathological conditions including Gaucher disease, a lysosomal storage disorder resulting from a functional deficiency of glucocerebrosidase (GC). Several distinct point mutations in GC, such as N370S, G202R, and L444P, target this enzyme for degradation in the endoplasmic reticulum (ER) and lead to reduced concentration of GC in the lysosome. Sawkar *et al.* (p 235) demonstrate that the lysosomal concentration of these GC mutants can be increased by temperature reduction or through the use of chemical chaperones.

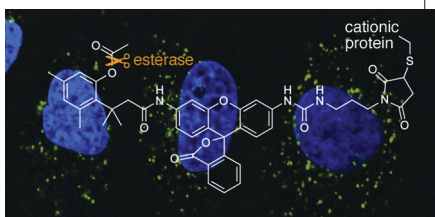
The GC mutants suffer from destabilization at neutral pH, leading to their degradation in the ER before relocation to the lysosome. However, these mutants are relatively stable at the low pH of the

lysosome, suggesting that if the folding predicament could be resolved and the mutant proteins could be trafficked to the lysosome, their activity might be partially retained. The researchers demonstrate that there is increased activity and lysosomal distribution of the N370S and G202R mutants in fibroblasts grown at 30 °C. Likewise, when chemical chaperones, or small molecules that can assist protein folding in the ER, were added to these cells, similar increases in N370S and G202R enzymatic activity and lysosomal localization were observed. These results demonstrate that the protein folding environment of the ER can be manipulated by small molecules and/or temperature regulation, offering potential new strategies for developing treatments for Gaucher disease and related protein folding disorders.

Promoting the Pro-Fluorophore

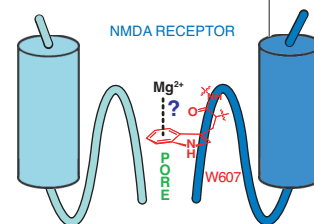
Fluorescent molecules are powerful tools for exploration of biological processes. The desire to understand these processes at increasingly intricate levels of detail has inspired innovative modifications to fluorescent compounds that expand their utility as molecular probes. In this spirit, Lavis *et al.* (p 252) have generated versatile, rhodamine-derived fluorogenic labels, termed “pro-fluorophores”, that have controllable fluorescence properties and a bioconjugation handle. The pro-fluorophores are modified at one of the rhodamine nitrogens with a “trimethyl lock”, or an enzyme-reactive chemical modification that masks the fluorescence of the molecule, and at the other with a urea moiety that enables placement of a functional group for conjugation to biomolecules of interest. Synthesis and characterization of several model pro-fluorophores facilitated the creation of a compound that was stable in aqueous media, was highly reactive to porcine liver esterase, and, when added to HeLa cells, diffused into the cytoplasm where it was enzymatically hydrolyzed to yield a fluorescent compound.

In a final derivatization, a maleimide group was installed in conjunction with the urea functionality to enable conjugation of the pro-fluorophore to a thiol. The compound was conjugated to an RNase A derivative, and the high stability and low background fluorescence of the fluorogenic label allowed for time-lapse imaging of endocytosis of the conjugate by HeLa cells (the authors have provided a movie for online viewing). The chemistry behind this versatile new molecular tool allows for variations at the enzyme reactive and bioconjugation sites, enabling the design of molecular probes tailored to desired areas of interest.



No Nonsense Technology

The NMDA receptor is a member of the ligand-gated ion channels that responds to the excitatory neurotransmitter glutamate. NMDA receptors are structurally and chemically complex proteins that contain three binding sites, one of which is occupied by a Mg^{2+} ion that blocks the channel pore. It has been proposed that the Mg^{2+} engages in a cation- π interaction with a conserved tryptophan residue in the binding site. Using unnatural amino acid mutagenesis, McMenimen *et al.* (p 227) demonstrate that this is not the case; rather, the tryptophan acts as a structural element in the channel blocking role of the Mg^{2+} .



Nonsense suppression technology was used to replace the tryptophan residue with several unnatural residues in order to explore the relationship between the tryptophan and the Mg^{2+} ion. Fluorinated tryptophans and 2-naphthylalanine have smaller cation- π interaction abilities than tryptophan, yet the IC_{50} values and the voltage dependence of the channel block were not significantly altered with these amino acids. Moreover, comparison of tyrosine with the non-aromatic cyclohexylalanine also resulted in an essentially identical Mg^{2+} block. Collectively, these data indicate that it is the size, shape, and hydrophobicity of the tryptophan that aptly shapes the channel for effective Mg^{2+} blockade. This study is the first application of nonsense suppression technology to the NMDA receptor and promises to lead the way for new revelations into the structure and function of this fascinating receptor.

Published online May 19, 2006

10.1021/cb600192q CCC: \$33.50

© 2006 American Chemical Society

Spotlight

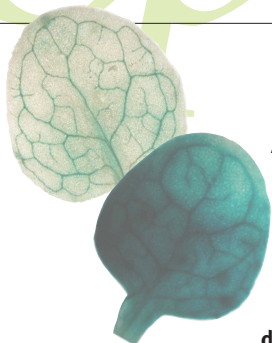


Image reprinted with permission from AAAS

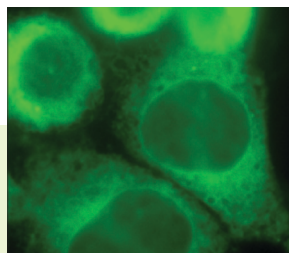
Arabidopsis Resists Hormone Stimulation

Plants and animals have mechanisms to detect the presence of a pathogen. In plants, a 22 amino acid peptide derived from the N-terminus of flagellin (flg22) triggers a rapid down regulation of a subset of genes, in part through a posttranscriptional mechanism. One such mechanism is RNA silencing, an mRNA degradation process mediated by short 20–24 nucleotide short interfering (siRNA) and microRNA (miRNA). These short RNAs bind to target mRNAs and mediate their cleavage. Until recently, it was not clear if miRNAs, which modulate developmental processes, were also involved in antibacterial resistance. Now, Navarro *et al.* (*Science* 2006, 312, 436–439) show that the flg22 induces the generation of a miRNA that targets mRNAs of receptors for auxin, a plant hormone, resulting in reduced susceptibility to bacterial infection.

To investigate the mechanism by which flg22 functions in bacterial resistance, transgenic

Arabidopsis plants expressing viral genes that suppress miRNA- and siRNA-mediated processes were treated with flg22. Analysis of the mRNA levels from treated and untreated plant cells showed that a subset of mRNAs TIR1, AFB2 and AFB3 (but not AFB1) were more prevalent in these transgenic plants. These F-box proteins are auxin receptors. The repression of TIR1 and the AFB proteins coincides with increases in the level of miR393, a conserved miRNA. This miRNA-mediated repression leads to the down-regulation of auxin signaling pathways implicated in disease susceptibility. The miR393 pathway works in parallel with transcriptional repression of auxin receptors to ensure a rapid and robust immune response to the attaching bacteria. Precisely how auxin promotes disease susceptibility is not clear.

These data show for the first time that miRNAs also contribute to antibacterial resistance in plants. It will be interesting to determine if other stress-induced processes also utilize miRNAs to control cellular responses. EJ

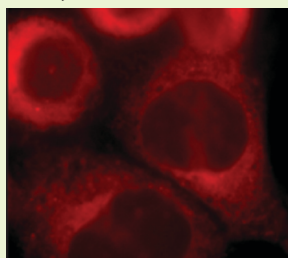


Stressed Out Cells

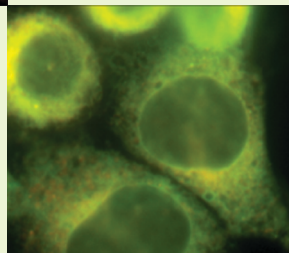
Humans aren't the only ones who get "stressed out". Cells get stressed out under a variety of circumstances, such as glucose deprivation, altered glycosylation, or accumulation of unfolded proteins.

Fortunately, the cell has intricate systems in place to deal with such stress, including the unfolded protein response (UPR). The UPR initiates three signaling pathways, one of which is mediated by the endoplasmic reticulum (ER) transmembrane protein kinase and endoribonuclease inositol-requiring enzyme 1 α (IRE1 α). Members of the BCL-2 protein family, known for their intimate involvement in apoptosis, are also found in the ER, and various agents that induce apoptosis also induce the UPR. To further define the role of BCL-2 proteins in the cellular stress response, Hetz *et al.* (*Science* 2006, 312, 572–576) have examined the involvement of the proapoptotic BCL-2 proteins BAK and BAX in UPR signaling events.

The researchers created BAX and BAK double knockout (DKO) mice and explored the effects on UPR signaling pathways.



Reprinted with permission from AAAS



Comparison of DKO and wild-type cells using a variety of experiments including protein expression and phosphorylation analysis, coimmunoprecipitation, protein mutagenesis, and small interfering RNA revealed that BAX and BAK modulate the IRE1 α signaling pathway by affecting X-box-binding protein 1 (XBP-1) expression, a transcriptional activator of UPR-related genes that is activated by IRE1 α . Furthermore, it was discovered that BAK expression at the ER membrane augmented IRE1 α signaling, BAK and BAX interacted directly with IRE1 α through their BH3 and BH1 domains, and this interaction was enhanced in cells undergoing ER stress suggesting that BAX and BAK may stabilize the active form of IRE1 α .

The authors propose that BCL-2 proteins may function not only as proapoptotic molecules but also as modulators of ER homeostasis that link stress signals to the apoptotic circuitry in cells. This study provides a tangible connection between the UPR and proteins involved in apoptosis and will contribute to further deciphering of these two fundamental cellular processes. EG

Toxicity is in the CARDS

Mycoplasma pneumoniae is responsible for community-acquired pneumonia and various other airway disorders. This unique species of bacterium is characterized by the lack of a cellular wall, making them resistant to antibiotics that disrupt cell wall synthesis such as penicillins. In addition, like viruses, they are dependent on host cells for virulence. These unusual characteristics have posed unique challenges in the investigation and treatment of *M. pneumoniae* infections, and many of the molecular mechanisms involved in their pathology remain a mystery. Now, Kannan and Baseman (*PNAS* 2006, 103, 6724–6729) report the identification of a virulence factor seemingly responsible for respiratory injury associated with *M. pneumoniae*.

In the search for agents responsible for *M. pneumoniae* virulence, affinity chromatography using a human lung protein led to the identification of a 68 kDa protein designated community-acquired respiratory distress syndrome toxin (CARDS TX). Analysis of the primary sequence revealed that CARDS TX shares sequence similarities to pertussis toxin (PTX), hinting that, like PTX, CARDS TX may be an ADP-ribosylating toxin (ADPRT). Indeed, experiments

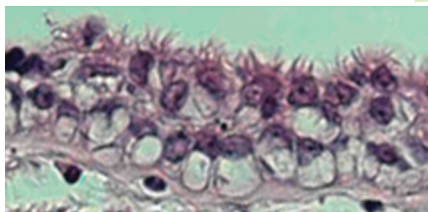


Image reprinted with permission from PNAS

demonstrated that CARDS TX possesses ADP-ribosyltransferase (ART) activity, modifying an overlapping but distinct set of proteins than that of PTX. Treatment of mammalian cell cultures and baboon tracheal rings with the toxin revealed that CARDS TX causes

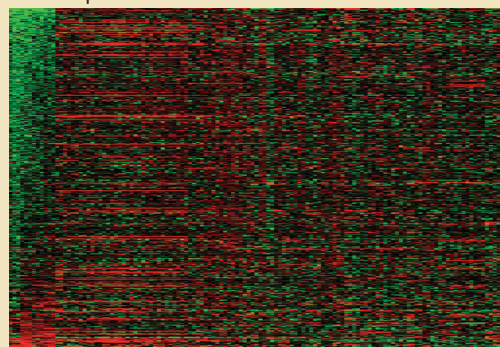
the characteristic cytoplasmic vacuolization observed in infected specimens. Notably, infected individuals seroconvert to CARDS TX, providing further evidence that CARDS TX is the virulent factor in *M. pneumoniae*. The authors propose that intimate contact between the mycoplasma and the host cell in the early stages of infection could enable release of CARDS TX into target cells, leading to ADP ribosylation, vacuolization, and eventual cytotoxicity. Identification of this mycoplasma-associated toxin provides an explanation for the pathology of *M. pneumoniae* and will facilitate diagnostic, preventative, and treatment strategies for *M. pneumoniae* infections. **EG**

(Poly)Combing through the Stem Cell Genome

The extraordinary ability of stem cells to differentiate into most cell types has stimulated much research spanning the basic biology behind their metamorphosis to their tantalizing clinical potential in regenerative medicine. Recently, intriguing connections have been made between stem cell regulation and chromatin structure. Now, four studies tackle genome-wide analyses of chromatin structure to provide insight into the role of chromatin in stem cell regulation.

Chromatin, the structural building block of chromosomes, is made up largely of chromosomal DNA and proteins called histones. Posttranslational modification of histones, such as methylation or acetylation of specific residues, modulates chromatin structure, and these modifications are part of the gene regulation machinery. The polycomb group (PcG) of proteins are chromatin-binding proteins

and are integral pieces of this machinery. PcG proteins form two major complexes, termed polycomb repressive complexes (PRCs). PRC2 binds to sites of transcriptional repression and modifies chromatin



Reprinted with permission from Cell Press

structure through epigenetic changes, or “heritable” changes that do not involve alteration of DNA sequence, by catalyzing histone H3 lysine-27 (H3K27) methylation. Trimethylation of H3K27

(H3K27me3) in turn provides a landscape that recruits PRC1, which facilitates chromatin reorganization, including oligomerization and condensation, reinforcing transcriptional repression. Investigation into the identity and function of the specific genes that PRCs target will help elucidate mechanisms of gene regulation and define the pathways involved in embryonic development.

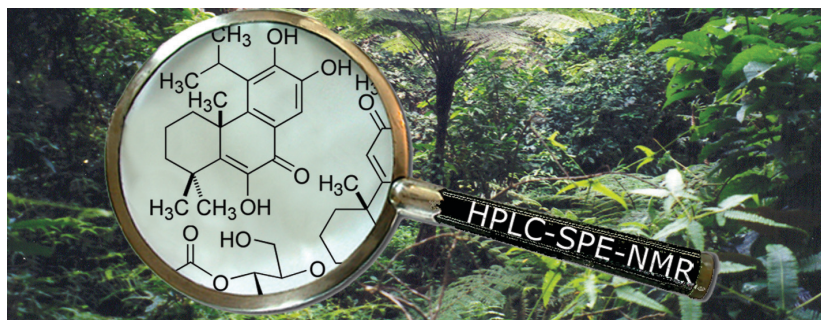
To this end, Tolhuis *et al.* (*Nature Genetics*, published online April 20, 2006, doi:10.1038/ng1792), Boyer *et al.* (*Nature*, published online April 19, 2006, doi:10.1038/ng04733), and Lee *et al.* (*Cell* 2006, 125, 301–313) have mapped binding patterns of PcG proteins in *Drosophila melanogaster*, mouse, and

(continued on page 190)

Extracting Extract Structures

Natural products are an incredibly valuable resource for discovery of potential new medicines and molecular tools for biological exploration. One major hurdle in natural products research is the tedious isolation and structure determination processes that are often required. Clarkson *et al.* (*J. Nat. Prod.* 2006, 69, 527–530) have developed a new technique, high-performance liquid chromatography–solid-phase extraction–nuclear magnetic resonance (HPLC–SPE–NMR) that enables the rapid structure determination of constituents of natural product extracts.

The researchers used HPLC–SPE–NMR to analyze the petroleum ether root extract from *Harpagophytum procumbens*, a plant native to South Africa whose large roots are used as an anti-inflammatory agent and to stimulate digestion. From 11 major and minor HPLC peaks analyzed, two novel structures were identified. Structure determination was facilitated by several key features of the HPLC–SPE–NMR technique. First, extensive 2D NMR data can be obtained from all HPLC peaks without interruption of the mobile phase flow, as in direct, stopped-flow HPLC–NMR. In addition, multiple SPE trappings can be conducted, dramatically improving signal-to-noise ratios



in the NMR analysis. Finally, the entire analysis can be performed under inert conditions, facilitating structure elucidation of air- and moisture-sensitive compounds. The novel compounds were found to contain an uncommon diterpene skeleton, named chinane, with an unusually placed isopropyl group on the terpene ring system. On the basis of comparison of their NMR data with that of other diterpenes, the authors propose that the chinane structure may exist in other known compounds whose structures have been misinterpreted. This new technique could revolutionize the structure elucidation process in natural products research and revitalize industrial drug discovery programs based on natural products by providing accelerated access to chemical diversity of biological sources. EG

(Poly)Combing through the Stem Cell Genome, *continued*

human embryonic stem (ES) cells, respectively, and Bernstein *et al.* (*Cell* 2006, 125, 315–326) have mapped histone methylation patterns in mouse ES cells. In all three species, PcG proteins strategically bind to sites of transcriptional repression of developmentally important genes, including genes that regulate early and late developmental processes in *D. melanogaster*, such as ectoderm development and eye morphogenesis, and genes involved in mammalian cell differentiation, cell-fate commitment, organogenesis, hematopoiesis, and neurogenesis. In addition, there is a strong correlation between PcG binding and sites of H3K27me₃ that substantially decreases upon cell differentiation. This cross-species, genome-wide data demonstrates that polycomb complexes target developmental genes and are a critical component of the gene silencing

machinery. This implicates chromatin structure as an important regulator in ES cell pluripotency.

Remarkably, a unique characteristic of ES cell chromatin appears to allow for the reversal of gene silencing upon appropriate developmental signals. While H3K27 methylation is a transcriptional repression signal, methylation of H3 lysine 4 (H3K4) signifies transcriptional activation. Upon analysis of histone methylation patterns across the mouse ES cell genome, Bernstein *et al.* identified a specific pattern in the vicinity of many developmental genes, termed “bivalent domains” that consists of large regions of H3K27 methylation harboring smaller regions of H3K4 methylation. The implicated genes are largely silenced in ES cells, consistent with the H3K27 mark and with the findings of the other studies. The authors propose that the coincident H3K4

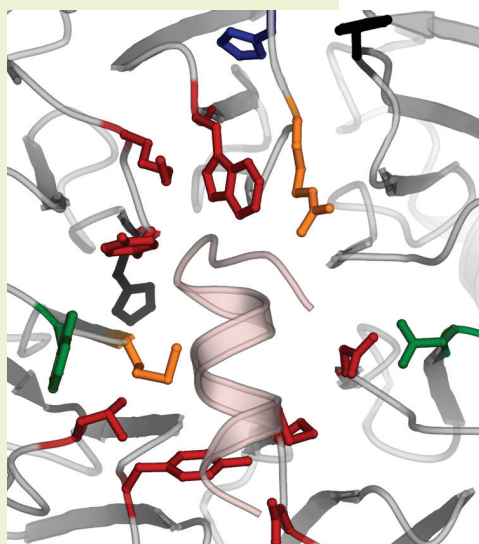
methylation keeps these critical genes poised for activation at later developmental stages. Interestingly, strong correlations between genome sequence and histone methylation were also observed, highlighting the importance of DNA sequence in defining the initial epigenetic state, which is then altered as embryonic development progresses.

Taken together, these studies point to a model of ES cell regulation whereby PcG proteins and histone methylation function to control transcription of developmentally relevant genes, silencing them until developmental cues call for their activation. Further elucidation of the molecular details of this intricate regulation system will enhance our understanding of myriad developmental processes, including stem cell differentiation, embryonic development, tissue homeostasis, aging, and oncogenesis. EG

G Protein Hot Spots

Heterotrimeric guanine nucleotide-binding proteins (G proteins) mediate diverse physiological processes, and small molecule modulators of these proteins have broad therapeutic potential. G protein coupled receptor activation results in the release of G protein $\beta\gamma$ subunits that then participate in interactions with multiple downstream effector molecules. Many of these effector interactions occur at a hotspot for protein-protein interactions on the β -subunit surface. Bonacci *et al.* (*Science*, 2006, 312, 443–446) use molecular modeling to discover small molecules that interact with this surface and demonstrate that these compounds can differentially affect G protein function.

Virtual docking of a structurally diverse small molecule library to the interaction hotspot led to the identification of several molecules, including M119 and M201, that bound to distinct subsurfaces of the hotspot. While M119 inhibited G $\beta\gamma$ -dependent phospholipase C $\beta 3$ (PLC $\beta 3$) and phosphoinositide 3-kinase activation, M201 potentiated these interactions. In addition, cellular assays demonstrated that M119 attenuated peptide-induced, G protein-mediated calcium increases while M201 had no effect. Moreover, *in vivo* studies showed that M119 increased the sensitivity of mice to morphine, an expected effect from a compound that blocks PLC $\beta 3$ activity. Given the diversity of interactions in which G proteins participate, the ability to differentially manipulate G protein function with small molecules is a powerful strategy. EG



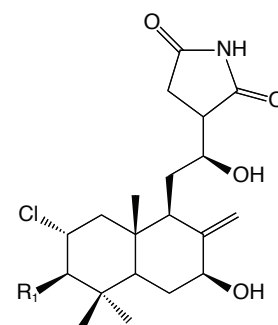
Reprinted with permission from *Biochemistry*, 2005, 44, 10593–10604.

Elongate No More

Some marine mollusks secrete small molecules that inhibit biological processes such as protein translation. Mechanistic studies of translation have benefited significantly from the use of these ligands, many of which target specific steps in protein synthesis. Furthermore, several inhibitors of translation have been used as anticancer drugs and are in clinical trials. Recent studies showed that chlorinated lissoclimides from the marine mollusk *Pleurobranchus forskalii* affect translation. Now, Robert *et al.* (*RNA* 2006, 12, 717–724) find that two members of this family of molecules, chlorolissoclimide and dichlorolissoclimide, are potent inhibitors of eukaryotic translation elongation.

Chlorolissoclimide and dichlorolissoclimide are toxic bicyclic diterpene alkaloids and are members of a larger family of diterpenoids called labdanes. Investigation into the mechanism of inhibition of these compounds revealed that they inhibit the elongation phase of translation, as opposed to preventing translation initiation. Furthermore, unlike other translation elongation inhibitors such as phyllanthoside and nagilactone C, which disrupt polyribosomes, chlorolissoclimide and dichlorolissoclimide block translation elongation by stalling the ribosome on the mRNA. The authors note that there is a structural similarity between the lissoclimides and cycloheximide, a translation inhibitor that also inhibits translocation by interfering with the release of ribosomes from polysomes.

Diterpenoids from the labdane class have been isolated from a variety of plants and animals, and they have been found to possess inhibitory activity against a wide range of eukaryotic organisms, including *Trypanosoma*



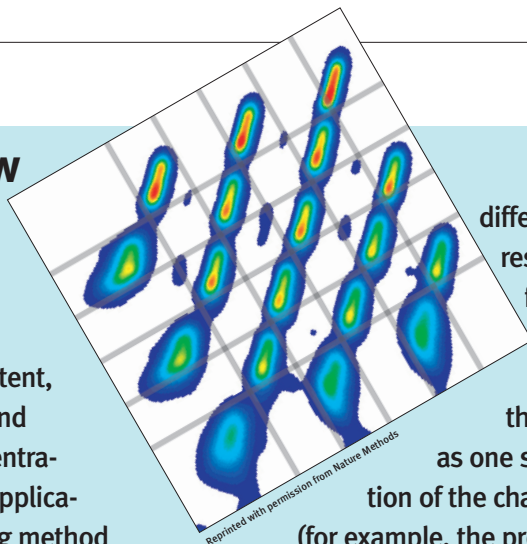
Reprinted with permission from the RNA Society

cruxi, algae, and cancer cells. The authors propose that, although more research is required into the mechanisms of action of these compounds, diverse species may use inhibition of protein synthesis as a defense mechanism against predators. EG

Go with the Flow

Flow cytometry is a powerful tool for analysis of many cellular properties, including cell surface markers, DNA content, calcium flux, apoptosis, and intracellular protein concentrations. The wide range of applications make this a tempting method for large-scale investigations such as drug screening and cellular profiling, but increasing the capacity of the technology has been hampered by corresponding increases in reagent expense, inadequate sample throughput, and variability in sample labeling. Now, Krutzik and Nolan (*Nature Methods* 2006, 3, 361–368) present a cell-based multiplexing approach, termed fluorescent cell barcoding (FCB), that overcomes these limitations and enables a wide range of high-throughput flow cytometry applications.

Flow cytometry permits multiparameter analysis in complex cell populations. For example, subpopulations of white blood cells in a heterogeneous sample can each be analyzed for the presence of multiple cell surface markers. FCB takes this multiparameter capacity to an even higher level by exploiting the ability of flow cytometers to discriminate between samples labeled with different intensities of the same fluorophore. FCB combines the use of multiple fluorophores at multiple intensities per fluorophore so that each sample acquires a unique fluorescent signature, or barcode. For example, use of two fluorophores at six



different intensities each results in 36 unique fluorescent barcodes. The uniquely bar-coded samples can then be run together as one sample in the evaluation of the characteristics of interest

(for example, the presence of specific phosphorylated proteins), resulting in significant reductions in reagent cost, sample variability, and time on the instrument. Deconvolution of the data after the samples are run enables analysis of the original samples based on their FCB signature.

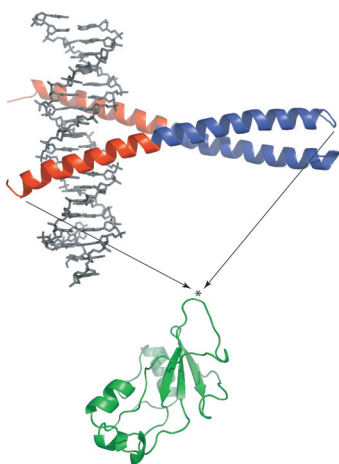
The feasibility of using the FCB platform in high-throughput applications was demonstrated with two large-scale assays. First, a drug screening assay was performed in which three FCB markers were used to barcode 96 samples (analogous to use of a 96-well plate), and 70 small molecule inhibitors were screened for inhibition of cytokine-induced protein phosphorylation events. Data analysis rapidly revealed compound selectivity for specific signaling pathways. Second, a cellular profiling assay was run where a heterogeneous population of mouse splenocytes was analyzed for differential response to cytokine stimulation, and cell-type specific effects were promptly discerned. Technical advancements in this powerful technology, including innovative fluorescent markers and more sophisticated data analysis methods, will further contribute to its utility in high-throughput applications. EG

Protein Tug-of-War

The design of proteins whose functions can be manipulated under specific conditions is an innovative strategy for the development of macromolecules with novel sensor capabilities. Ha *et al.* (*J. Mol. Biol.*, 2006, 357, 1058–1062) have created a fusion protein that can be induced either to be catalytically active or to bind DNA, depending on the conditions.

The fusion protein, termed BG, consists of a target protein, the catalytically active bacterial ribonuclease barnase (Bn), and an 'inserted' protein, the DNA binding domain of GCN4. The GCN4 sequence is strategically integrated at a surface-exposed loop in Bn. In the absence of the GCN4 ligand, a DNA oligonucleotide called AP-1, the Bn domain of the chimera is more stable than the GCN4 sequence. Bn is folded and catalytically active, while GCN4 is largely disordered. In contrast, in the presence of AP-1, the GCN4 portion is more stable than the Bn region and a thermodynamic 'tug-of-war' ensues, wherein folding of the ligand binding domain of GCN4 effectively splits Bn in two and abolishes enzymatic activity.

BG is a model for the design of additional chimeric proteins that possess novel sensor capabilities. For example, a potential therapeutic agent could be designed by inserting a different binding domain into Bn. The enzymatic activity of the new chimera would be controlled by binding of a specific ligand, which may be present only in certain cell types. EG



Reprinted with permission from the *Journal of Molecular Biology*

Illuminating the Drug Delivery Pathway

As if creating effective inhibitors for potential drug targets isn't hard enough, appropriate formulation and bioavailability are two additional major hurdles in crossing the drug discovery finish line. One method over these hurdles is conjugation of the potential drug to a molecular transporter that yields a water-soluble yet membrane-permeable entity. Once inside the cell, however, the compound likely needs to be released from the transporter so that it can perform its intended biological activity. Jones *et al.* (*JACS*, published online April 25, 2006, doi:10.1021/ja0586283) have created a controllable releasable linker system that allows for liberation of the molecule after cell entry, and they have developed a cellular assay that enables measurement of conjugate uptake and release of the molecule.

The conjugate design incorporates specific features that make it an attractive drug delivery system. First, the linker that connects the potential drug, or cargo, to the transporter (a D-octarginine

molecule that has previously been shown to effectively transport small molecules into cells and tissue) contains two key functional groups, a carbonate and a disulfide. Upon entry of the conjugate into the reducing environment of the cell's cytoplasm, the disulfide bond is cleaved. The exposed free thiol reacts with the carbonate, resulting in liberation of the cargo. Furthermore, conjugate stability can be increased from hours to days simply by increasing the linker length, providing a tunable system. To test the effectiveness of the conjugate, luciferin was employed as the cargo, and luciferase-transfected cells were treated with the conjugate. Conjugate exposure resulted in an initial increase in luminescence followed by a gradual decay over the course of several minutes, and luminescence was demonstrated to be dependent on intracellular release of luciferin. This releasable luciferin conjugate enables the exploration of other transporters and linkers for evaluation of delivery, release, and target interaction of potential drugs. EG

UPCOMING CONFERENCES

Nucleic Acids, GRC

June 4–9, 2006
Newport, RI

Nucleic Acid Enzyme, FASEB

June 10–15, 2006
Saxton River, VT

Single Molecule Approaches to Biology, GRC

June 18–23, 2006
New London, NH

Biological Methylation, FASEB

June 24–29, 2006
Saxton River, VT

Ubiquitin & Cellular Regulation, FASEB

July 22–27, 2006
Saxton River, VT

Natural Products, GRC

July 23–28, 2006
Tilton, NH

Technology Transfer: From Society to the Lab and Classroom

David P. Martinsen*

American Chemical Society, 1155 16th Street, NW, Washington, D.C. 20036

Sunday morning at 8 am might not seem like the best bet to find an interesting session at an American Chemical Society (ACS) meeting, but at the Chemical Information (CINF) Division program in Atlanta, this turned out to be the case. Every year or so, the CINFers present what amounts to a technology review, and this one, organized by Andrea Twiss-Brooks of the University of Chicago and Erja Kajosalu from MIT, was exciting on a number of counts. There were a total of eight speakers in the symposium on Social Software and Chemical Information, covering topics from social software, Web 2.0, classroom/educational applications, open access, Wikipedia, blogs, and webinars. It would be impossible to highlight all of the topics covered, so I will only attempt to give a few highlights. If you wish, you can skip my comments altogether, and jump to the last paragraph for the *coup de grâce*.

The most interesting impression to me is that so much is happening; it is difficult to keep abreast of everything. The other impression I got is that while we usually think of technology being transferred from the laboratory to society, this technology is doing just the opposite, moving from society-at-large to the scientific and professional worlds.

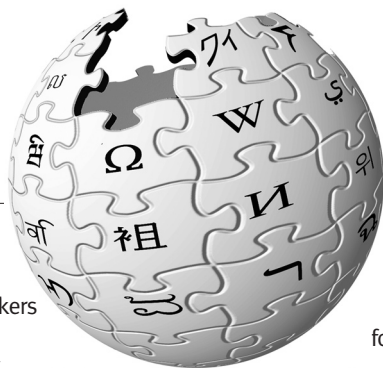
The first talk presented a laundry list of new technologies that together comprise what is coming to be known as social software. The speaker, Beth Thomsett-Scott from the University of North Texas, polled the audience to determine how many were familiar with each technology

she mentioned. Nearly all had heard of the initial topics she discussed, such as RSS, Wiki, and Flickr. However, by her last slides, only two or three raised their hands. Things like jybe (join your browser with everyone), furl (frame uniform resource locator), and clicker (a wireless device that could be used for students to respond to questions from a teacher) had not yet reached the attention of most in the audience. These are interesting technologies to examine, but the question with all of these tools is how many will actually make it into the mainstream.

Teri Vogel of UCSD gave an overview of RSS as a unifying example of Web 2.0 (*If you don't yet know what Web 2.0 is, don't worry, it's not that well defined. Check Wikipedia for the current definition.*) Vogel quotes a figure of 5–10% of Internet users currently utilizing RSS. If you are not one of those, this talk included a short tutorial on how to use RSS. Please see the last paragraph for more information. Vogel also included many examples of web sites, including libraries, publishers, and government organizations that are using RSS. She stated a couple of times that one could use RSS to poll many different sites and then read the feeds at your leisure. One wonders if, after keeping track of new emerging technologies, learning how to use new technologies, and then connecting to a host of RSS feeds, there will ever be leisure time to read the content of the feeds (*Note: if you don't know what RSS is, there seems to be an ever-growing list of meanings for the acronym. Perhaps an RSS feed of new RSS meanings of RSS would help.*)

*To whom correspondence may be addressed.
E-mail: d_martinsen@acs.org.

Published online May 19, 2006
10.1021/cb600195z CCC: \$33.50
©2006 American Chemical Society



WIKIPEDIA *The Free Encyclopedia*

© Wikimedia – Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc. (<http://www.wikipedia.org>)

Two of the other speakers discussed technology to make audio recording of classes and/or meetings for later use by the students or others who might have missed the class or presentation or might wish to review some of the material covered. Brian Lynch, of St. Francis Xavier University, pointed out that in Australia it is standard practice to record every lecture and post them on the Internet within an hour. Jeremy Garritano described the application of audio recording technology into classrooms in Purdue. While the technology is fairly straightforward, difficulties were encountered in scalability, location of microphones to capture the lecturer (but not background noise), and delivery of the audio files in an easily usable fashion.

Blogging was the subject of two talks as well. Barbara Greenman of University of Colorado Boulder and Randy Reichardt of the University of Alberta, talked about how blogging is changing the way classes are run, changing the way students collaborate, and could potentially change the scientific publication process as well.

Wikipedia has emerged as a common tool that many use regularly for at least an initial attempt to find information about a new subject. In fact, several of the morning's speakers used Wikipedia as a reference to define the new terms about which they were speaking. Concerns over this resource as a reliable source for information were captured in the title of the talk by Martin Walker of SUNY Potsdam, "Wikipedia: Social revolution or information disaster?" As might be expected from one of the 29 editors of the Wikipedia chemistry project, Walker's answer was on the side of the revolution. He acknowledged the potential for problems. For example, anybody can edit a Wikipedia entry and entries don't undergo a formal peer review. He even gave examples that have been seen on the

chemistry portal: an entry for barium chloride had invalid data on solubility, reactivity, and toxicity for six weeks; an entry was made for a fictitious molecule. However,

these instances are rare. Some types of vandalism can be detected automatically by Wikipedia. With alerting tools in place, the portal editors are quickly notified of any changes, legitimate changes, changes introduced by well-intentioned but uninformed individuals as well as those introduced by malicious miscreants. This allows errors to be discovered relatively quickly and repaired. Quoting from a *Nature* study (<http://dx.doi.org/10.1038/438900a>), Walker concluded that Wikipedia is right most of the time and represents a real success of the Open Access movement. He noted that none of the other speakers referenced *Encyclopedia Britannica* for their definitions, since they aren't available there. It will be interesting to see how the level of effort in the growth and maintenance of Wikipedia will continue over time. For now, the result is impressive.

The last talk in the morning session was presented via WebEx. The speaker, Jonathan Coffman from Wyeth Pharma, did not make the trip to Atlanta but presented his talk via speakerphone and Internet. The slides were controlled by Coffman in New Hampshire, viewed on a laptop in the meeting room, and projected onto the screen. The speakerphone, situated next to the microphone, provided the audio. Coffman's topic was how the ACS Biotechnology Secretariat has used WebEx technology to hold a number of remote symposia, and he discussed the potential benefits of the technology if it were to be more widely adopted for use in ACS National and Regional Meetings. BIOT's use of remote symposia was focused mainly on how to maintain membership in the BIOT Secretariat, as well as to reach out

to a larger number of scientists than were able to attend the symposia in person. The talk went off without a hitch, and there was not really any hindrance to the exchange of information by Coffman's presence virtually. Even the question and answer part of the talk was not too much different from normal. I don't know how much of the questions from the audience Coffman could hear, but as is usual, the session chair repeated the question, and she was seated next to the speaker phone.

I will briefly mention the afternoon session on scholarly publishing, since besides my involvement as a co-organizer, there is a tie-in with the theme of the morning session. George Whitesides opened the session on The Nuts and Bolts of Scholarly Publishing with a very nuts and bolts talk on authoring a paper. With over 900 publications to his credit, he is certainly qualified to address the subject. He advocated a very methodical approach to writing an article for publication, encouraging those present to consider the process of writing to be an integral part of the research process itself. Whitesides simply stated that if you don't publish your research, what is the point of having done it?

The following talks dealt with the peer review process, ethical questions, and the process of turning the author's material into published articles. These may be thought of as fairly mundane topics, but given the recent examples of fraudulent publications, they should, perhaps, be given greater consideration. As was emphasized in each of the talks, the sheer volume of manuscripts submitted for publication place ever increasing burdens on the entire system. If the increasing pressure for publishing, along with advanced tools capable of manipulating graphics and generating high volumes of reasonable looking data, results in an increase in the number of scientists willing to cut corners, the entire enterprise is at risk. Journal editors may one day use technology to help them spot the most

blatant forms of fraud. These talks, as well as similar presentations at other venues, indicate that journal editors and publishers are starting to look for ways to get the word out that ethics is something to be considered in training scientists for their careers.

The last talk in the afternoon was given by *ACS Chemical Biology's* (ACS CB) own Sarah Tegen and Evelyn Jabri. Their topic was the way in which the World Wide Web is being exploited in the publishing world, with examples from ACS CB as well as from other journals. This presentation showed how some of the examples discussed in the morning session are beginning to appear in one way or another in a variety of publications. Through examples from *Amazon*, *Science*, *Nature*, and ACS CB, features to find related information, evaluate content, and organize content were demonstrated. Interactive features were also demonstrated, although this is probably the last area to be incorporated routinely within the scientific publishing community.

In the spirit of "practicing what you preach", the entire symposium was recorded, using iPod technology, and the recordings, along with PowerPoint and PDF files of the presentations, have been posted on the CINF web site, under the "Technical Session" link for Atlanta at "Meetings" at <http://www.acscinf.org>. Brian Lynch was responsible for the recording and described the procedure during his talk. So you don't need to take my word for it, you can listen for yourself. The last speaker in the session illustrated of the power of the technology. As already mentioned, the speaker himself was not present in person. Yet the WebEx mechanism which brought his presentation, live, to the audience was recorded in the same manner, and there are no essential differences between his presentation and those of the other speakers. To highlight the data from both Lynch and Coffman, the posting of these presentations makes the session available to the 95% of ACS members who didn't

make it to the meeting or those who were attending one of the other 66 sessions occurring at the same time. Whether this really substitutes for attendance at a meeting, with the opportunity for face-to-face networking at receptions, meals, even informal hallway conversations, can be debated. What can't be debated, though, is that technology is making an impact on the way we interact with each other, and this impact is felt in our scientific interaction as well. Many of the technologies discussed in this session have exploded in the culture at large but have yet to gain wide acceptance in our professional world. To echo the words of Andrea Twiss-Brooks in her opening comments, the next generation, the "digital natives", could already



Image courtesy of Getty Images.

be using technologies which haven't even made it onto our radar. Many organizations, including universities, libraries, societies, and publishers, are experimenting with emerging technologies, not knowing which will capture the mind of the scientific user. The challenge for any organization is whether to jump in on the leading, and experimental, edge, knowing that some experiments will fail, or to join in later, once a technology has been proven, and play "catch-up". Whatever decision one makes, there are risks. The one thing we can say is that whatever the future brings, someone will have predicted it. We just don't know who. So stay tuned.

Small Molecule Signaling in *Caenorhabditis elegans*

Frank C. Schroeder*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

ABSTRACT Whereas the *C. elegans* genome was sequenced many years ago, the role of small molecule signals in its biology is still poorly understood. A recent publication reports the identification of two steroidal signaling molecules that regulate *C. elegans* reproductive development and dauer diapause via the nuclear receptor DAF-12. The two compounds, named dafachronic acids, represent the first endogenous ligands identified for any of the 284 nuclear receptors in *C. elegans*.

Because of its short life span and genetic tractability, the nematode *Caenorhabditis elegans* has long been a pet organism of geneticists and molecular biologists. As one of the first complex organisms whose genomes were sequenced, many of its physiological pathways show analogies to corresponding pathways in higher animals, with interesting implications for research related to human disease. Of particular interest are signaling pathways regulating development that control growth, reproductive maturation, and ultimately life span in *C. elegans* and thus might hold cues for the regulation of analogous endocrine signaling in higher organisms. Surprisingly, very little is known about the role that small molecules play in *C. elegans* endocrine signaling. In a beautiful example for the use of genetic information to deduce structure and function of a small molecule signal, Motola *et al.* (1) have now identified two steroidal hormones, the dafachronic acids **1** and **2** (Figure 1), as the long-anticipated endogenous ligands of the nuclear receptor DAF-12.

The life cycle of *C. elegans*, which normally develops through four larval stages before reaching adulthood, can include a unique phase of metabolic diapause, which seemingly allows the larvae to put aging on hold. Under unfavorable environmental conditions, development arrests prior to reaching reproductive maturity, and the larvae enter the so-called “dauer” stage (from the German “dauer” for durable), an alternative, nonfeeding larval stage which

can persist for up to 2 months, compared to a normal life span for *C. elegans* of 14 days (2). Upon return to favorable conditions, dauer larvae resume feeding and continue normal development into adulthood. The discovery of such a well-defined phase of metabolic diapause suggested the presence of an elaborate circuitry for its control and for coordination of cellular programs required for the associated stage transitions throughout the organism (3). In fact, studies of the genes involved in dauer formation have provided tantalizing insights in regulation of metabolism, reproductive development, and life span of *C. elegans*. Genetic screens for mutants that either cannot attain the dauer stage or form dauer larvae constitutively have identified about three dozen genes directly involved in dauer stage control (2). A main switch in the signaling pathway controlling dauer formation appears to be the nuclear receptor DAF-12 (DAuer Formation). Loss of *daf-12* as well as certain *daf-12* mutations result in larvae that cannot attain the dauer stage and further develop into heterochronic phenotypes, indicating that *daf-12* is required for dauer formation as well as for proper developmental timing (3).

DAF-12 is one of at least 284 nuclear receptors in *C. elegans*, many of which have been shown to serve important roles in reproductive development and metabolism (4); however, prior to the identification of the dafachronic acids by Motola *et al.*, not a single ligand that regulates their functions had been identified. In the case of DAF-12,

*To whom correspondence should be addressed.
E-mail: Frank_Schroeder@hms.harvard.edu.

Published online May 19, 2006
10.1021/cb600173t CCC: \$33.50

© 2006 by American Chemical Society

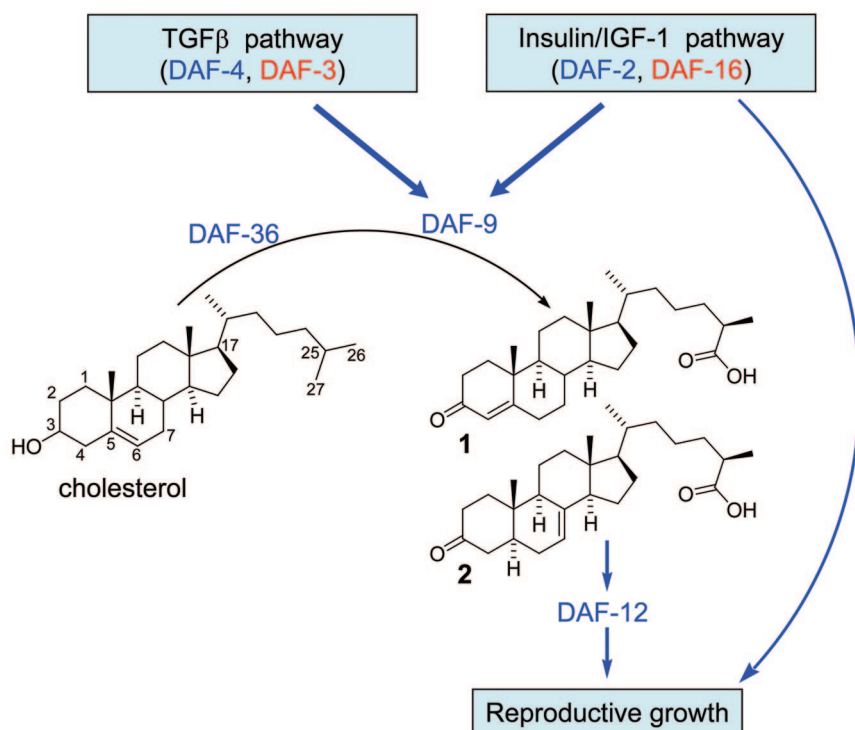


Figure 1. Regulation of reproductive growth in *C. elegans* via DAF-12. Under favorable environmental conditions, signaling through TGFβ and insulin/IGF-1 pathways activates genes including *daf-9* required for the synthesis of Δ⁴- and Δ⁷-dafachronic acids (1 and 2), which bind to DAF-12 promoting reproductive growth. Under unfavorable conditions, *daf-9* is inactive; thus synthesis of the dafachronic acids ceases and the resulting unliganded DAF-12 induces dauer diapause (not shown). DAF-36 functions as a Rieske-like oxygenase, which introduces the Δ⁷ double bond in the synthesis of 2 (8).

strong evidence had accumulated suggesting the presence of a small molecule ligand that would inhibit its dauer-promoting activity and trigger its functions relating to reproductive development. Two important signaling pathways converge on DAF-12, the insulin/IGF-1 (Insulin-like Growth Factor 1) pathway acting through DAF-2 and the transcription factor DAF-16, a worm-ortholog of FOXO, and the TGFβ (Transforming Growth Factor β) pathway, acting through DAF-4 and the Smad transcription factor DAF-3 (Figure 1). These pathways had been shown to act cell nonautonomously, suggesting that they directly or indirectly induce the production of a small-molecule ligand of DAF-12. Of particular

significance was the finding that a CYP2 cytochrome-P450 enzyme, DAF-9, which is expressed only in a few specific cell types, acts cell nonautonomously directly upstream from DAF-12 (5). Furthermore, *daf-12* mutants mapping to the predicted binding sites of the presumed hormone closely mimic phenotypes of *daf-9* mutants (6). As mammalian CYP2 enzymes can metabolize steroid hormones and the vertebrate DAF-12 orthologs represent steroid hormone receptors (3), it seemed likely that DAF-9 would participate in the synthesis of a steroidal ligand of DAF-12. Additional evidence showed that lipophilic extracts from wild-type worms can rescue worms that are dauer-constitutive as a result of a *daf-9*

defect (7), suggesting the presence of the suspected steroid hormone(s) in these extracts.

In order to identify the suspected small molecule hormone, a traditional natural products chemist might have resorted to activity-guided fractionation of the lipophilic worm extract, with the intention of isolating the active component(s) in pure form to determine their structure *via* spectroscopic methods. However, Motola *et al.* pursued a different approach. As a first step, they screened a variety of commercially available steroids for DAF-12 binding using a GAL4-DAF-12 cotransfection assay, which identified a 3-keto sterol, 3-ketolithocholic acid, as a weak activator of DAF-12. As the corresponding alcohol, lithocholic acid, was not active, it was concluded that a 3-keto functionality was essential. To confirm the relevance of a 3-keto functionalization and to further explore the role of DAF-9, Motola *et al.* incubated a series of 3-keto-sterols with *Sf9* microsomes containing DAF-9 and screened the resulting extracts for their ability to rescue *daf-9(-)* worms. Extracts from the DAF-9 microsomes incubated with 4-cholesten-3-one or lathosterone, two cholesterol metabolites known to occur in *C. elegans*, resulted in 100% rescue of the *daf-9(-)* mutants, producing a phenotype indistinguishable from wild-type adults, whereas controls exposing *daf-9(-)* mutants to unmetabolized 4-cholesten-3-one or lathosterone did not recover. These results suggested that *daf-9* encodes an enzyme converting the 3-ketosterols 4-cholesten-3-one and lathosterone into steroidal hormones which then activate DAF-12. Additional chemical characterization of the 4-cholesten-3-one or lathosterone metabolites obtained from DAF-9 microsomes revealed that DAF-9 hydroxylates and then further oxidized these substrates in a non-stereoselective manner at C-26/C-27 in the side chain, producing two diastereomers of the dafachronic acids 1 and 2 (from dauer formation and hetero-

chronic). Motola *et al.* confirmed the presence of the dafachronic acids in lipophilic wild-type worm extracts *via* a short fractionation scheme. Using synthetic samples of **1**, the authors showed that the (25*S*)-diastereomers of **1** and **2** in fact constitute the ligands that activate DAF-12 in cotransfection assays and rescue dauer-constitutive *daf-9* mutants.

Both **1** and **2** are potently active. (25*S*)- Δ^4 -Dafachronic acid (**1**) completely rescued *daf-9(-)* worms at concentrations of 250 nM, and it appears that (25*S*)- Δ^7 -dafachronic acid (**2**), which has not yet been synthesized, might exhibit even higher specific activity. As expected, the dafachronic acids act downstream of the insulin/IGF-1 and TGF β pathways; however, insulin/IGF-1 signaling might also act downstream of the dafachronic acids, or *via* a parallel pathway, as (25*S*)- Δ^4 -dafachronic acid (**1**) did not fully rescue certain *daf-2* mutants. It seems likely that the two dafachronic acids have somewhat distinct biological function, which will have to be explored in more detail once (25*S*)- Δ^7 -dafachronic acid (**2**) becomes synthetically available. Furthermore, the question remains whether the much less active (25*R*)-diastereomers of **1** and **2** are indeed produced *in vivo* and, if so, what possible function they might serve. There is strong evidence that DAF-12 not only is involved in the critical decision between dauer phase and reproductive growth but also plays an important role in the control of adult life span *via* DAF-16 (**8**). Future synthetic availability of both ligands and possibly of derivatives that act as DAF-12 agonists or antagonists will greatly aid further exploration of DAF-12 biological function.

The identification of the dafachronic acids represents only the beginning of molecular endocrinology in *C. elegans*, and a chemist might well ask, why did it take so long? As one of the best-studied higher life forms on earth, next to fruit flies and *Arabidopsis*, our understanding of *C. elegans*

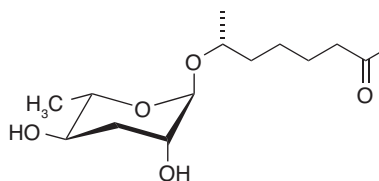


Figure 2. Structure of daumone, **3**.

small molecule chemistry seems highly underdeveloped. Just recently work by Jeong *et al.* (**9**) addressed another long-standing question in *C. elegans* signaling: What are the environmental cues that trigger entry and exit from the dauer stage? Over 20 years ago, Riddle and Golden (**10**) had shown that a group of unknown small molecules released from *C. elegans* constitute one primary cue for dauer entry. Jeong *et al.* now identified a glycoside of the dideoxysugar ascarylose, which they termed “daumone” (**3**), as the dauer pheromone; however, the exact mechanism of daumone signaling is not yet understood.

Still, there are many more observations in *C. elegans* biology that suggest the presence of small molecule signals, whose identification would contribute greatly to the understanding of corresponding pathways. For example, germline control of adult longevity *via* DAF-12 and DAF-16 might involve additional small molecule signals (**8**). And, as Motola *et al.* point out, the identification of the dafachronic acids as ligands of the nuclear receptor DAF-12 still leaves 283 of 284 nuclear receptors in *C. elegans* as orphans, providing plenty of opportunity for small molecule chemists.

REFERENCES

1. Motola, D. L., Cummins, C. L., Rottiers, V., Sharma, K. K., Li, T., Li, Y., Suino-Powell, K., Xu, H. E., Auchus, R. J., Antebi, A., and Mangelsdorf, D. J. (2006) Identification of Ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*, *Cell* **124**, 1209–1223.
2. Riddle, D. L., and Albert, P. S. (1997) in *C. elegans II* (Riddle, D. L., Blumenthal, T., Meyer, B. J., and Priess, J. R., Eds.), pp 739–768, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

3. Antebi, A., Yeh, W. H., Tait, D., Hedgecock, E. M., and Riddle, D. L. (2000) *daf-12* encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C. elegans*, *Genes Dev.* **14**, 1512–1527.
4. Gissendanner, C. R., Crossgrove, K., Kraus, K. A., Maina, C. V., and Sluder, A. E. (2004) Expression and function of conserved nuclear receptor genes in *Caenorhabditis elegans*, *Dev. Biol.* **266**, 399–416.
5. Mak, H. Y., and Ruvkin, G. (2004) Intercellular signaling of reproductive development by the *C. elegans* DAF-9 cytochrome P450, *Development* **131**, 1777–1786.
6. Jia, K., Albert, P. S., and Riddle, D. L. (2002) DAF-9, a cytochrome P450 regulating *C. elegans* larval development and adult longevity, *Development* **129**, 221–231.
7. Gill, M. S., Held, J. M., Fisher, A. L., Gibson, B. W., and Lithgow, G. J. (2004) Lipophilic regulator of a developmental switch in *Caenorhabditis elegans*, *Aging Cell* **3**, 413–421.
8. Berman, J. R., and Kenyon, C. (2006) Germ-cell loss extends *C. elegans* life span through regulation of DAF-16 by *kri-1* and lipophilic hormone signaling, *Cell* **124**, 1055–1068.
9. Jeong, P. Y., Jung, M., Yim, Y. H., Kim, H., Park, M., Hong, E., Lee, W., Kim, Y. H., Kim, K., and Paik, Y. K. (2005) Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone, *Nature* **433**, 541–545.
10. Golden, J. W., and Riddle, D. L. (1984) A *Caenorhabditis elegans* dauer-inducing pheromone and an antagonistic component of the food supply, *J. Chem. Ecol.* **10**, 1265–1280.

Reclamation of Proteins from the Cellular Scrap Heap

Jason E. Gestwicki*

Department of Pathology and the Life Sciences Institute, University of Michigan, 210 Washtenaw Avenue, Ann Arbor, Michigan 48109-2216

Newly synthesized polypeptides that are unable to fold in a timely manner are targeted to the proteasome. While this quality control checkpoint is intended to avoid harmful accumulation of misfolded intermediates, premature disposal can also cause severe loss-of-function defects. In these cases, the levels of folded protein can sometimes be restored by treatment with that protein's ligands. Because these compounds mimic the productive folding activity of the cell's own chaperones, they are termed chemical chaperones. However, the mechanism used by these molecules to rescue otherwise doomed polypeptides has been uncertain. On page 235 of this issue, Sawkar *et al.* (1) provide our best glimpse of chemical chaperones in action. Using unstable glucocerebrosidase (GC) mutants as a model, Sawkar *et al.* show that inhibitors leverage their binding energy to favor productive folding in the environment of the endoplasmic reticulum. This allows the enzyme to escape capture by the quality control machinery. These important insights should facilitate the design of small molecules that snatch specific proteins from the cellular scrap heap. However, this nontraditional drug strategy requires a shift in thinking for chemists and biologists alike; counterintuitive concepts, such as active site inhibitors that essentially behave as agonists at the cellular level, will need to be mastered. The more we understand about the activity of chemical chaperones, the greater potential

they have to treat the growing number of disorders recognized, broadly, as misfolding diseases.

Most newly synthesized proteins require the assistance of cellular chaperones to interpret the folding information contained in their primary sequence. For example, peptides expressed into the lumen of the endoplasmic reticulum (ER) are subject to a large family of heat shock proteins, cochaperones, lectins, membrane-spanning translocons, and ubiquitin-conjugating enzymes (for recent reviews see refs 2 and 3). If prolonged failures are detected, this machinery also mediates the ER-associated degradation (ERAD) that eliminates the offending polypeptide. The choice between continued folding and terminal disposal is central to protein homeostasis and critical to the ongoing health of the cell: when misfolded intermediates are left unattended, they can disrupt neighboring proteins and nucleate cytotoxic aggregates.

Mutations that leave a protein susceptible to misfolding are associated with numerous diseases (4, 5), including the neurodegenerative disorders (*e.g.*, Alzheimer's and Parkinson's diseases), peripheral amyloidoses, cystic fibrosis, and certain lysosomal storage disorders (*e.g.*, Gaucher disease, see below). Some of these diseases, such as Alzheimer's, are characterized by the formation of protease-resistant aggregates and an associated gain-of-function toxicity. Others are caused by loss-of-function defects; pathology arises

ABSTRACT A growing number of diseases have been associated with protein misfolding. Thus, strategies that use small molecules to adjust folding tendencies have therapeutic potential. However, progress in this area has been hampered by an insufficient description of the molecular underpinnings of protein instability within the cell. In a recent report, a chemical approach was taken to probe the mechanism by which Gaucher disease associated mutations in glucocerebrosidase destabilize that enzyme and lead to its destruction. These studies provide a blueprint for the design of "chemical chaperones" for the exploration of cellular protein homeostasis and the treatment of misfolding diseases.

*To whom correspondence may be addressed.
E-mail: gestwick@umich.edu.

Published online May 19, 2006
10.1021/cb001784 CCC: \$33.50
© 2006 by American Chemical Society

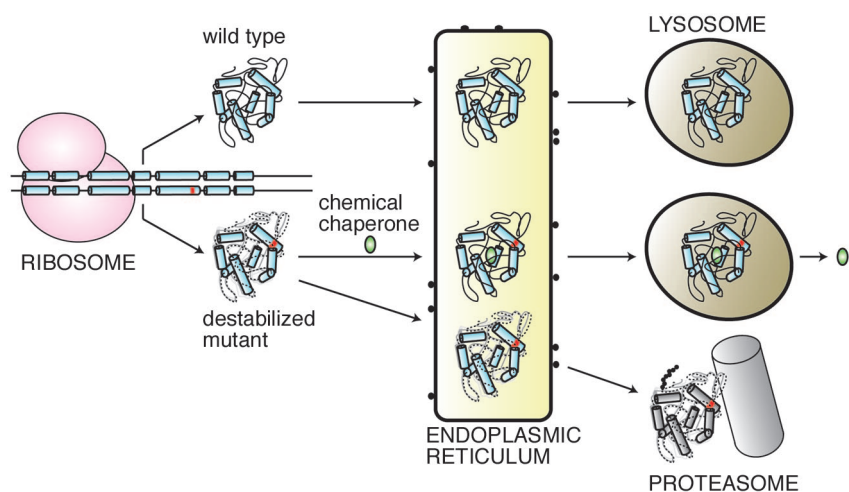


Figure 1. Model for chemical chaperone-assisted stabilization of glucocerebrosidase. Mutants harboring destabilizing substitutions (red) fail to fold properly in the neutral ER lumen. These polypeptides are destined for removal by the proteasome. However, binding to chemical chaperones rescues these mutants and permits trafficking to the lysosome. After arrival in this compartment, the chemical chaperone is no longer required for folding and its dissociation restores enzymatic activity.

from the diminished activity of an essential protein. In perhaps the best-known example of this type, reduced cell surface levels of the cystic fibrosis transmembrane conductance regulator (CFTR) underlie pathology of that disease (6).

Gaucher disease is one of approximately 40 lysosomal storage disorders in which defects arise from incomplete catabolism of glycosphingolipids (7). In Gaucher patients, harmful accumulation of glucosylceramide is caused by loss-of-function mutations in a specific lysosomal GC. While some of these mutants appear to encode enzymatically inactive products, other common mutations may simply destabilize the enzyme and mark it for ERAD. Because these later defects are distal from the active site, they do not directly impinge on enzymatic function. This suggests that, if the proteins were safeguarded through the ER's quality control machinery, they would retain a high degree of enzymatic function. Consistent with this hypothesis, osmolytes that generally favor protein folding, such as dimethyl

sulfoxide (DMSO) and glycerol, have been found to rescue lysosomal GC activity (8).

Chemical chaperones were developed as a pharmacologically superior alternative to osmolytes (for recent reviews see refs 9–12). These small molecules are designed to specifically interact with enzymes and protect against misfolding. For example, in 1999, a competitive inhibitor of α -galactosidase was found to enhance total enzyme activity in cells harboring an unstable mutant (13). However, proper use of these reagents poses an interesting conundrum: how can an inhibitor be used to *enhance* an enzyme's function? Our understanding of this apparent contradiction would directly benefit from a deeper mechanistic understanding of how chemical chaperones work. For example, how much of the lost folding energy needs to be restored to avoid ERAD? In what subcellular compartment must chaperoning occur?

A key step forward in our understanding of this process is reported by Jeffery Kelly's laboratory (1). The Kelly laboratory has previ-

ously established that nonyl deoxynojirimycin, a glucocerebrosidase inhibitor, can rescue lysosomal GC activity in a Gaucher disease model (14). Now, for the first time, the Kelly laboratory has quantified the energetics of an unstable GC variant, N370S. They found that N370S is thermally unstable at the neutral pH of the ER but more stable (and partially functional) in the lysosome's acidic environment. Importantly, nonyl deoxynojirimycin was able to partly restore folding energy at neutral pH. This result strongly supports a model (Figure 1) in which membrane-permeable ligands directly facilitate folding of nascent polypeptide in the ER. Following successful trafficking, drug is released and enzymatic function restored. In addition to these mechanistic insights, the Kelly laboratory also reports that drug treatment reverses trafficking defects and restores lysosomal enzymatic activity in fibroblasts derived from Gaucher disease patients expressing the N370S mutation. These physiologically significant findings affirm the potential therapeutic value of chemical chaperones. Finally, the mechanistic and cellular details gleaned from these studies provide the foundation for optimizing the activity of chemical chaperones. For example, next generation molecules might be designed to bind selectively at neutral pH with rapid dissociation in acidic compartments (e.g., using protonatable groups with an appropriate pK_a). These compounds might be expected to have a superior therapeutic window, owing to their combination of high chaperone activity and low inhibitory potential.

In addition to their promise as therapeutics for misfolding diseases, chemical chaperones might become important reagents for exploring the fundamental biology of protein homeostasis. Currently, inhibitors of cellular chaperones are among the best tools for these studies (15). These drugs rapidly inactivate the refolding machinery and reveal the biological consequences of

cellular chaperone function. For example, geldanamycin, an inhibitor of heat shock protein 90 (Hsp90), was used to define this protein's substantial role in p53 regulation (16). Chemical chaperones complement these reagents by enhancing the folding energy of specific targets. Used in combination, cellular chaperone inhibitors and chemical chaperones might illuminate how the decision is made to refold or degrade proteins.

REFERENCES

1. Sawkar, A. R. (2006) *ACS Chem. Biol.* **1**, 235–251.
2. Ahner, A., and Brodsky, J. L. (2004) Checkpoints in ER-associated degradation: excuse me, which way to the proteasome?, *Trends Cell. Biol.* **14**, 474–478.
3. Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005) ERAD: the long road to destruction, *Nat. Cell. Biol.* **7**, 766–772.
4. Selkoe, D. (2003) Folding proteins in fatal ways, *Nature* **426**, 900–904.
5. Stefani, M., and Dobson, C. M. (2003) Protein aggregation and aggregation toxicity: new insights into protein folding, misfolding diseases and biological evolution, *J. Mol. Med.* **81**, 678–699.
6. Kopito, R. R. (1999) Biosynthesis and degradation of CFTR, *Physiol. Rev.* **79**, S167–S173.
7. Sawkar, A. R., D'Haese, W., and Kelly, J. W. (2006) Therapeutic strategies to ameliorate lysosomal storage disorders—a focus on Gaucher disease, *Cell. Mol. Life Sci.*, published online Mar 29, <http://dx.doi.org/10.1007/s00018-005-5437-0>.
8. Brown, C. R., HongBrown, L. Q., and Welch, W. J. (1997) Correcting temperature-sensitive protein folding defects, *J. Clin. Invest.* **99**, 1432–1444.
9. Ulloa-Aguirre, A., Janovick, J. A., Brothers, S. P., and Conn, P. M. (2004) Pharmacologic rescue of conformationally-defective proteins: implications for the treatment of human disease, *Traffic* **5**, 821–837.
10. Perlmutter, D. H. (2002) Chemical chaperones: a pharmacological strategy for disorders of protein folding and trafficking, *Pediatr. Res.* **52**, 832–836.
11. Bemier, V., Lagace, M., Bichet, D. G., and Bouvier, M. (2004) Pharmacological chaperones: potential treatment for conformational diseases, *Trends Endocrinol. Metab.* **15**, 222–228.
12. Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein misfolding diseases, *Nature* **426**, 905–909.
13. Fan, J. Q., Ishii, S., Asano, N., and Suzuki, Y. (1999) Accelerated transport and a maturation of lysosomal α -galactosidase A in Fabry lymphoblasts by an enzyme inhibitor, *Nat. Med.* **5**, 112–115.
14. Sawkar, A. R., Adamski-Wemer, S. L., Cheng, W.-C., Wong, C.-H., Beutler, E., Zimmer, K.-P., and Kelly, J. W. (2005) Gaucher disease-associated glucocerebrosidases show mutation-dependent chemical chaperoning profiles, *Chem. Biol.* **12**, 1235–1244.
15. Smith, D. F., Whitesell, L., and Katsanis, E. (1998) Molecular chaperones: biology and prospects for pharmacological intervention, *Pharm. Rev.* **50**, 493–513.
16. Galigniana, M. D., Harrell, J. M., O-Hagen, H. M., Ljungman, M., and Pratt, W. B. (2004) Hsp90-binding immunophilins link p53 to dynein during p53 transport to the nucleus, *J. Biol. Chem.* **279**, 22483–22489.

For a Healthy Histone Code, a Little SUMO in the Tail Keeps the Acetyl Away

Jorge A. Iñiguez-Lluhí*

Department of Pharmacology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0632

ABSTRACT Chemical modification of histones through a growing number of post-translational mechanisms is an integral part of transcription. A recent report provides exciting new evidence that conjugation of the ubiquitin-like protein SUMO to histones opposes acetylation and establishes SUMOylation as an important histone mark linked to transcriptional repression.

Histones are physically and figuratively central to the compact packaging of eukaryotic genomes. All DNA-based transactions such as replication, transcription, recombination, and repair require direct access to the nucleic acids, and this is often accompanied by chemical modification of histones, many of which target lysine residues within their N and C terminal tails (1). Over the past few years, the list of such modifications has continued to grow not only in number but also in the size of the modifying group. The complexity of such modifications and their interrelations suggests the existence of a “histone code” where distinct patterns of modification serve specific roles (2). Although the role of histone modification by small chemical groups such as acetyl and methyl is becoming better understood, the function and mechanism of action of other modifications remain poorly defined. In general, acetylation of histone lysine residues is associated with actively transcribed chromatin, and the mechanism of action of transcriptional activators often involves the recruitment of complexes harboring acetylase activity. In vertebrates and fission yeast, methylation of specific lysine residues is instrumental in the establishment of repressive heterochromatic domains by serving as a recruitment site for heterochromatin protein 1. In contrast, no such repressive mark had been identified in budding yeast. A potential answer to this apparent imbalance is provided in a recent report in the April issue of *Genes and Development* (3). Through an elegant combination of chemical,

biochemical, and genetic approaches, Nathan and co-workers provide evidence that SUMOylation of histones antagonizes their acetylation and serves as an evolutionarily conserved repressive mark.

The conjugation of SUMO to target proteins follows a pathway analogous to that of ubiquitination (Figure 1) and requires dedicated E1 activating (SAE1/SAE2) and E2 conjugating (UBC9) enzymes. UBC9 interacts directly with substrates to catalyze the formation of an isopeptide bond between the C-terminus of SUMO and the amino group of the target lysine. This step is facilitated by E3 SUMO ligases. SUMOylation is reversible and specific isopeptidases release the SUMO moiety. SUMOylation appears to regulate target proteins by altering their interaction with other proteins, their intrinsic function, or localization. SUMOylation can also compete with other lysine targeted modifications such as ubiquitination or acetylation (4). The function of SUMOylation has been substantially examined in the context of sequence-specific transcription factors, and for the most part this modification leads to inhibition of transcription (5). In many cases, SUMOylation exerts its effects in a striking promoter context dependent manner such that effective inhibition depends on recruitment of SUMOylated factors to multiple independent sites on promoters (6). SUMOylation however, extends to other components of the transcriptional machinery, and Shiio *et al.* (7) described the SUMOylation of mammalian histones in 2003.

*To whom correspondence should be addressed.
E-mail: iniguez@umich.edu.

Published online May 19, 2006
10.1021/cb600188m CCC: \$33.50

© 2006 by American Chemical Society

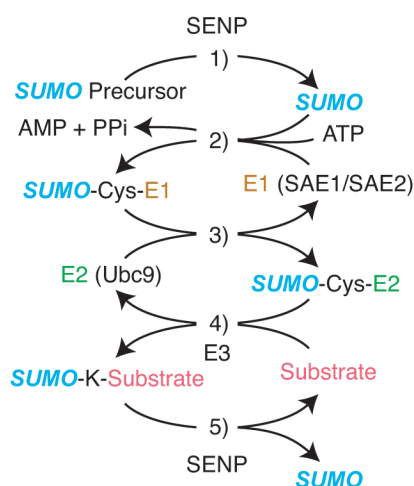


Figure 1. SUMOylation pathway. 1) Initial processing of SUMO precursor by SUMO-specific proteases (SENP) removes C terminal residues to generate a new GlyGly C terminus. 2) ATP-dependent activation of SUMO by the SUMO-specific E1 leads to the formation of a thioester bond between the C terminus of SUMO and a cysteine in the SAE2 subunit. 3) The SUMO moiety is transferred to the SUMO E2 ligase UBC9 through a trans-esterification reaction. 4) Ubc9-catalyzed conjugation of SUMO to substrate leads to the formation of an isopeptide bond between the C terminus of SUMO and the amino group of the target lysine. This step is enhanced by E3 ligases. 5) SUMO conjugation is reversible through the isopeptidase activity of SUMO-specific proteases.

The current paper by Nathan *et al.* (3) provides compelling evidence that this modification plays important roles in the control of transcription by antagonizing histone acetylation (Figure 2). This group demonstrated that in budding yeast all four core histones (H2A, H2B, H3, and H4) are SUMOylated in a manner that depends genetically on an intact SUMOylation machinery. Although the extent of modification is less than a few percent, they detected multiple modified forms, suggestive of multisite modification or poly-SUMO chain formation. Notably, the modification does not appear to be equal for all histones since the H2AZ histone variant, which is usually associated with active transcription, is SUMOylated to a substantially lower extent. Mass spectrometry analysis of protease digests identified lysines 6 and 7 in H2B and K 126 in H2A as sites of modification. Mutagenesis studies also implicated K16/17 in H2B as well as all five lysines in the N terminal region of H4. The heterogeneous pattern of modification and the lack of similarity to the canonical SUMOylation consensus indicated that histone SUMOylation can occur at multiple lysine residues without strict sequence requirements. This suggests that histone SUMOylation may rely on recruitment of UBC9

through adaptor E3 proteins. Interestingly, many of the identified SUMOylation sites are also targeted for acetylation.

Chromatin immunoprecipitation (ChIP) analysis using antibodies specific for individual modifications has allowed the mapping of the distribution and temporal evolution of acetylation and methylation histone marks in the genome. SUMOylation, however, does not lend itself easily to this approach. The authors, however, devised a clever sequential immunoprecipitation and affinity isolation scheme (ChDIP) to gauge the presence of SUMO conjugates associated with individual histones, which in the case of H2B correspond mainly to SUMOylation of H2B

itself. Consistent with a role of histone SUMOylation in repression, this analysis revealed that SUMOylated H2B is present throughout the genome but is over-represented by ~2-fold at the transcriptionally silent telomeric regions where positive marks such as acetylation and ubiquitination are less frequent.

A more direct analysis of the functional consequences of histone SUMOylation came from the expression of H2B mutants bearing alanine substitutions at the four N-terminal lysines that are targeted for SUMOylation. These substitutions lead to a substantial but not complete loss of SUMO from H2B. If histone SUMOylation participates in transcriptional repression, loss of this modification should lead to derepression. Indeed, yeast expressing this H2B mutant displayed enhanced basal transcription of multiple genes under repressive conditions. By the same argument, enhanced SUMOylation of histones should lead to repression. To mimic persistently SUMOylated histones, the researchers

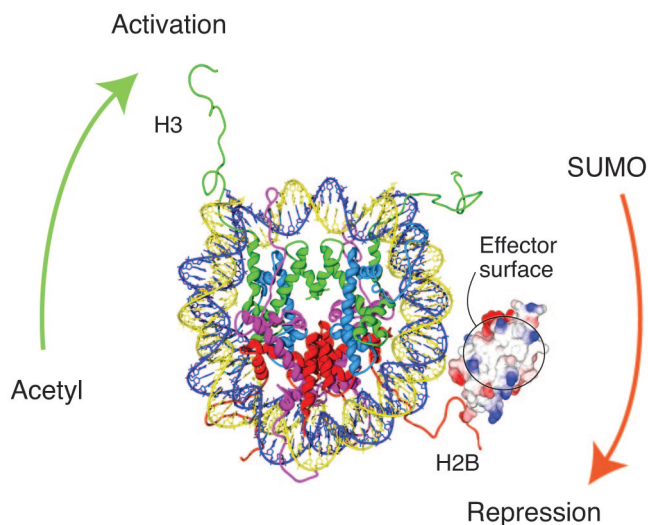


Figure 2. SUMOylation antagonizes acetylation. Model of a nucleosome core showing a single surface rendered SUMO molecule conjugated to K6 of one of the H2B N terminal tails. The basic surface in SUMO essential for its transcriptional repressive function is circled.

expressed a non-cleavable colinear fusion between the C terminus of SUMO and the N terminus of H2B or H4. Consistent with an inhibitory role of SUMOylation, this manipulation led to substantial reductions in the expression of the galactose-inducible gene Gal1. The loss of activity, however, could be a trivial consequence of steric hindrance or non-specific alterations in the function of the histones. To address this issue, the authors took advantage of recent data that identified a critical effector surface in SUMO essential for its transcriptional inhibitory function (8). Substitution of basic residues by acidic ones within this region (circled in Figure 2) eliminated the repressive effects of the SUMO-H2B fusion, indicating that the same effector surface of SUMO is required for its repressive effects in the context of both histones and transcription factors.

Since SUMOylation and acetylation sites overlap substantially, the authors examined the interplay between these modifications during the switch from repression to activation that accompanies an experimental shift in carbon source supply. A time course analysis indicated a reciprocal pattern of modifications where a transient loss of histone SUMOylation parallels enhanced acetylation both when the bulk of the histones are analyzed or at a specific promoter. These data argue for a counterregulatory role for SUMOylation in opposition to acetylation (Figure 2). Consistent with this model, reduction on the SUMOylation capacity of the cells by inactivation of Ubc9 or genetic disruption of E3 ligases leads to enhanced bulk and gene-specific acetylation of histones.

Like many good experiments that push a field forward, these new results raise a number of questions and provide new opportunities. The mechanism of antagonism between SUMOylation and acetylation remains obscure. Although both modifications target overlapping lysine residues,

acetylation is much more prevalent than SUMOylation. This argues that a direct competition mechanism is likely insufficient to account for their antagonism. This may not be so surprising since in many cases SUMO has important regulatory effects at disturbingly low stoichiometries. Like a discreet but influential member of a royal court, SUMO appears to be uncannily able to exert substantial influence without being detected. Nevertheless, the proteins that “read” the signal provided by SUMOylation of histones are likely to do so by binding to the critical effector surface in SUMO. Recent structural and functional data indicates that this surface is the site of binding of an emerging group of SUMO binding motifs (9).

Given the repressive role of histone SUMOylation, repressor functions within sequence-specific factors could operate by recruiting UBC9 and proteins with E3 activity towards histones. Studies in mammalian cells have indicated that UBC9 and members of the PIAS family of SUMO E3 ligases associate with multiple transcription factors that are themselves SUMOylated. In many cases though, SUMO E3s have repressive effects that depend on their E3 activity but are independent of the SUMOylation of the factor with which they interact. Histones may be one of the targets of such E3s. Given the intrinsic repressive function of SUMO and the growing number of proteins within transcription complexes that are SUMOylated, it will be important to determine the relative contributions of histone SUMOylation to repression. In contrast to a potential role of SUMO E3s in repression, recruitment of a SUMO-specific protease to erase the SUMO mark on histones could contribute to mechanisms of transcriptional activation. It is amply clear though that alterations in histone modifications have profound physiological and pharmacological consequences. Given the interplay between acetylation and SUMOylation and the potential of HDAC

inhibitors in the treatment of cancer (10), small molecule modulators of SUMOylation machinery may prove an attractive arena for pharmacological exploration.

Chemical complexity is at the center of biological processes, and the post-translational modifications that decorate histones provide a clear example on how biological systems exploit reversible chemical variations to control dynamically complex systems. Acetylation has commanded much of the recent “action”. Given its heft, it seems appropriate for SUMO to provide the “equal and opposite reaction” to maintain a well balanced code.

REFERENCES

1. van Attikum, H., and Gasser, S. M. (2005) The histone code at DNA breaks: a guide to repair? *Nat. Rev. Mol. Cell Biol.* **6**, 757–765.
2. Margueron, R., Trojer, P., and Reinberg, D. (2005) The key to development: interpreting the histone code? *Curr. Opin. Genet. Dev.* **15**, 163–176.
3. Nathan, D., Ingvarsdottir, K., Stemer, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B., Johnson, E. S., and Berger, S. L. (2006) Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications, *Genes Dev.* **20**, 966–976.
4. Johnson, E. S. (2004) Protein modification by SUMO, *Annu. Rev. Biochem.* **73**, 355–382.
5. Hay, R. T. (2005) SUMO: A history of modification, *Mol. Cell* **18**, 1–12.
6. Holmstrom, S., Van Antwerp, M. E., and Iniguez-Lluhi, J. A. (2003) Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15758–15763.
7. Shiio, Y., and Eisenman, R. N. (2003) Histone sumoylation is associated with transcriptional repression, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13225–13230.
8. Chupreta, S., Holstrom, S., Subramanian, L., and Iniguez-Lluhi, J. A. (2005) A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties, *Mol. Cell Biol.* **25**, 4272–4282.
9. Reverter, D., and Lima, C. D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex, *Nature* **435**, 687–692.
10. Drummond, D.C., Noble, C. O., Kirpotin, D. B., Guo, A., Scott, G. K., and Benz, C. C. (2005) Clinical development of histone deacetylase inhibitors as anticancer agents, *Annu. Rev. Pharmacol. Toxicol.* **45**, 495–528.

Act Globally, Think Locally: Systems Biology Addresses the PDZ Domain

Mark R. Spaller*

Department of Chemistry, Wayne State University, Detroit, Michigan 48202

Chances are you have a favorite protein (or know someone who does), upon which research attention is generously lavished. This typically manifests itself through the use of one or more highly focused techniques, singularly trained upon that special protein, in order to thoroughly characterize the chemical or biological behavior of interest. But that modus operandi has been changing, as evidenced by the rising tide of literature disclosing research that capitalizes on array technologies, in which collections of molecules are simultaneously screened and analyzed. One of the latest contributions to this trend is a report by Stiffler *et al.* (1), and on display is, among the newer of the high-throughput platforms, the protein microarray. Fittingly, it has been commissioned to investigate a structural class that is itself of relatively recent vintage, the PDZ domain.

So why pay attention to the PDZ domain? Simply put, because of the expansive range of cellular signaling functions performed by polypeptides that possess one or more of these units. Mammalian synapses are teeming with proteins bearing PDZ domains, the latter congregating with various membrane receptors, ion channels, and cell adhesion molecules, as well as those of cytoplasmic origin, with resultant activities that range from protein localization and clustering to trafficking (2). In addition to their responsibilities to neuronal cells, PDZ domains find employment in roles that are further removed, such as in the regulation of

drug transporters (3) and the organization of proximal tubules in the kidney (4). When one reflects upon how rapidly and widely these domains have infiltrated cellular biology, it should not elicit much surprise to learn that there is significant potential for therapeutic discovery (5).

At the biophysical root of these events, PDZ domains mediate specific protein–protein interactions. This behavior, coupled to their modular nature and modest size of approximately 90 residues, places the PDZ class into the larger fraternity of protein interaction domains (6). Considering how populated this domain family has become (Table 1), library-based approaches will be indispensable if we seek timely progress in determining the precise functional contributions its members make to the well-being of a cell. PDZ domains are elements in a large mosaic of binding networks that crisscross the cell both temporally and spatially. In

ABSTRACT Interaction networks, cartography and mapping, wiring and circuitry, whichever metaphor is invoked, the cardinal questions regarding cellular proteins are the same: What are they? How much is there? What do they do, and to whom do they do it? One of the more recent proteomics tools to pursue these lines of inquiry is the protein microarray, and a current report has unleashed this formidable technique upon a target of considerable biological interest, the PDZ domain family of signaling molecules.

TABLE 1. Numbers of selected protein interaction domains assigned within the human and mouse genomes^a

Domain	<i>Homo sapiens</i>	<i>Mus musculus</i>
PDZ	918	771
PTB	141	115
SH2	352	323
SH3	894	738

^aAll values are extracted from a table prepared by Bhattacharyya *et al.* (11).

*To whom correspondence should be addressed.
E-mail: mspaller@chem.wayne.edu.

Published online May 19, 2006
10.1021/cb600191y CCC: \$33.50

© 2006 by American Chemical Society

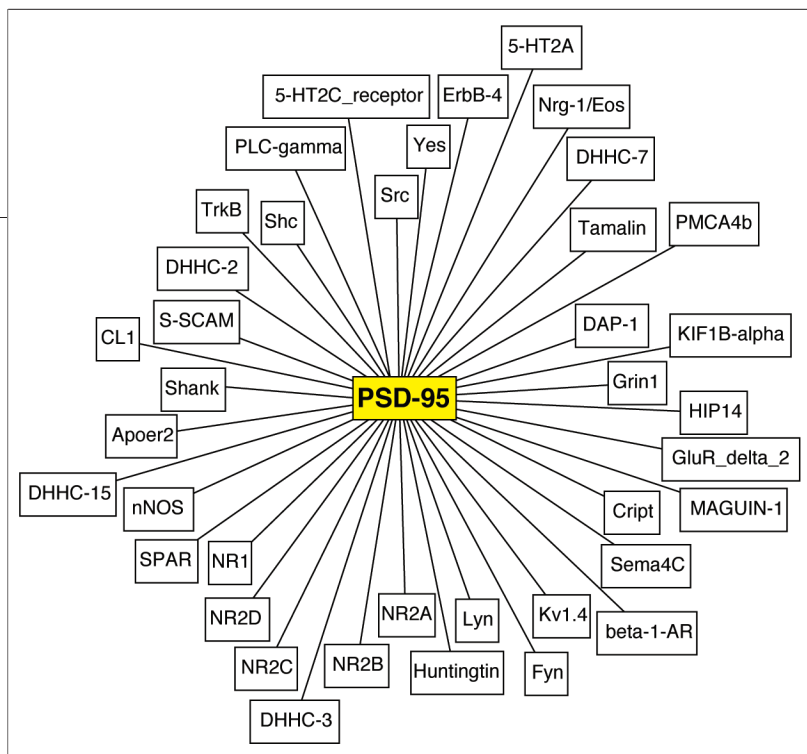


Figure 1. Protein–protein interaction map for PSD-95. The data derive from a search using the Biomolecular Interaction Network Database (BIND, (16)) and results graphed and edited with Cytoscape (17).

conjunction with a multitude of other proteins, they are the focus of various mapping efforts that represent the source of one pipeline that feeds into systems biology.

While this Big Picture discipline is still in its foundational stage (7), systems biology will need to assimilate mountain-ranges worth of data that pertain to the function, structure, and concentrations of cellular constituents over time. Chief among the challenges is correctly ascertaining all possible protein–protein interactions of biological consequence. Cellular network theory (8) has been summoned to help process this data and, to the extent possible, form the basis for computational tools to project the multidimensional information in a manner that is both intellectually and visually palatable for human consumption.

How is information like this obtained? To date, the identification of associating proteins on the large scale, and the subsequent mapping of their interaction networks, has relied primarily upon one of two experimental approaches: yeast two-hybrid assays and co-affinity purification with a bait protein, followed by mass spectroscopic analysis (co-AP/MS) (9).

Much data has been generated, but there are nagging questions about the accuracy of the predicted protein–protein interactions. Rates of false positives and false negatives are uncomfortably high, and the overlap of predicted interactions between some studies using the same methods is surprisingly low.

Technical and procedural improvements in experimental execution may diminish the

incidence of spurious results, but one approach to make the data less ambiguous and more reliable is a divide-and-conquer strategy: rather than confront a multidomain polypeptide in its entirety, carry out a dissection and tackle the domains individually. Take, for example, an archetypical PDZ domain protein, postsynaptic density-95 kDa (PSD-95). Constructing a simple map that reveals the observed binary interactions, while remaining fairly low on the scale of network complexity, still provides a surfeit of partners to examine (Figure 1). The task becomes simpler when deciding to concentrate on specific domains within PSD-95, such as the three of five that are of the PDZ variety (Figure 2).

This domain-centered approach allows for clearer experimental design and data analysis in both low- and high-throughput investigations (10, 11). An assortment of studies, each employing distinct methods to conduct a library-based screening of PDZ domain interactions, has been published (Table 2). Given the contrasting nature of each of the techniques, any direct comparisons between them must be qualified. For example, the enormous peptide diversity

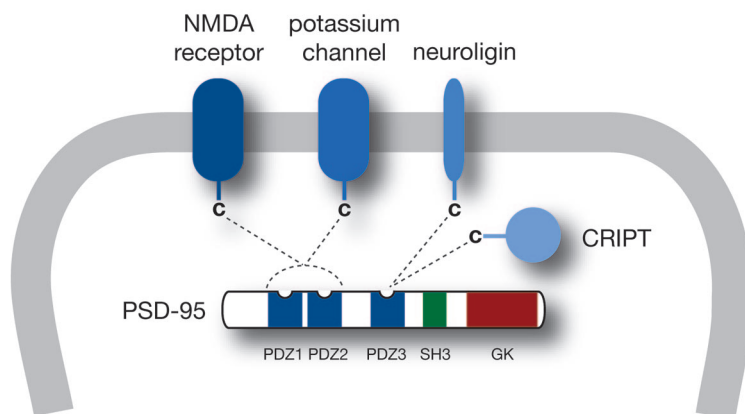


Figure 2. Minimized protein-binding scheme for PSD-95 near the postsynaptic membrane. The ‘c’ appendage denotes free carboxylate termini of proteins that interact with specified PDZ domains. All three PDZ domains depicted here are contained within the study by Stiffler *et al.* (1), as are the C-terminal sequences of the NMDA receptor (NMDAR2A), the potassium channel (Kv1.4), and CRIP1.

TABLE 2. Selected library approaches to probing PDZ domain interactions^a

Reference	Immobilized/Arrayed Component	Soluble Component
Songyang <i>et al.</i> (18)	PDZ domains [10]; <i>GST fusions on glutathione-bound agarose resin</i>	domain partners [2.5×10^7] ^b ; <i>peptide mixtures</i>
Fuh <i>et al.</i> (19)	PDZ domains [2]; <i>GST fusions on microtiter plates</i>	domain partners [2×10^{10}]; <i>peptides on M13 phage</i>
Wiedemann <i>et al.</i> (20)	domain partners [6223]; <i>peptides on cellulose membranes</i>	PDZ domains [3]; <i>GST fusions</i>
Fam <i>et al.</i> (21)	PDZ domains [96]; <i>His₆-S-Tag fusions on nylon membranes</i>	domain partners [1]; <i>polypeptide as GST fusion</i>
Stiffler <i>et al.</i> (1)	PDZ domains [26]; <i>thioredoxin-His₆ fusions on glass slides</i>	domain partners [20]; <i>N-fluorescently labeled peptides</i>

^aThe values in square brackets equal the number of different constructs or ligands used. ^bReported as total degeneracy.

of the phage display approach is striking, but they represent non-biased, random sequences; other studies employ far fewer peptides, but those are known PDZ binding motifs. In addition to these explorations that begin with a discrete, spatial separation of the arrayed and non-arrayed components, a report has recently been published that utilizes the yeast two-hybrid assay to evaluate five different PDZ domains when challenged with a random peptide library (12).

One commonality that benefits these disparate studies is the use of a relatively small ligand that is recognized by the PDZ domain; namely, a C-terminal-exposed peptide is surrogate for the full-sized endogenous binding partner. Aside from a few exceptional cases, PDZ domains seek out the very carboxylate end of their partner proteins; further, it appears that less than 10 (and sometimes only five or six) of those consecutive terminal residues are necessary and sufficient to capture maximal binding. Thus, one may efficiently mimic an array of protein–protein interactions in the technically more convenient manner of a protein–peptide ensemble.

The work of Stiffler *et al.* (1) builds upon such a peptide-based approach, aiming for

the development and validation of a protein microarray with which to study the binding and selectivity properties of PDZ domain proteins. This report follows closely that of another from the same laboratory that mapped the interaction profile of human Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains (13). As with this predecessor study, the PDZ domain investigation uses the same methodological blueprint, although the scope is smaller and more in keeping with the nature of a pilot study.

The broad objectives are twofold: identify, then quantify. The first requirement is to prepare and use the protein microarray to determine which of 26 mouse recombinant PDZ domain constructs binds to 20 fluorescent peptides whose sequences were derived from known partner proteins. The second task is to then conduct a solution-based fluorescence polarization (FP) assay for each of those same 520 possible interactions and correlate the measured affinity constants with the microarray responses.

In the development of the materials for use in such an array, uniformity and homogeneity are the dual watchwords. All PDZ domains are subcloned in identical

fashion as thioredoxin–polyhistidine fusion proteins, and experimental confirmation is made of sample purity and monomeric disposition. On the partner side, decapeptides are designed and synthesized so as to encompass the presumed binding epitope. To each, an identical tripeptide segment, NNG, is affixed to promote solubility, followed by the rhodamine tag 5(6)-TAMRA. Taken together, both domain and peptide partner compo-

ponents reflect a level of diversity based on the different binding motifs that PDZ domains can accommodate. Importantly, the array incorporates both single and dual domains, the latter being structures that closely border one another in their shared polypeptide, and which may exhibit forms of naturally-occurring cooperative behavior in their structural or conformational forms as well as in their binding.

Without recounting the many issues of technical development and quality control that arise during the construction and use of the actual PDZ domain microarray, suffice it to say that the investigators are cognizant of all the right concerns, and admirably address as many as reasonably possible. The level of rigor exhibited sets a standard that would be well to emulate in future investigations by those developing or applying protein microarrays. Their efforts are repaid in that the combination of the microarray and FP assay data reveals several important conclusions for ligands of typical interest (those that possess moderate-to-high affinity): (a) false-positive and false-negative rates are each held to about 14%, quite reasonable in the face of much higher error seen with other

high-throughput methods; (b) 85% of previously documented interactions are rediscovered; and (c) new and unreported associations are detected and quantified.

In addition to mapping out the signaling networks that engage PDZ domains, another obvious yet key application of this technology includes its use in screening combinatorial libraries and determining the selectivity profiles of new or established inhibitory ligands. Another avenue of interest would be to examine the effect of “macromolecular crowding” conditions, those that mimic the excluded volume characteristic of a cell’s fluid environment, upon PDZ domain interactions (14). This phenomenon, in conjunction with the prospect that various proteins may exist in conformational ensembles with distinct binding behaviors (15), may result in discrepancies between how interactions occur *in vivo* and in the artificial conditions common to *in vitro* binding assays. The crowding effect has in fact been little studied even in nonarray settings, but the robust nature of the PDZ domain microarray’s performance as revealed in this study suggests that the technique might be appropriate for this application.

“All politics is local” is an adage of statecraft, but the same can be said of biomolecular interactions. Microarrays are a means to an end, and data collected on a global scale must eventually be subjected to microscopic evaluation. Aside from the specific details of what has been learned in this successful investigation, the larger accomplishment of this report should be to foster future screening efforts that will cast a wider net and encompass a greater number of the known PDZ domains. The resulting outpouring of reliable data from such an exercise would dramatically accelerate the rate at which PDZ-mediated interactions are studied and provide us with the next generation of proteins upon which to lavish our attention.

REFERENCES

1. Stiffler, M. A., Grantcharova, V. P., Sevecka, M., and Macbeath, G. (2006) Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays, *J. Am. Chem. Soc.* **128**, 5913–5922.
2. Kim, E., and Sheng, M. (2004) PDZ domain proteins of synapses, *Nat. Rev. Neurosci.* **5**, 771–781.
3. Kato, Y., Watanabe, C., and Tsuji, A. (2006) Regulation of drug transporters by PDZ adaptor proteins and nuclear receptors, *Eur. J. Pharm. Sci.* **27**, 487–500.
4. Biber, J., Gislser, S. M., Hernando, N., and Murer, H. (2005) Protein/protein interactions (PDZ) in proximal tubules, *J. Membr. Biol.* **203**, 111–118.
5. Dev, K. K. (2004) Making protein interactions druggable: targeting PDZ domains, *Nat. Rev. Drug Discovery* **3**, 1047–1056.
6. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains, *Science* **300**, 445–452.
7. Kirschner, M. W. (2005) The meaning of systems biology, *Cell* **121**, 503–504.
8. Albert, R. (2005) Scale-free networks in cell biology, *J. Cell Sci.* **118**, 4947–4957.
9. Uetz, P., and Finley, R. L., Jr. (2005) From protein networks to biological systems, *FEBS Lett.* **579**, 1821–1827.
10. Santonico, E., Castagnoli, L., and Cesareni, G. (2005) Methods to reveal domain networks, *Drug Discovery Today* **10**, 1111–1117.
11. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits, *Annu. Rev. Biochem.*, published online Mar 17, <http://dx.doi.org/10.1146/annurev.biochem.75.103004.142710>.
12. Song, E., Gao, S., Tian, R., Ma, S., Huang, H., Guo, J., Li, Y., Zhang, L., and Gao, Y. (2006) A high-efficiency strategy for binding property characterization of peptide-binding domains, *Mol. Cell. Proteomics*, published online Apr 23, <http://dx.doi.org/10.1074/mcp.M600072-MCP200>.
13. Jones, R. B., Gordus, A., Krall, J. A., and MacBeath, G. (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays, *Nature* **439**, 168–174.
14. Minton, A. P. (2005) Influence of macromolecular crowding upon the stability and state of association of proteins: predictions and observations, *J. Pharm. Sci.* **94**, 1668–1675.
15. Teague, S. J. (2003) Implications of protein flexibility for drug discovery, *Nat. Rev. Drug Discovery* **2**, 527–541.
16. Alfarano, C., Andrade, C. E., Anthony, K., Bahroos, N., Bajec, M., Bantoft, K., Betel, D., Bobechko, B., Boutillier, K., Burgess, E., Buzadzija, K., Caverio, R., D’Abreo, C., Donaldson, I., Dorairajoo, D., Dumontier, M. J., Dumontier, M. R., Earles, V., Farrall, R., Feldman, H., Garderman, E., Gong, Y., Gonzaga, R., Grytsan, V., Gryz, E., Gu, V., Haldorsen, E., Halupa, A., Haw, R., Hrvovic, A., Hurrell, L., Isserlin, R., Jack, F., Juma, F., Khan, A., Kon, T., Konopinsky, S., Le, V., Lee, E., Ling, S., Magidin, M., Moniakis, J., Montojo, J., Moore, S., Muskat, B., Ng, I., Paraiso, J. P., Parker, B., Pintilie, G., Pirone, R., Salama, J. J., Sgro, S., Shan, T., Shu, Y., Siew, J., Skinner, D., Snyder, K., Stasiuk, R., Strumpf, D., Tuekam, B., Tao, S., Wang, Z., White, M., Willis, R., Wolting, C., Wong, S., Wrong, A., Xin, C., Yao, R., Yates, B., Zhang, S., Zheng, K., Pawson, T., Ouellette, B. F., and Hogue, C. W. (2005) The Biomolecular Interaction Network Database and related tools 2005 update, *Nucleic Acids Res.* **33**, D418–D424.
17. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* **13**, 2498–2504.
18. Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M., and Cantley, L. C. (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains, *Science* **275**, 73–77.
19. Fuh, G., Pisabarro, M. T., Li, Y., Quan, C., Lasky, L. A., and Sidhu, S. S. (2000) Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display, *J. Biol. Chem.* **275**, 21486–21491.
20. Wiedemann, U., Boisguerin, P., Leben, R., Leitner, D., Krause, G., Moelling, K., Volkmer-Engert, R., and Oschkinat, H. (2004) Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides, *J. Mol. Biol.* **343**, 703–718.
21. Fam, S. R., Paquet, M., Castleberry, A. M., Oller, H., Lee, C. J., Traynelis, S. F., Smith, Y., Yun, C. C., and Hall, R. A. (2005) P2Y1 receptor signaling is controlled by interaction with the PDZ scaffold NHERF-2, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8042–8047.

Autophagy and Neurodegeneration

Congcong He and Daniel J. Klionsky*

Life Sciences Institute, and Departments of Molecular, Cellular and Developmental Biology and Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109

Removal of unwanted proteins is crucial for the maintenance of cellular homeostasis, especially in postmitotic cells, such as neurons, since the accumulation of abnormal proteins will not be diluted by cell division. Indeed, misfolded protein aggregation characterizes most neurodegenerative diseases (1). Eukaryotic cells primarily utilize two mechanisms for protein clearance, the ubiquitin–proteasome pathway and the autophagy–lysosome pathway. Compared with the proteasome, which recognizes and degrades protein substrates conjugated to a polyubiquitin chain, autophagy is a less selective clearance mechanism (Figure 1). Autophagy is induced under conditions of physiological stress, such as starvation. Autophagy also occurs at a basal level in normal conditions. Yet, little is known about the physiological importance of constitutive autophagy in nondividing cells. Two recent papers published in *Nature* by Hara *et al.* (2) and Komatsu *et al.* (3) proposed a neuroprotective role of basal autophagy, based on their striking findings that mice with autophagy defects specifically in neurons formed intracellular inclusions in the brain and developed neurodegenerative symptoms, even without the expression of any disease-causing, aggregate-prone proteins.

The autophagic process requires the concerted coordination of proteins encoded by *Atg* (autophagy-related) genes (4). The ubiquitous knockout of *Atg5* or *Atg7*, two essential genes for autophagy, causes early postnatal lethality in mice. Accordingly, Hara *et al.* (2) and Komatsu *et al.* (3) used a

Cre-loxP system to disrupt *Atg5* and *Atg7*, respectively, in the mouse central nervous system (CNS) by expressing Cre recombinase driven by a CNS-specific promoter. This approach allowed the authors to investigate the role of autophagy in the CNS in adult animals and study long-term neurodegenerative progression. The neuron-specific, autophagy-deficient mice (referred to as *Atg5^{-/-}* and *Atg7^{-/-}*) were born normally but developed progressive behavioral defects resembling those of aging-related neurodegeneration in humans and mouse models, for example, limb-clasping, tremor, and ataxic walking pattern. The behavioral phenotypes were accompanied by mass death of Purkinje cells and the presence of ubiquitin-positive inclusion bodies in certain regions of the brain including the cerebral cortex, cerebellum, and hypothalamus, which is thought to be the pathological hallmark of neurodegeneration. In addition, the autophagy-deficient mice had a dramatically decreased survival rate; most animals died within 28 weeks. In *Atg7^{-/-}* mice, polyubiquitinated proteins accumulated and aggregated into inclusions despite apparently normal proteasome function, which suggests that basal autophagy may be essential in clearing some misfolded proteins that are beyond the degradative capability of the proteasome. Thus, basal autophagy may be of more biological importance than was previously recognized.

If defects in autophagy alone are sufficient to elicit protein aggregation and degeneration in normal neurons

ABSTRACT Autophagy mediates the bulk degradation of cytosolic proteins and organelles. Recent studies using neuron-specific knockout mouse models demonstrate that autophagy deficiency leads to protein aggregation and neurodegeneration, even in the absence of disease-related aggregate-prone proteins.

*To whom correspondence should be addressed.
E-mail: klionsky@umich.edu.

Published online May 19, 2006
10.1021/cb600182h CCC: \$33.50
© 2006 by American Chemical Society

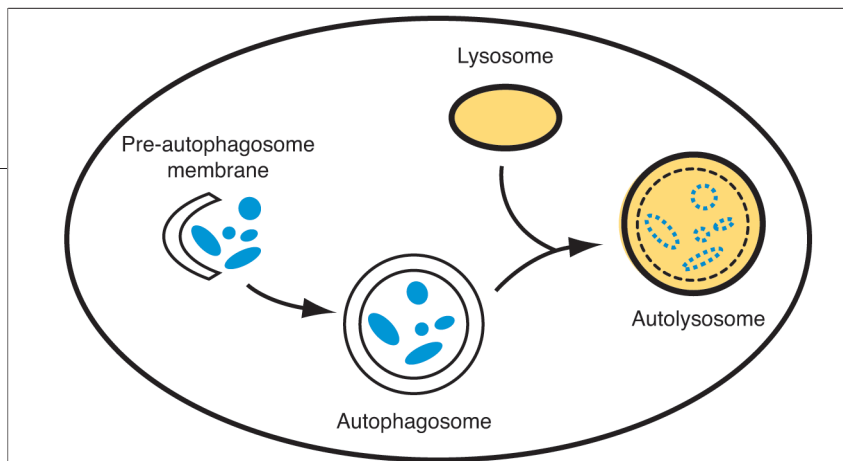


Figure 1. Schematic representation of autophagy in a mammalian cell. A portion of cytosol destined for degradation is sequestered in a double-membrane vesicle termed an autophagosome, by elongation of a small pre-autophagosome membrane structure that derives from an unknown origin. The autophagosome then fuses with the lysosome to form an autolysosome. The inner vesicle is degraded by resident hydrolases together with its cargoes.

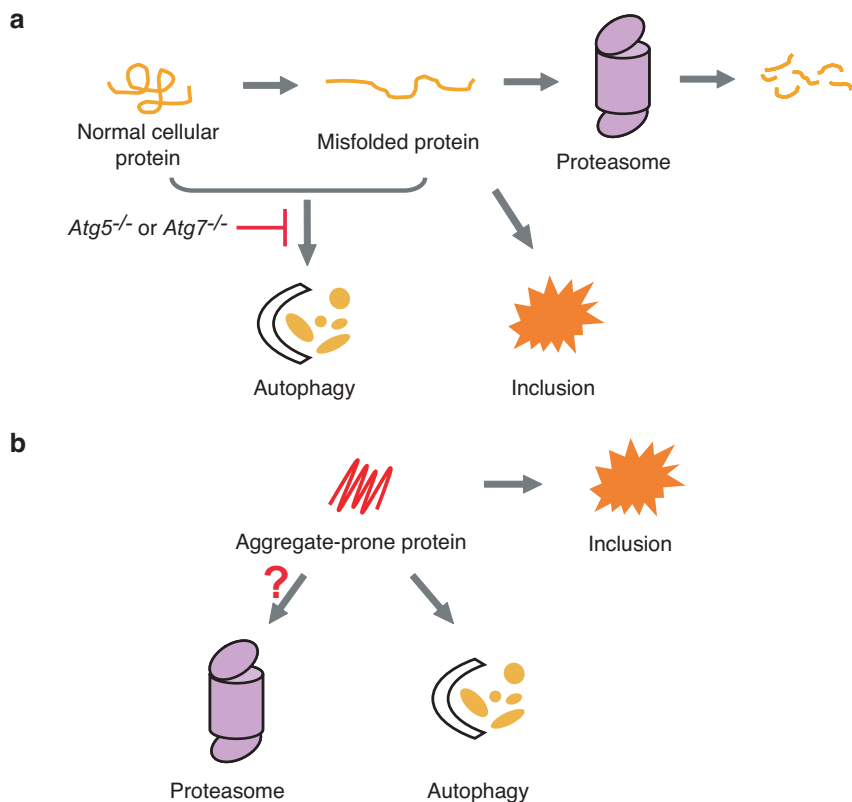
(Figure 2, panel a), it is interesting to consider cells that are burdened with aggregate-prone proteins related to human diseases, for example, α -synuclein and expanded polyglutamine proteins. In a line of inducible transgenic mice, neurodegeneration could be promisingly reversed when

the expression of the harmful protein is shut off (5). This observation implies a protective role of protein clearance. Increasing evidence indicates autophagy is actively involved in this process (6, 7). For example, chemical inhibition of autophagy at the autophagosome formation or autophago-

some-lysosome fusion stages is associated with enhanced protein aggregation and death in cell models (8). Conversely, rapamycin, a chemical inducer of autophagy, reduces aggregate formation and improves behavioral tasks in transgenic animal models (9, 10).

Recent studies suggest that the cytosolic early protein species (monomer or oligomer) in the aggregation process, rather than the inclusion itself, are the main source of toxicity (1, 11). In *Atg5^{-/-}* neurons and hepatocytes, inclusions formed after diffuse-ubiquitinated proteins had accumulated to a high level. This observation supports the idea that inclusions may be a consequence rather than the cause of pathogenesis. The other intriguing hypothesis that stems from this finding is that autophagy may protect against neurode-

Figure 2. Hypothetical models of clearance mechanisms of normal proteins and disease-related aggregate-prone proteins. a) In normal neurons, misfolded proteins are ubiquitinated and actively degraded by proteasomes. The basal level of autophagy also constitutively degrades excess or misfolded proteins, maintaining a favorable environment for various cellular functions. In the neurons of *Atg5^{-/-}* or *Atg7^{-/-}* mice where autophagy is blocked, an elevated level of misfolded proteins accumulates in the cell and forms inclusions, leading to neurodegeneration and cell death. b) In patients or transgenic animals with neurodegenerative disorders, aggregate-prone proteins, for example, expanded polyglutamine proteins, express and adopt aberrant structures that easily aggregate into inclusions that are highly resistant to proteases. Autophagy may be a primary mechanism by which cells clear diffuse monomeric or oligomeric forms of aggregate-prone proteins, although a role for the proteasome in clearance cannot be ruled out.



generation by reducing the level of potentially toxic diffuse protein species rather than the inclusions per se (Figure 2, panel b). In fact, the size of inclusions usually exceeds that of typical mammalian autophagosomes.

There is some controversy regarding the ability of the proteasome to degrade these amyloidogenic proteins (12–14). In some diseases, the proteasomal function is significantly impaired in the presence of aggregate-prone proteins (15, 16). In addition, it has been found that proteasomes are unable to cleave within expanded polyglutamine sequences (17), which may result in steric hindrance and inhibition of the proteasome.

Despite recent progress, there are many open questions to be answered with regard to misfolded protein clearance and neurodegeneration. For example, it is not clear if, or how, the proteasome and autophagy coordinate to accomplish the task of protein quality control. In addition, we do not know exactly the forms of aggregate-prone proteins (diffuse or inclusion) degraded through autophagy, or the molecular components and mechanism of autophagic degradation of these proteins. Since it is suggested that toxicity has been initiated before aggregates are formed (18), it will be helpful to have additional data that identify the factors/pathways that regulate the clearance of abnormal proteins before or during early stages of the aggregation process, as these may serve as potential therapeutic targets in neurodegenerative disorders. The studies by Hara *et al.* (2) and Komatsu *et al.* (3) provide a direct causative connection between impaired autophagy and neurodegeneration in the absence of genetic

mutations that predispose an individual to disease and, thus, unveil a critical role of basal autophagic clearance. Neurons may be especially sensitive to proteolytic stress due to their quiescent state and highly specific function. The neuron-specific, autophagy-deficient mouse models will be useful in further investigating the role of autophagy in neurological diseases, specifically, in protection against misfolded protein accumulation and aggregation.

Acknowledgment: This article is dedicated to the memory of Cecile M. Pickart.

REFERENCES

- Ross, C. A., and Poirier, M. A. (2004) Protein aggregation and neurodegenerative disease. *Nat. Med.* **10** (Suppl.), S10–S17.
- Hara, T., Nakamura, K., Matsui, M., Yamamoto, A., Nakahara, Y., Suzuki-Migishima, R., Yokoyama, M., Mishima, K., Saito, I., Okano, H., and Mizushima, N. (2006) Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature*, published online April 19, <http://dx.doi.org/10.1038/nature04724>.
- Komatsu, M., Waguri, S., Chiba, T., Murata, S., Iwata, J. I., Tanida, I., Ueno, T., Koike, M., Uchiyama, Y., Kominami, E., and Tanaka, K. (2006) Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature*, published online April 19, <http://dx.doi.org/10.1038/nature04723>.
- Levine, B., and Klionsky, D. J. (2004) Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev. Cell* **6**, 463–477.
- Yamamoto, A., Lucas, J. J., and Hen, R. (2000) Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell* **101**, 57–66.
- Bjorkoy, G., Lamark, T., Brech, A., Outzen, H., Perander, M., Overvatn, A., Stenmark, H., and Johansen, T. (2005) p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death. *J. Cell Biol.* **171**, 603–614.
- Qin, Z. H., Wang, Y., Kegel, K. B., Kazantsev, A., Apostol, B. L., Thompson, L. M., Yoder, J., Aronin, N., and DiFiglia, M. (2003) Autophagy regulates the processing of amino terminal huntingtin fragments. *Hum. Mol. Genet.* **12**, 3231–3244.
- Ravikumar, B., Duden, R., and Rubinsztein, D. C. (2002) Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy. *Hum. Mol. Genet.* **11**, 1107–1117.
- Ravikumar, B., Vacher, C., Berger, Z., Davies, J. E., Luo, S., Oroz, L. G., Scaravilli, F., Easton, D. F., Duden, R., O'Kane, C. J., and Rubinsztein, D. C. (2004) Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat. Genet.* **36**, 585–595.
- Berger, Z., Ravikumar, B., Menzies, F. M., Oroz, L. G., Underwood, B. R., Pangalos, M. N., Schmitt, I., Wullner, U., Evert, B. O., O'Kane, C. J., and Rubinsztein, D. C. (2006) Rapamycin alleviates toxicity of different aggregate-prone proteins. *Hum. Mol. Genet.* **15**, 433–442.
- Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810.
- Webb, J. L., Ravikumar, B., Atkins, J., Skepper, J. N., and Rubinsztein, D. C. (2003) α -synuclein is degraded by both autophagy and the proteasome. *J. Biol. Chem.* **278**, 25009–25013.
- Ross, C. A., and Pickart, C. M. (2004) The ubiquitin-proteasome pathway in Parkinson's disease and other neurodegenerative diseases. *Trends Cell Biol.* **14**, 703–711.
- Martin-Aparicio, E., Yamamoto, A., Hernandez, F., Hen, R., Avila, J., and Lucas, J. J. (2001) Proteasomal-dependent aggregate reversal and absence of cell death in a conditional mouse model of Huntington's disease. *J. Neurosci.* **21**, 8772–8781.
- Bence, N. F., Sampat, R. M., and Kopito, R. R. (2001) Impairment of the ubiquitin-proteasome system by protein aggregation. *Science* **292**, 1552–1555.
- Bennett, E. J., Bence, N. F., Jayakumar, R., and Kopito, R. R. (2005) Global impairment of the ubiquitin-proteasome system by nuclear or cytoplasmic protein aggregates precedes inclusion body formation. *Mol. Cell* **17**, 351–365.
- Venkatraman, P., Wetzel, R., Tanaka, M., Nukina, N., and Goldberg, A. L. (2004) Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. *Mol. Cell* **14**, 95–104.
- Ross, C. A., and Poirier, M. A. (2005) Opinion: What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol.* **6**, 891–898.

Pin1 Flips Alzheimer's Switch

Felicia A. Etzkorn*

Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061

ABSTRACT The biochemical processes leading to Alzheimer's disease are just now being elucidated. A recent study shows that a peptidyl-prolyl isomerase, Pin1, specifically regulates the degradation of amyloid precursor protein (APP). An alternative model for Pin1 regulation of APP processing is also proposed.

Alzheimer's disease strikes fear in the hearts (or more appropriately, in the *minds*) of scientists and intellectuals, perhaps more so than heart disease or cancer, because it destroys the brain before it kills. Despite intense research in the past 30 years, changes in biochemical regulation leading to the development of Alzheimer's have only begun to yield to modern techniques. In a recent issue of *Nature*, Lucia Pastorino, Anyang Sun, and co-workers (1) report that the peptidyl-prolyl isomerase (PPIase), Pin1, specifically regulates the amyloid precursor protein (APP) degradation product ratio. In the presence of Pin1, less of the bad actor amyloid peptide A β 42 is produced and more of the presumably benign α -APPs are produced. In the absence of Pin1, significantly more A β 42 is produced. A β 42 is known as the precursor for amyloid plaque formation. Pin1 has been hypothesized to protect against neurodegeneration in Alzheimer's diseased brains (2). Pin1 restores the function of phosphorylated Tau protein, also significant in Alzheimer's disease, by binding specifically to pThr231-Pro (2).

As the only known phosphorylation-dependent PPIase in humans (Figure 1), and the only human PPIase with a WW domain (3, 4), Pin1 is uniquely situated to regulate a number of dynamic cellular signaling processes (5). Pin1 specifically catalyzes isomerization of phosphoThr- or phosphoSer-Pro motifs in a number of cell cycle proteins (6).

PPIases are capable only of restoring the equilibrium between cis and trans Xaa-Pro amide conformations; they do not make or

break bonds. This equilibration is accelerated by PPIases to the tune of 10^6 , while a significant rate of isomerization takes place thermally, *i.e.*, at ambient or body temperatures (7). These data have led many to question the significance of PPIase's endogenous roles in biology. The other families of PPIases, the cyclophilins and FKBP, are best known for their role in activating the immunosuppressive prodrugs cyclosporine A and FK506 (8), while their endogenous roles are probably as catalysts of protein folding (9). As a cell cycle regulator, Pin1 differs from the kinases, phosphatases, histone acetyl transferases, and histone deacetylases precisely in that it does not make or break bonds. Pin1 is unique because it regulates the cell cycle by a conformational switch mechanism (10).

The significance of this role, and Pin1 does appear to be important in biological processes including mitosis (5), must arise from the need to reestablish chemical equilibria rapidly after the equilibrium is upset by another process. Thus, phosphorylation or dephosphorylation could apply such pressure to one or the other side of the equilibrium by LeChatelier's principle. Another possibility is that binding phenomena could "take out" one of the substrates, cis or trans, upsetting the equilibrium and calling Pin1 into play. Pastorino *et al.* have identified just such a substrate relevant to Alzheimer's disease.

Pin1 was shown to bind APP from mitotic cells through the single pThr668-Pro669 site by immunoprecipitation (1). This binding was mediated by the WW domain, as demonstrated by ^{15}N - ^1H correlated NMR

*To whom correspondence should be addressed.
E-mail: fetzkorn@vt.edu.

Published online May 19, 2006
10.1021/cb600171g CCC: \$33.50
© 2006 by American Chemical Society

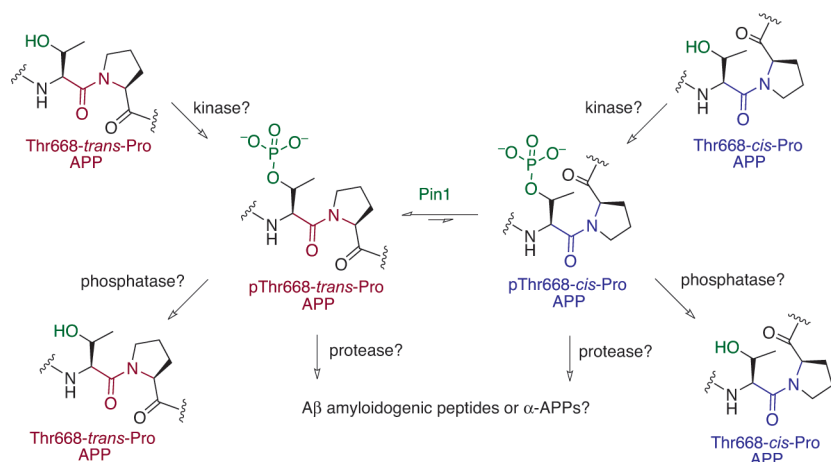


Figure 1. Alternative model to explain the central role of Pin1 in processing APP based on equilibrium upset by a kinase or phosphatase. APP must rapidly cross the central equilibration catalyzed by Pin1 to avoid processing to amyloidogenic products. This model assumes a conformation-specific kinase or phosphatase to upset the conformational equilibrium. Whether the left (magenta) or the right (blue) side of the scheme leads to amyloidogenic processing remains to be seen.

(HSQC) on a 21-residue phosphopeptide (G659–Q679). The peptide was specifically labeled with ^{15}N at Glu670, just after the Pro. The WW domain bound cis and trans conformers of the peptide in the same ratio that exists free in solution. These results differ significantly from X-ray studies that show only the trans conformation of a Cdc25 peptide bound to the WW domain (4). It may be that only the trans conformation of the Cdc25 peptide binds to the WW domain, while the APP peptide can bind in both cis and trans forms. Or perhaps the crystal packing forces permitted only the trans form of Cdc25 to be seen in complex with the WW domain. Solution phase NMR may allow both forms to be detected. Yet, the binding phenomenon was not the only interaction of Pin1 with APP observed.

The APP peptide is isomerized by the catalytic site of Pin1 specifically at pThr668-Pro (Figure 1) as measured by ROESY of the cis–trans exchange cross peaks at two mixing times (1). The rate of cis to trans isomerization was measured to be 5- and 15-fold faster than the rate in the trans to cis direction, depending upon the cross peak integrated. (The uncertainty may be in the integration of the trans–trans peak on the diagonal, which is not well resolved.) Most significantly, the Pin1 catalyzed rate is on the order of 10^5 -fold faster than the

thermal isomerization rate. Pin1 isomerization of an Alzheimer's disease relevant substrate has clearly been demonstrated (although "atomic resolution" is debatable since a single ^{15}N was measured). Moreover, both APP substrate binding to the WW domain and isomerization by the catalytic domain have been established, though not necessarily correlated. An intriguing question remains, which comes first, binding or catalysis?

Cellular and *in vivo* results cinch the deal. Pin1 was found co-localized with APP, and they coprecipitated with a Pin1-specific antibody (1). Pin1 and APP co-localize at the plasma membrane and in clathrin-coated intracellular vesicles. Overexpression of Pin1 reduces A β secretion in a phospho-Thr668 dependent manner. Pin1 depleted cells (asynchronous CHO-APP and Pin1 $^{-/-}$ H4 breast cancer) show increased APP processing and A β secretion, while mitotic CHO-APP cells expressing Pin1 and Pin1 $^{+/+}$ H4 cells show ~ 3 -fold higher secretion of α -APPs and ~ 7 -fold less A β secretion. Pin1 knockout mice showed increased levels of insoluble A β peptides, particularly of the major toxic peptide A β 42, in an age-dependent manner leading to plaque formation. Tissue culture of mice brains with immunogold electron microscopy with anti-human A β 42 antibodies showed A β 42

localized in multivesicular bodies in dorsal medial cortical neurons. In mice overexpressing familial Alzheimer's disease (FAD), the effects of Pin1 were also age dependent.

On the basis of their results, the authors have proposed a model for Pin1 regulation of APP processing and prevention of Alzheimer's disease. In their model (Figure 5, panels g and h of ref 1), Pin1 activity decreases the levels of amyloidogenic A β peptides by decreasing the concentration of pThr668-cis-Pro APP (1). In the absence of active Pin1, concentrations of cis APP are higher resulting in high levels of amyloidogenic A β peptides. There is little in their results to presume that the cis form of APP is the bad actor however. The relative rates of cis–trans vs trans–cis isomerization are not so very different (5- to 10-fold) compared with the Pin1 acceleration over the uncatalyzed rate (10^5 -fold). Again, Pin1 can act only to reestablish an equilibrium that has been upset from outside. The relative rates of the two isomerizations tell us nothing about which side of the cis–trans equilibrium is overstocked at any given time. The real significance of these results is in the difference between the effect of the presence or absence of Pin1 on APP processing and the identification of pThr668-Pro as the exact site of substrate isomerization.

The key to unlocking this story may lie upstream or downstream of Pin1. What upsets the equilibrium in the first place? If APP processing depletes the concentration of the cis (or the trans) form, resulting specifically in amyloidogenic A β peptides, Pin1 would accelerate the formation of the depleted (bad) species. This does not appear to be the case based on the results of Pastorino *et al.* Another explanation must be found. I propose an alternative model for Pin1 regulation of APP processing (Figure 1). If a kinase phosphorylates only the cis (or only the trans) form of Thr668-Pro in APP, Pin1 might be essential in accelerating the reequilibration away from the bad form, whichever that is. Alternatively, a phosphatase might deplete the concentration of the bad form of phosphoThr668-Pro APP, drawing it away from amyloidogenic processing, and Pin1 could feed the phosphatase more of its “good” substrate. Conformationally locked substrates provide one promising way to distinguish between these possibilities (11).

This work by Pastorino *et al.* represents a major advance in explaining the detailed molecular regulation of amyloid precursor protein processing by the PPlase activity of Pin1. Their results set the stage for serious advancement toward the prevention and treatment of Alzheimer’s and other amyloid-based diseases.

REFERENCES

1. Pastorino, L., Sun, A., Lu, P. J., Zhou, X. Z., Balastik, M., Finn, G., Wulf, G., Lim, J., Li, S. H., Li, X., Xia, W., Nicholson, L. K., Lu, K. P. (2006) The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production, *Nature* 440, 528–534.
2. Lu, P. J., Wulf, G., Zhou, X. Z., Davies, P., Lu, K. P. (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein, *Nature* 399, 784–788.
3. Ranganathan, R., Lu, K. P., Hunter, T., Noel, J. P. (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent, *Cell* 89, 875–886.
4. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., Noel, J. P. (2000) Structural basis for phosphoserine-proline recognition by group IV WW domains, *Nat. Struct. Biol.* 7, 639–643.
5. Lu, K. P., Hanes, S. D., Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis, *Nature* 380, 544–547.
6. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., Lu, K. P. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: A potential mitotic regulatory mechanism, *Science* 278, 1957–1960.
7. Fischer, G. (1994) Peptidyl-Prolyl cis/trans Isomerases and Their Effectors, *Angew. Chem., Int. Ed. Engl.* 33, 1415–1436.
8. Etzkorn, F. A., Stolz, L. A., Chang, Z., Walsh, C. T. (1993) The role of the cyclosporin A-cyclophilin complex, FK506-FK506 binding protein and calcineurin in the inhibition of T-cell signal transduction, *Curr. Opin. Struct. Biol.* 3, 929–933.
9. Gothel, S. F., Marahiel, M. A. (1999) Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts, *Cell. Mol. Life Sci.* 55, 423–436.
10. Shen, M., Stukenberg, P. T., Kirschner, M. W., Lu, K. P. (1998) The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins, *Genes Dev.* 12, 706–720.
11. Wang, X. J., Xu, B., Mullins, A. B., Neiler, F. K., Etzkorn, F. A. (2004) Conformationally Locked Isostere of PhosphoSer-cis-Pro Inhibits Pin1 23-Fold Better than PhosphoSer-trans-Pro Isostere, *J. Am. Chem. Soc.* 126, 15533–15542.

Following the Path of the Virus: The Exploitation of Host DNA Repair Mechanisms by Retroviruses

Johanna A. Smith and René Daniel*

Division of Infectious Diseases—Center for Human Virology, Thomas Jefferson University, 1020 Locust Street, Philadelphia, Pennsylvania 19107

Viruses are obligatory parasites lacking certain biochemical components, which are necessary for self-replication (1, 2). Retroviruses are a class of enveloped viruses that store their genetic material in the form of RNA. These viruses are unique because they must convert their RNA into DNA *via* the process of reverse transcription. Infection is initiated by the recognition of viral proteins by receptors present on the surface of target cells. Once inside the infected cell, reverse transcription is catalyzed by the retroviral enzyme reverse transcriptase (RT). RT reverse transcribes a single-stranded viral RNA molecule into viral DNA (3). The DNA product of reverse transcription can then be inserted into the genome of the host during the process of integration, which depends on the retroviral enzyme integrase (IN) as well as host cellular cofactors (4). Upon the successful addition of viral DNA to the host genome, the virus can replicate. From the stably transduced host genome, transcription of viral genes ensues in order to generate polypeptides as well as two copies of the viral genome in the form of RNA. These components are exported from the nucleus to the cytoplasm and assemble into new viral particles, which are then released from the host cell (1, 2, 4).

It is important to recognize the similarities between certain chemical mechanisms that occur in retroviral infections and transposition (5). Transposable elements (or transposons) are segments of DNA that can shift around within a genome. Completion of the genome project revealed that around 45% of the human genome consists of transposable elements (6). Retrotransposons are a subset of these mobile genetic elements that spread from one location to another *via* an mRNA intermediate (6–8). The mRNA is reverse transcribed by their self-encoded RT, producing a cDNA copy. Transposon DNA is then inserted into a

ABSTRACT Numerous host cellular cofactors are involved in the life cycle of retroviruses. Importantly, DNA repair machinery of infected cells is activated by retroviruses and retroviral vectors during the process of integration and host cell DNA repair proteins are employed to create a fully integrated provirus. The full delineation of these repair mechanisms that are triggered by retroviruses also has implications outside of the field of retrovirology. It will undoubtedly be of interest to developers of gene therapy and will also further facilitate our understanding of DNA repair and cancer. This review gives a brief summary of the accomplishments in the field of DNA repair and retroviral integration and the opportunities that this area of science provides with regards to the elucidation of repair mechanisms, in the context of retroviral infection.

*To whom correspondence should be addressed.
E-mail: Rene.Daniel@jefferson.edu.

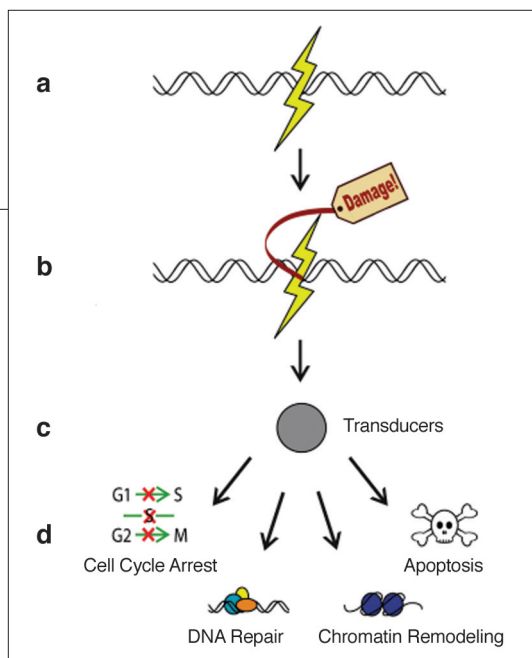
Received for review March 23, 2006
and accepted May 2, 2006

Published online May 19, 2006
10.1021/cb600131q CCC: \$33.50

© 2006 American Chemical Society

Figure 1. The DNA damage response signal transduction cascade. a) DNA damage resulting from exogenous sources (IR, UV light, chemicals) or endogenous sources (meiosis, V(D)J recombination, stalled replication forks, byproducts of cellular metabolism).

b) Sensor proteins detect damage and pass the signal along to c) transducers which relay the signals to effector proteins which execute d) the appropriate response: cell cycle arrest, DNA repair, chromatin reconstitution, or apoptosis.



new genomic location. The processes of genomic DNA cleavage, as well as the ligation reaction of the newly broken ends to the retrotransposon genetic material, are very similar to retroviral DNA integration. Retrotransposons have been reported to integrate into endogenous breaks in DNA (9) and, furthermore, to employ host DNA repair processes (10). Like retrotransposons, recent data indicate that retroviral DNA integration is completed with the aid of host cellular proteins that are involved in double-strand DNA break repair (11–17).

Cellular responses to double-strand DNA breaks

All living cells continually suffer DNA damage from a variety of exogenous as well as endogenous sources. These may result in single-strand breaks (SSBs), double-strand breaks (DSBs), or bulky DNA lesions. Internal processes which lead to DSBs, which are the most hazardous form of DNA damage, include stalled DNA replication forks and reactive byproducts of cellular metabolism (reactive oxidation species). Exogenous agents known to cause DSB include ionizing radiation, as well as some chemotherapeutic drugs (18, 19). DSBs are deleterious to the cell because if left unrepaired they are often responsible for the formation of neoplasms as well as cell death.

Organisms have evolved complex mechanisms to repair chromosomal insults when they take place to maintain high fidelity of genetic information transmission. Cells respond to damage to their genetic material by activating an elaborate DNA damage response pathway which can be divided into the following three

stages: (1) sensing of damage, (2) transduction of the information that the damage has occurred, and (3) the activation of effectors that execute the appropriate response (18, 19). This sequence of events is depicted (Figure 1). When activated, the effectors may carry out one or more of the following events: cell cycle arrest, the physical DNA repair, chromatin remodeling, and apoptosis (when damage is irreparable). It should be emphasized that Figure 1 is highly simplified, and the interactions that occur subsequent to DNA damage are orchestrated in an exceedingly complex and intertwined network.

Three members of the phosphatidylinositol 3-kinase-like (PIKK) protein family are of most importance in cellular responses to DSBs. These mammalian PIKK proteins are DNA-dependent protein kinase (DNA-PK), ataxia-telangiectasia mutated (ATM), and ataxia-telangiectasia and Rad3-related (ATR) (20, 21). The activation of these protein kinases is crucial for the execution of the phosphorylation events that follow (21, 22). DNA-PK and ATM are primarily activated in response to DSB, while ATR has traditionally been viewed to play a distinct role in the presence of SSBs, and other forms of DNA damage. DNA-PK is vital for both V(D)J recombination and DSB repair and also plays a role in the induction of apoptosis (23).

The magnitude of the importance of the functionality of these PIKK proteins is apparent by the detrimental consequences that result when their functional wild-type forms are inadequately expressed. Their significance may also be revealed by the fact that all eukaryotes contain homologues to ATM and ATR (24). Individuals carrying a homozygous mutation in the ATM gene display the severe disorder ataxia-telangiectasia. These patients have high rates of cancer, genomic instability, and cerebellar degeneration. Individuals with ataxia telangiectasia are hypersensitive to ionizing radiation (IR) and do not display the proper function of cellular checkpoints (20, 21). ATR deficiencies are embryonic lethal in mice, indicating that ATR is essential for organism survival (25, 26). Until recently there was no known disease caused by inadequacies of ATR expression. However, ATR-Seckel syndrome is now recognized as an autosomal recessive disorder caused by ATR mutations and is marked by developmental delay, microcephaly, growth and mental retardation, dwarfism, characteristic facial features, and DNA repair deficiencies (27). Animals with mutant DNA-PK are

immunodeficient, defective in DSB repair, and sensitive to ionizing radiation (28).

Cell Cycle Arrest. One outcome of the robust and multifaceted signal transduction cascade is cell cycle arrest (Figure 1). This delay of cell cycle progression is imperative in order to allow adequate time for DNA repair (29). The signaling kinases of primary concern with regards to cell cycle checkpoints are ATM and ATR.

In the event of DSB damage, inactive ATM homodimers auto-phosphorylate on Ser1981, releasing the two macromolecules into active monomers (30). ATM is recruited to the lesion site by the Mre11/Rad50/Nbs1 (MRN) sensor complex, which enables ATM to relay the damage signal to appropriate effectors (31–36). Phosphorylation of p53 by ATM may induce

any of the three cell-cycle checkpoints (G1/S, S, and or G2/M) (36). In response to SSBs, stalled replication forks, or bulky lesions, ATR is the primary inducer of checkpoint activation. ATR is recruited to damage sites by ATRIP (ATR-interacting protein) to single-stranded DNA which is coated with RPA (replication protein A) (37–40) to induce the G2/M checkpoint (38, 41). In addition, ATR is involved in DSB repair, where ATM and the MRN complex are first localized to DSB sites, and then alter the lesion to a structure that recruits the ATR-ATRIP complex (Figure 2, (37–39).

DNA Repair. Whether or not checkpoint activation has occurred, DNA repair must take place in order to maintain a cell's genetic integrity. Two mechanisms have evolved to repair DSB in mammalian cells when they take place, either during tightly regulated cellular processes or as a result of ill-fated DNA damage events.

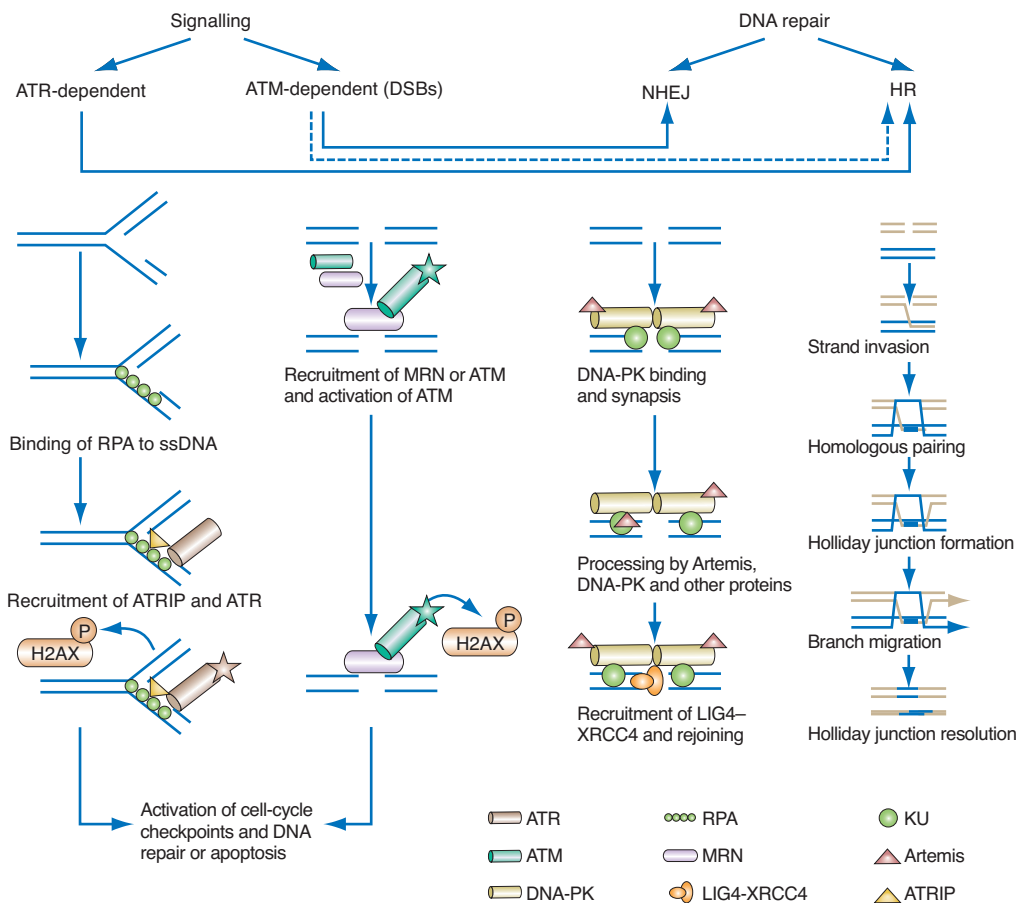


Figure 2. The roles of PIKK kinases in cellular DNA damage response. Details are given in the text.

These mechanisms are homologous recombination (HR) and nonhomologous end joining (NHEJ) (Figure 2) (42, 43). Although HR is a more faithful means of DNA repair to be conducted because the sister chromatid is used as a template, NHEJ is thought to repair the majority of DSB in mammalian cells. To begin this process, two Ku heterodimers bind to either end of the DSB (Figure 2). This heterodimer consists of the proteins Ku70 and Ku80 (44, 45). It is understood that next the DNA-PK catalytic subunit (DNA-PKcs) along with the protein Artemis is recruited to the broken ends at the sites of Ku in the form of a complex (DNA-PKcs/Artemis complex) (44, 45). Artemis is required for endonuclease activity (46–48) to function if there are any DNA overhangs to be cleaved. Once blunt ends have been created, the XRCC4/Ligase 4 complex is brought within close proximity of Ku. This second complex is

The final outcome of the DNA damage signal transduction cascade is apoptosis.

responsible for ligating the trimmed ends to complete NHEJ, although details have yet to be elucidated (45).

ATM also seems to be physically involved in DNA repair (44, 45). Its kinase activity is activated by free DNA ends *in vitro* (49) and it cosediments with damaged DNA in cell-free systems (50). In addition, ATM phosphorylates DNA repair proteins such as BRCA1, which plays a role in HR (51).

Chromatin Remodeling. Chromatin remodeling is a third possible outcome of the DNA damage response (Figure 1). This process is an important component of DNA repair, although much remains to be discovered in this realm of molecular biology. Histones are proteins, which organize the genome by compacting DNA into ordered structures (52). They also play a major role in gene expression. Dynamic modifications occur to these organizational proteins in order to render the DNA, with which they associate, accessible to other proteins when appropriate. Histone tails may be modified by methylation, acetylation, phosphorylation, ubiquitination, or ADP-ribosylation, which are all reversible reactions (53).

One well-characterized change in histone structure in association with DSB repair is H2AX phosphorylation. The histone H2AX is quickly phosphorylated on Ser139 (54) in the vicinity of 2×10^6 base pairs around the DSB (55). The principal kinase to phosphorylate H2AX at DSB sites is ATM (56, 57). The phosphorylated form of this histone is designated as γ -H2AX and is considered a marker for DSBs, where distinct large nuclear domains, or foci, appear and can be visualized by

immunostaining assays (54). Assembly of DNA repair complexes to the site of genomic insults is vital and has been thought to be conducted and/or retained by γ -H2AX (58–60).

The reversible acetylation reaction significantly affects transcriptional regulation. The role of acetylation in other DNA metabolic reactions is just beginning to be unveiled. In the budding yeast *Saccharomyces cerevisiae*, the histone H4 is acetylated by the Esa1 histone acetyltransferase (HAT) and is required for NHEJ (61). The Esa1 mammalian homologue, TIP60, is involved in both DNA repair and apoptosis. It is also of interest that the Sin3p/Rpd3p histone deacetylase (HDAC) complex modulates DSB repair

in yeast (62). These novel findings suggest that such histone alterations may serve an imperative function in the DNA repair of other eukaryotes as well (61).

In fact, the mammalian protein, TRRAP (transactivation-transformation domain-associated protein), which is a constituent of several HAT complexes, has been recently suggested to function in DSB repair. Importantly, TRRAP associates with the MRN complex, although this association did not display any acetyltransferase activity. The authors advocate that TRRAP may serve as a bridge between the processes of DSB signaling, repair, and chromatin remodeling (63). Metnase is another protein that may link chromosomal structure to DNA repair and the integration of foreign DNA. This protein methylates histone H3, causing the chromatin configuration to open therefore facilitating repair (64).

Apoptosis. The final outcome of the DNA damage signal transduction cascade (depicted in Figure 1) is apoptosis (or programmed cell death). In response to IR, DNA-PK preferentially induces apoptosis (65) by phosphorylating Ser15 of p53, a protein responsible for cell cycle succession (23). BRCA1 is phosphorylated at Ser1423 and 1524 in damage responses and, in some cases, serves as a tumor suppressor by inducing apoptosis (66). When DNA damage is too severe and/or unregulated, uncontrolled cell proliferation could arise as a result of the compromised genome. Therefore, the role of apoptosis is central in organism survival, because of its function in cancer prevention.

It has been shown recently that DSB DNA repair is involved in replication of diverse viruses and transposons, including adenoviruses, herpes simplex virus 1, SV40, and the *Sleeping Beauty* element (for recent reviews see refs 57, 67, 68, and 69–71). In addition to the above-mentioned viruses, DNA repair plays a critical role in the life cycle of retroviruses and retroviral vectors. In this article, we review recent evidence that suggests that retroviral DNA integration triggers host cellular DSB DNA repair mechanisms, which then complete the process of integration.

Integration and postintegration repair

Retroviral DNA integration begins with cleavage of a dinucleotide from 3' viral DNA ends (Figure 3, panel a). This stage is called 3'-end processing and is necessary in order to expose hydroxyl groups at the 3'-ends. The processed viral ends then ligate to host DNA during the second stage, denoted as joining, *via* a nucleophilic

KEYWORDS

Integration: Process of insertion of retroviral DNA into host cell genome.

Postintegration repair: A process that finalizes the integration of retroviral DNA into host genome.

Retrovirus: An RNA virus that employs the virus enzyme reverse transcriptase to transcribe the RNA genome into DNA, which then integrates into the host cell's genome.

Double strand DNA break: A discontinuity in DNA that involves both DNA strands.

Nonhomologous end joining: A DNA repair system that involves a large number of proteins and is the major mammalian system for repair of double-strand DNA breaks.

attack on staggered phosphodiester bonds in complementary strands of host cell DNA (1, 2, 72). Both processing and joining are catalyzed by the retroviral enzyme IN (73, 74). Host DNA is cut five base pairs apart in the case of HIV-1 (75), creating short gaps on opposite strands of complementary host DNA 5'-ends (Figure 3, panel a). These gaps between viral and genomic DNA must be filled, their corresponding ends ligated, and the chromatin structure reconstituted, in order to complete the third stage of integration, recently denoted as postintegration repair. This final stage has been shown by multiple groups to be mediated by host cellular DSB DNA repair mechanisms.

Evidence for the role of PIKK kinases in integration

In 1999, it was first observed that cells deficient in components of the NHEJ pathway display reduced rates of retroviral infection (11). It was shown that retroviral infection efficiencies of DNA-PKcs-, Ku80-, and XRCC4-deficient fibroblasts were 5–10-fold lower than that of control fibroblasts and that 40–50% of DNA-Pkcs-deficient pre-B cells underwent apoptosis following infection (11). Cell death was dependent upon the presence of the active IN in the infecting virion. (11, 13, 14). On the basis of this evidence, it has been suggested that NHEJ is necessary to complete the retroviral DNA integration, and a failure of this process results in cell death. The NHEJ requirement can be apparently extended to all retroviruses and retroviral vectors, since it was observed for such diverse vectors as avian sarcoma virus (ASV)-based vectors, murine leukemia virus (MLV), and HIV-1-based vectors.

Subsequently, it had been demonstrated that the depletion of Ku80 in human cells decreases replication and stable integration by an HIV-1 strain (76). Consistently, Waninger *et al.* demonstrated that by downregulating Ku80 by the utilization of both ribozymes (short RNA molecules that can cleave target mRNA) and siRNA (short interfering RNA used to knock down gene expression in a sequence specific fashion) techniques, HIV-1 infection is significantly decreased (77). An extensive work by the O'Connor laboratory demonstrated that NHEJ proteins, including DNA-PKcs, Ku80, Ku70, XRCC4, and ligase IV are required for efficient transduction by both HIV-1-based and MLV-based vectors (16). Finally, Downs *et al.* showed that yeast deficient in either Ku subunit display a decrease in Ty1 (a retrotransposon with self-encoded

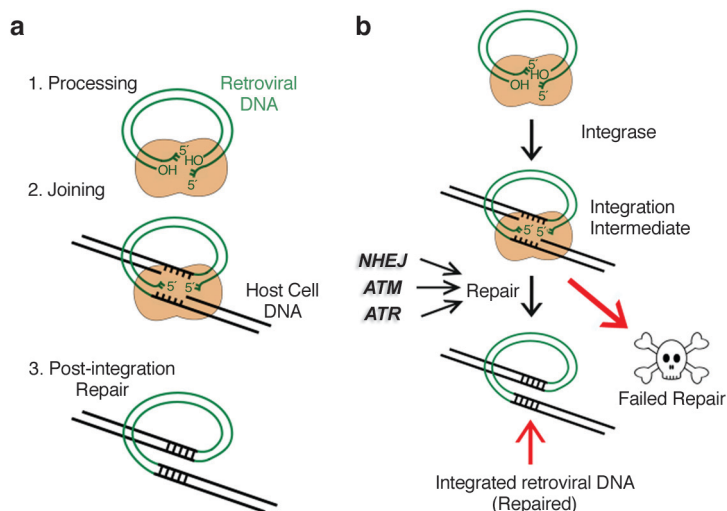


Figure 3. Integration and postintegration repair. a) Retroviral DNA integration: tan, the viral enzyme IN; green, viral DNA; black, host cell DNA. (1) Processing, IN excises a dinucleotide from 3' viral ends. (2) Joining, IN joins 3' viral ends to host cell DNA in a concerted cleavage-ligation reaction. (3) Postintegration repair, host cellular DNA repair mechanisms facilitate the repair and gap-filling to create a fully integrated provirus. b) A model for the roles of DSB repair proteins and pathways in postintegration (already defined in a). A failure of postintegration repair results in cell death, reducing the efficiency of retroviral transduction and replication.

integrase) transposition by over 80% (15). These Ty elements require Ku, the DNA binding protein of DNA-PK, to efficiently integrate. Together, these data are consistent with the hypothesis that NHEJ plays a role in integration (15).

An alternative hypothesis for the role of DNA-PK in the retroviral life cycle involves events that occur prior to integration. It was found that Ku facilitates the circularization of linear viral DNA in human cells (76, 78). Therefore it has been suggested that cell death in NHEJ-deficient cells is due to the presence of free viral DNA ends. However, upregulation of the HR repair protein, Rad52, inhibits integration and suppresses 2-LTR circle formation but does not promote apoptosis (16). Therefore, the majority of evidence supports the hypothesis that the death of NHEJ-deficient cells is due to failed postintegration repair.

Recently it has been shown by Lau *et al.* (17) that ATM-deficient cells are transduced about 5-fold lower than that of control cells when infected with pseudotyped HIV-1 based vectors. Most importantly however, in this study, the O'Connor group showed that a novel pharmacological ATM inhibitor blocks HIV-1 replication.

Although a strong link between host cellular DNA damage response proteins and retroviral integration exists, the exact mechanism(s) involved remains to be discovered.

The small molecule KU-55933, a specific and efficient inhibitor of ATM, suppresses both wild-type and drug-resistant strains of HIV-1 without cytotoxicity effects. Although it was first thought that ATM played only a minor role in integration (14), ATM thus appears to be more significant than previously believed. Consistent with Lau *et al.*, we have recently observed a substantial reduction in the percentage of ataxia telangiectasia primary fibroblast cells that are stably transduced with HIV-1 based vectors, when compared to that of control fibroblasts (Smith and Daniel, unpublished data). In addition, it was demonstrated that an ATM substrate, the histone H2AX isoform, is phosphorylated at the sites of retroviral DNA integration (79).

It should also be noted that a third PIKK kinase, ATR, has been implicated in retroviral integration. It was shown that the efficiency of ASV and HIV-1 transduction and stable integration is dramatically decreased upon expression of a dominant negative ATR mutant product (12). Finally, consistent with the observation above, we have shown that the ATM/ATR inhibitor caffeine and caffeine-related methylxanthines inhibit HIV-1 replication in primary human cells, likely at the postintegration repair step of the HIV-1 life cycle (80).

All three PIKK kinases thus appear to play a role in postintegration repair (Figure 3, panel b). The presented evidence indicates that the cell senses the integration event as a DSB, therefore triggering the cellular DNA damage response to repair the break created by the virus. This viral exploitation of host cellular DNA repair pathways is then required for the completion of integration.

Although all three of the PIKK kinases discussed have been implicated in HIV-1 integration, some labs have not observed a significant role of ATM or ATR in HIV integration (81, 82). Lau *et al.* suggests some possible explanations for these discrepancies (17). One is that studies in the field of DNA damage response in retroviral integration have been conducted using a wide variety of retroviruses and retroviral-based vectors and reporter gene readouts could be inaccurately interpreted. In addition, the inhibitors used have low specificity and could be targeting more than one of these kinases simultaneously, so that

the interpretation of data could be confused. Differing cell types could also lead to contradictory results. Lau *et al.* (17) also points out that siRNA techniques appear to alter readouts. Even control siRNA was shown to upregulate infection as in the published work by DeHart *et al.* (81), which may indicate that the conclusions drawn based on infection assays conducted on siRNA-transfected cells could be misinterpreted. One must consider these potential problems meticulously, in order to conduct precise experiments and derive accurate conclusions.

Mechanistic aspects of cellular DSB repair proteins in postintegration repair

Although a strong link between host cellular DNA damage response proteins and retroviral integration exists, the exact mechanism(s) involved remains to be discovered. Host cellular proteins could serve for one or more of the following functions in retroviral integration: nuclease, polymerase, and/or ligase activity, as well as the rearrangement of chromatin structure, and cell cycle checkpoint induction. The hypothetical outcomes of the 5'-end joining process are shown (Figure 4). Infection of cells deficient of DSB repair components could result in any of the first four outcomes, resulting in cell death, or at least the failure of postintegration repair.

At first glance, one may question the geometry of the break. It may seem counterintuitive that integration induces a DSB response, because each strand of viral DNA is joined at one end to host DNA. However, when the bigger picture is taken into consideration, it may not seem as unlikely. For example, the cuts made in host DNA may release tension from chromatin wrapped around histones at the site of the break, triggering a widespread DSB response. We do not know the activity and location of the IN protein at this stage. It may still be keeping the ends in close proximity to one another at the catalytic site of the cleavage and strand transfer reactions. Given that we do not know the kinetics of IN dispersal from the break site, it is interesting to postulate the activities that may depend upon the location of IN. IN may depart as soon as cleavage and joining are complete, exposing both viral to host DNA junctions at once. The same DNA processing/repair mechanisms could be happening at both ends at once. On the other hand, IN may release one junction before the other, leaving the vicinity exposed to repair factors at different time points. This may lead to alternative DNA repair

KEYWORDS

DNA-PK: DNA-dependent protein kinase. A serine/threonine protein kinase that is activated by double-strand DNA breaks and controls nonhomologous end joining (NHEJ).

ATM: A DNA-PK-related kinase. ATM stands for ataxia telangiectasia mutated, since its mutation is the underlying cause of ataxia telangiectasia. ATM is involved in repair of double-strand DNA breaks.

ATR: ATR stands for ATM and Rad3 related. A large serine/threonine kinase that plays a crucial role in DNA repair, including collapsed replication forks and double-strand DNA breaks.

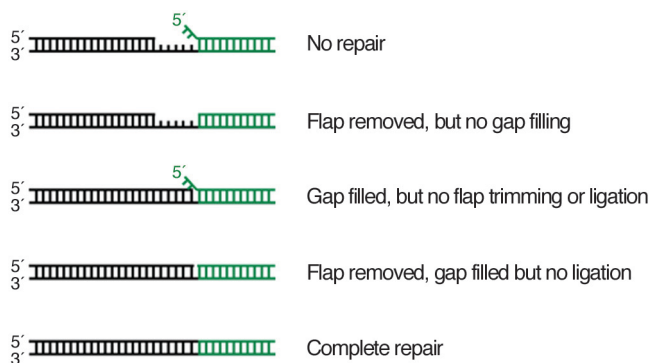


Figure 4. Hypothetical outcomes of the 5'-end joining process. Green, viral DNA; black, host cell DNA. See text for details. Deficiency of host cellular DNA repair proteins, which are required for efficient postintegration repair, may result in any of the first four outcomes. Cells containing all required DNA repair components complete postintegration repair, as depicted in the last outcome.

pathways working at either end. It is conceivable that one end could be repaired by a faster repair pathway, while at the other end, repair could be facilitated by a slower acting pathway.

One step of postintegration repair is trimming of 5' viral DNA ends (Figure 4). The DNA-PK–Artemis complex is capable of endonuclease activity once DNA-PK autophosphorylates the complex (45, 48), and this could be the nuclease expedited during postintegration repair to excise the 5'-viral overhangs.

Another step of postintegration repair is the filling of the five base pair gap. In mammals, two members of the Pol X family of polymerases, Pol λ and Pol μ , appear to be involved in NHEJ when gaps must be filled (83–85). One or both of these polymerases may be utilized in the gap-filling process.

Ligation of the newly synthesized ends to host cell 3'-ends is the final stage of postintegration repair, and this may depend on the XRCC4/Ligase 4 complex.

ATM and ATR could mediate retroviral integration by triggering cell cycle checkpoints to allow adequate time for postintegration repair. If mediating the cell cycle arrest were the role of ATM/ATR, it would seem possible that these kinases would only be required for infection of dividing cells. However, we have shown that terminally differentiated, nondividing human neurons and macrophages are susceptible to decreased infection rates in the presence of the ATM/ATR inhibitor caffeine (86). We have also shown that growth-arrested cells expressing a dominant mutant ATR exhibit

decreased transduction efficiency by retroviral vectors, suggesting that the role of ATR/ATM is not mediated by regulation of cell cycle checkpoints (86). Therefore, it appears that ATM and ATR play a more direct role in postintegration repair, possibly by phosphorylation and activation of other DNA repair proteins at the integration sites.

Changes in chromatin structure may activate DSB repair (87) as mentioned previously. Histone H2AX is phosphorylated at sites of retroviral integration, but this response is not essential for retroviral transduction (79). It had been suggested previously that γ -H2AX could play an anchoring role, by holding broken DNA ends together (88). This proposed role

for γ -H2AX is consistent with our finding that H2AX phosphorylation is not necessary for efficient postintegration repair (see above). In the context of integration, it is feasible that this anchoring function would be superfluous, because host DNA ends are tethered together *via* the viral DNA in between. One could also speculate that γ -H2AX could have a primary role in the recruitment of a fast acting DNA repair pathway, and in its absence, another, slower acting pathway could assume the responsibility of repair.

Ultimately, the study of retroviral DNA integration will lead to a further understanding of the DNA damage response to DSBs, which will have ramifications in the fields of DNA damage, oncology, and gene therapy, in addition to future therapeutic advances in the fight against AIDS.

New opportunities

DNA integration offers a unique context from which to study DNA repair and chromosomal structure. Retroviral infection offers a mild means to introduce a potentially detrimental DSB at random locations within a genome. The “dose” of breaks introduced to a particular cell population may be highly controlled, by exposing the cells to specific viral concentrations. This gentle induction of DSB may give us deeper insights into what happens when eukaryotic DSBs are encountered.

This system is also unique because a viral tag is provided at the site of the randomly chosen location of the DSB. Because we know the sequence of the

virus/vector that is utilized, molecular biology techniques can be used to find the site of the break and isolate the surrounding cellular proteins, which are involved in the repair and restructuring of the lesion. Chromatin immunoprecipitation (ChIP) techniques have been designed for one such function. By using antibodies to specific DNA repair proteins and/or chromosomal structural proteins, we can identify the machinery that encompasses the integration site. By collection of the nuclear material at different time points postinfection, more may be elucidated about the DNA repair components that are involved as well as their kinetics of action. Much remains to be discovered about how exogenous DNA is incorporated into chromatin. For example a question that could be answered with ChIP techniques is how does this “naked” viral DNA become incorporated with the host cell chromatin structure? Does the viral DNA associate with certain histones in a particular order? What are the kinetics of this association? Is there any identifiable tag that exists on structural proteins of exogenous DNA that does not exist on genomic DNA? Are other DNA repair systems, such as base excision repair or nucleotide excision repair, involved in postintegration repair? The delineation of the integration process and the reorganization of chromosomal structure could lead us to some clues about how to locate latently infected cells, which are a major problem in treatment of HIV-1 infection.

This system also provides some distinct opportunities for the improvement of gene therapy. If we understand the components involved in increasing integration efficiency of retroviral-based vectors, we may come across methods to employ in gene therapy. Efficiency of gene delivery is one of the major problems with gene therapy. Knowing the mechanisms involved

in the integration process may enable us to increase the efficiency of stable integration, thereby increasing the efficacy of gene delivery into target cells. It should also be noted that independently created breaks in the host genome may influence selection sites of retroviral integration. Some retrotransposons have been reported to selectively integrate at DSB sites, bypassing their self-encoded mechanisms of producing breaks in host cell DNA (9, 10, 89). Metnase may play a role in this process because it facilitates integration of exogenous DNA into host cell DNA *via* exposing chromatin into more accessible structures and promoting NHEJ repair (64). Therefore, increased understanding of the role of DNA repair in retroviral integration and postintegration repair might also facilitate our abilities to target retroviral vectors to specific sites in the host cell genome.

The most obvious opportunity for scientific advancement in this field of study is a therapeutic advance in HIV treatment. Viral mutation rates cause a great deal of concern in the fight against AIDS. Viral proteins can quickly adapt to survive in the presence of what was initially a potent drug, targeting reverse transcriptase or protease. Because of the capability of the virus to mutate at high rates, HIV-1 infection has yet to be cured. A conceivable way to combat HIV-1 more effectively is to target host cellular proteins, which are not susceptible to the high mutation rates. Inhibitors targeting host cellular DNA repair factors, which are involved in postintegration repair, may thus increase our ability to control the HIV-1 infection.

Acknowledgment: The authors are grateful to E. Argyris, A. Maksymowych, and G. Nunnari of Thomas Jefferson University for reading the manuscript and critical comments. The authors also wish to thank L. E. X. Peters for his assistance with figure construction. This work has been supported by the USPHS Grant CA98090 and internal university funds to R.D.

REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Plainview, NY.
2. Flint, S. J., Enquist, L. W., Krug, R. M., Racaniello, V. R., and Skalka, A. M. (2000) *Principles of Virology. Molecular Biology, Pathogenesis, and Control*, ASM Press, Washington, DC.
3. Fuentes, G. M., Fay, P. J., and Bambara, R. A. (1996) Relationship between plus strand DNA synthesis removal of downstream segments of RNA by human immunodeficiency virus, murine leukemia virus and avian myeloblastoma virus reverse transcriptases, *Nucleic Acids Res.* **24**, 1719–1726.
4. Craigie, R. (2001) HIV integrase, a brief overview from chemistry to therapeutics, *J. Biol. Chem.* **276**, 23213–23216.
5. Roth, D. B., and Craig, N. L. (1998) VDJ recombination: a transposase goes to work, *Cell* **94**, 411–414.
6. Ostertag, E. M., and Kazazian, H. H., Jr. (2001) Biology of mammalian L1 retrotransposons, *Annu. Rev. Genet.* **35**, 501–538.
7. Prak, E. T., Dodson, A. W., Farkash, E. A., and Kazazian, H. H., Jr. (2003) Tracking an embryonic L1 retrotransposition event, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1832–1837.
8. Bestor, T. H. (2005) Transposons reanimated in mice, *Cell* **122**, 322–325.
9. El-Sawy, M., Kale, S. P., Dugan, C., Nguyen, T. Q., Belancio, V., Bruch, H., Roy-Engel, A. M., and Deininger, P. L. (2005) Nickel stimulates L1 retrotransposition by a post-transcriptional mechanism, *J. Mol. Biol.* **354**, 246–257.
10. Fujimoto, H., Hirukawa, Y., Tani, H., Matsuura, Y., Hashido, K., Tsuchida, K., Takada, N., Kobayashi, M., and Maekawa, H. (2004) Integration of the 5' end of the retrotransposon, R2Bm, can be complemented by homologous recombination, *Nucleic Acids Res.* **32**, 1555–1565.

11. Daniel, R., Katz, R. A., and Skalka, A. M. (1999) A role for DNA-PK in retroviral DNA integration, *Science* 284, 644–647.
12. Daniel, R., Kao, G., Taganov, K., Greger, J. G., Favorova, O., Merkel, G., Yen, T. J., Katz, R. A., and Skalka, A. M. (2003) Evidence that the retroviral DNA integration process triggers an ATR-dependent DNA damage response, *Proc. Natl. Acad. Sci. U.S.A.* 100, 4778–4783.
13. Daniel, R., Greger, J. G., Katz, R. A., Taganov, K. D., Wu, X., Kappes, J. C., and Skalka, A. M. (2004) Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the nonhomologous end joining repair pathway, *J. Virol.* 78, 8573–8581.
14. Daniel, R., Katz, R. A., Merkel, G., Hittle, J. C., Yen, T. J., and Skalka, A. M. (2001) Wortmannin potentiates integrase-mediated killing of lymphocytes and reduces the efficiency of stable transduction by retroviruses, *Mol. Cell. Biol.* 21, 1164–1172.
15. Downs, J. A., and Jackson, S. P. (1999) Involvement of DNA end-binding protein Ku in Ty element retrotransposition, *Mol. Cell. Biol.* 19, 6260–6268.
16. Lau, A., Kanaar, R., Jackson, S. P., and O'Connor, M. J. (2004) Suppression of retroviral infection by the RAD52 DNA repair protein, *EMBO J.* 23, 3421–3429.
17. Lau, A., Swinbank, K. M., Ahmed, P. S., Taylor, D. L., Jackson, S. P., Smith, G. C., and O'Connor, M. J. (2005) Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase, *Nat. Cell Biol.* 7, 493–500.
18. Hoeijmakers, J. H. (2001) Genome maintenance mechanisms for preventing cancer, *Nature* 411, 366–374.
19. Featherstone, C., and Jackson, S. P. (1999) DNA double-strand break repair, *Curr. Biol.* 9, R759–R761.
20. Shiloh, Y. (2001) ATM and ATR: networking cellular responses to DNA damage, *Curr. Opin. Genet. Dev.* 11, 71–77.
21. Durocher, D., and Jackson, S. P. (2001) DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr. Opin. Cell Biol.* 13, 225–231.
22. Falck, J., Coates, J., and Jackson, S. P. (2005) Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage, *Nature* 434, 605–611.
23. Burma, S., and Chen, D. J. (2004) Role of DNA-PK in the cellular response to DNA double-strand breaks, *DNA Repair* 3, 909–918.
24. Abraham, R. T. (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases, *Genes Dev.* 15, 2177–2196.
25. Brown, E. J., and Baltimore, D. (2000) ATR disruption leads to chromosomal fragmentation and early embryonic lethality, *Genes Dev.* 14, 397–402.
26. de Klein, A., Muijtjens, M., van Os, R., Verhoeven, Y., Smit, B., Carr, A. M., Lehmann, A. R., and Hoeijmakers, J. H. (2000) Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice, *Curr. Biol.* 10, 479–482.
27. O'Driscoll, M., Ruiz-Perez, V. L., Woods, C. G., Jeggo, P. A., and Goodship, J. A. (2003) A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome, *Nat. Genet.* 33, 497–501.
28. Collis, S. J., DeWeese, T. L., Jeggo, P. A., and Parker, A. R. (2005) The life and death of DNA-PK, *Oncogene* 24, 949–961.
29. Niida, H., and Nakanishi, M. (2006) DNA damage checkpoints in mammals, *Mutagenesis* 21, 3–9.
30. Bakkenist, C. J., and Kastan, M. B. (2003) DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation, *Nature* 421, 499–506.
31. Carson, C. T., Schwartz, R. A., Stracker, T. H., Liley, C. E., Lee, D. V., and Weitzman, M. D. (2003) The Mre11 complex is required for ATM activation and the G2/M checkpoint, *EMBO J.* 22, 6610–6620.
32. Lee, J. H., and Paull, T. T. (2004) Direct activation of the ATM protein kinase by the Mre11/Rad50/Nbs1 complex, *Science* 304, 93–96.
33. Lee, J. H., and Paull, T. T. (2005) ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex, *Science* 308, 551–554.
34. Uziel, T., Lerenthal, Y., Moyal, L., Andegeko, Y., Mittelman, L., and Shiloh, Y. (2003) Requirement of the MRN complex for ATM activation by DNA damage, *EMBO J.* 22, 5612–5621.
35. Kitagawa, R., Bakkenist, C. J., McKinnon, P. J., and Kastan, M. B. (2004) Phosphorylation of SMC1 is a critical downstream event in the ATM-NBS1-BCR1 pathway, *Genes Dev.* 18, 1423–1438.
36. Bartek, J., Falck, J., and Lukas, J. (2001) CHK2 kinase—a busy messenger, *Nat. Rev. Mol. Cell Biol.* 2, 877–886.
37. Zou, L., and Elledge, S. J. (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes, *Science* 300, 1542–1548.
38. Cortez, D., Guntuku, S., Qin, J., and Elledge, S. J. (2001) ATR and ATRIP: partners in checkpoint signaling, *Science* 294, 1713–1716.
39. Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G. C., Lukas, J., and Jackson, S. P. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks, *Nat. Cell Biol.* 8, 37–45.
40. Namiki, Y., and Zou, L. (2006) ATRIP associates with replication protein A-coated ssDNA through multiple interactions, *Proc. Natl. Acad. Sci. U.S.A.* 103, 580–585.
41. Zhao, H., and Piwnicka-Worms, H. (2001) ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1, *Mol. Cell. Biol.* 21, 4129–4139.
42. Cahill, D., Connor, B., and Carney, J. P. (2006) Mechanisms of eukaryotic DNA double strand break repair, *Front. Biosci.* 11, 1958–1976.
43. Pastwa, E., and Blasiak, J. (2003) Non-homologous DNA end joining, *Acta Biochim. Pol.* 50, 891–908.
44. Takata, M., Sasaki, M. S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A., and Takeda, S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells, *EMBO J.* 17, 5497–5508.
45. Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2004) The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination, *DNA Repair* 3, 817–826.
46. Pannicke, U., Ma, Y., Hopfner, K. P., Niewolik, D., Lieber, M. R., and Schwarz, K. (2004) Functional and biochemical dissection of the structure-specific nuclease ARTEMIS, *EMBO J.* 23, 1987–1997.
47. Moshous, D., Callebaut, I., de Chasseval, R., Comeo, B., Cavazzana-Calvo, M., Le Deist, F., Tezcan, I., Sanal, O., Bertrand, Y., Philippe, N., Fischer, A., and de Villartay, J. P. (2001) Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency, *Cell* 105, 177–186.
48. Ma, Y., Schwarz, K., and Lieber, M. R. (2005) The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps, *DNA Repair* 4, 845–851.
49. Gately, D. P., Hittle, J. C., Chan, G. K., and Yen, T. J. (1998) Characterization of ATM expression, localization, and associated DNA-dependent protein kinase activity, *Mol. Biol. Cell* 9, 2361–2374.
50. Costanzo, V., Robertson, K., Ying, C. Y., Kim, E., Avvedimento, E., Gottesman, M., Grieco, D., and Gautier, J. (2000) Reconstitution of an ATM-dependent checkpoint that inhibits chromosomal DNA replication following DNA damage, *Mol. Cell* 6, 649–659.
51. Cortez, D., Wang, Y., Qin, J., and Elledge, S. J. (1999) Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks, *Science* 286, 1162–1166.
52. Pruss, D., Hayes, J. J., and Wolffe, A. P. (1995) Nucleosomal anatomy—where are the histones? *Bioessays* 17, 161–170.
53. Shilatifard, A. (2006) Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression, *Annu. Rev. Biochem.*, posted online Feb 22, <http://dx.doi.org/10.1146/annurev.biochem.75.103004.142422>.

54. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., and Bonner, W. M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139, *J. Biol. Chem.* **273**, 5858–5868.
55. Rogakou, E. P., Boon, C., Redon, C., and Bonner, W. M. (1999) Megabase chromatin domains involved in DNA double-strand breaks *in vivo*, *J. Cell Biol.* **146**, 905–916.
56. Burma, S., Chen, B. P., Murphy, M., Kurimasa, A., and Chen, D. J. (2001) ATM phosphorylates histone H2AX in response to DNA double-strand breaks, *J. Biol. Chem.* **276**, 42462–42467.
57. Xu, Y. (2006) DNA damage: a trigger of innate immunity but a requirement for adaptive immune homeostasis, *Nat. Rev. Immunol.* **6**, 261–270.
58. Bassing, C. H., Chua, K. F., Sekiguchi, J., Suh, H., Whitlow, S. R., Fleming, J. C., Monroe, B. C., Ciccone, D. N., Yan, C., Vlasakova, K., Livingston, D. M., Ferguson, D. O., Scully, R., and Alt, F. W. (2002) Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8173–8178.
59. Celeste, A., Petersen, S., Romanienko, P. J., Fernandez-Capetillo, O., Chen, H. T., Sedelnikova, O. A., Reina-San-Martin, B., Coppola, V., Meffre, E., Difilippantonio, M. J., Redon, C., Pilch, D. R., Oлару, A., Eckhaus, M., Camerini-Otero, R. D., Tessarollo, L., Livak, F., Manova, K., Bonner, W. M., Nussenzweig, M. C., and Nussenzweig, A. (2002) Genomic instability in mice lacking histone H2AX, *Science* **296**, 922–927.
60. Celeste, A., Fernandez-Capetillo, O., Kruhlak, M. J., Pilch, D. R., Staudt, D. W., Lee, A., Bonner, R. F., Bonner, W. M., and Nussenzweig, A. (2003) Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks, *Nat. Cell Biol.* **5**, 675–679.
61. Bird, A. W., Yu, D. Y., Pray-Grant, M. G., Qiu, Q., Harmon, K. E., Megee, P. C., Grant, P. A., Smith, M. M., and Christman, M. F. (2002) Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair, *Nature* **419**, 411–415.
62. Jazayeri, A., McAnish, A. D., and Jackson, S. P. (2004) *Saccharomyces cerevisiae* Sin3p facilitates DNA double-strand break repair, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1644–1649.
63. Robert, F., Hardy, S., Nagy, Z., Baldeyron, C., Murr, R., Dery, U., Masson, J. Y., Papadopoulos, D., Herczeg, Z., and Tora, L. (2006) The transcriptional histone acetyltransferase cofactor TRRAP associates with the MRN repair complex and plays a role in DNA double-strand break repair, *Mol. Cell Biol.* **26**, 402–412.
64. Lee, S. H., Oshige, M., Durant, S. T., Rasila, K. K., Williamson, E. A., Ramsey, H., Kwan, L., Nickoloff, J. A., and Hromas, R. (2005) The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18075–18080.
65. Wang, S., Guo, M., Ouyang, H., Li, X., Cordon-Cardo, C., Kurimasa, A., Chen, D. J., Fuks, Z., Ling, C. C., and Li, G. C. (2000) The catalytic subunit of DNA-dependent protein kinase selectively regulates p53-dependent apoptosis but not cell-cycle arrest, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1584–1588.
66. Martin, S. A., and Ouchi, T. (2005) BRCA1 phosphorylation regulates caspase-3 activation in UV-induced apoptosis, *Cancer Res.* **65**, 10657–10662.
67. Sinclair, A., Yarranton, S., and Schelcher, C. (2006) DNA-damage response pathways triggered by viral replication, *Expert Rev. Mol. Med.* **8**, 1–11.
68. Weitzman, M. D., Carson, C. T., Schwartz, R. A., and Lilley, C. E. (2004) Interactions of viruses with the cellular DNA repair machinery, *DNA Repair* **3**, 1165–1173.
69. Izsvak, Z., Stuwe, E. E., Fiedler, D., Katzer, A., Jeggo, P. A., and Ivics, Z. (2004) Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells, *Mol. Cell* **13**, 279–290.
70. Lilley, C. E., Carson, C. T., Muotri, A. R., Gage, F. H., and Weitzman, M. D. (2005) DNA repair proteins affect the lifecycle of herpes simplex virus 1, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5844–5849.
71. Yant, S. R., and Kay, M. A. (2003) Nonhomologous-end-joining factors regulate DNA repair fidelity during Sleeping Beauty element transposition in mammalian cells, *Mol. Cell Biol.* **23**, 8505–8518.
72. Katz, R. A., and Skalka, A. M. (1994) The retroviral enzymes, *Annu. Rev. Biochem.* **63**, 133–173.
73. Craigie, R., Fujiwara, T., and Bushman, F. (1990) The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*, *Cell* **62**, 829–837.
74. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J., and Skalka, A. M. (1990) The avian retroviral IN protein is both necessary and sufficient for integrative recombination *in vitro*, *Cell* **63**, 87–95.
75. Vincent, K. A., York-Higgins, D., Quiroga, M., and Brown, P. O. (1990) Host sequences flanking the HIV provirus. *Nucleic Acids Res.* **18**, 6045–6047.
76. Jeanson, L., Subra, F., Vaganay, S., Hervy, M., Marangoni, E., Bourhis, J., and Mouscadet, J. F. (2002) Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells, *Virology* **300**, 100–108.
77. Waninger, S., Kuhlen, K., Hu, X., Chatterton, J. E., Wong-Staal, F., and Tang, H. (2004) Identification of cellular cofactors for human immunodeficiency virus replication *via* a ribozyme-based genomics approach, *J. Virol.* **78**, 12829–12837.
78. Li, L., Olvera, J. M., Yoder, K. E., Mitchell, R. S., Butler, S. L., Lieber, M., Martin, S. L., and Bushman, F. D. (2001) Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection, *EMBO J.* **20**, 3272–3281.
79. Daniel, R., Ramcharan, J., Rogakou, E., Taganov, K. D., Greger, J. G., Bonner, W., Nussenzweig, A., Katz, R. A., and Skalka, A. M. (2004) Histone H2AX is phosphorylated at sites of retroviral DNA integration but is dispensable for postintegration repair, *J. Biol. Chem.* **279**, 45810–45814.
80. Nunnari, G., Argyris, E., Fang, J., Mehlman, K. E., Pomerantz, R. J., and Daniel, R. (2005) Inhibition of HIV-1 replication by caffeine and caffeine-related methylxanthines, *Virology* **335**, 177–184.
81. Dehart, J. L., Andersen, J. L., Zimmerman, E. S., Ardon, O., An, D. S., Blackett, J., Kim, B., and Planelles, V. (2005) The ataxia telangiectasia-mutated and Rad3-related protein is dispensable for retroviral integration, *J. Virol.* **79**, 1389–1396.
82. Ariumi, Y., Turelli, P., Masutani, M., and Trono, D. (2005) DNA damage sensors ATM, ATR, DNA-PKcs, and PARP-1 are dispensable for human immunodeficiency virus type 1 integration, *J. Virol.* **79**, 2973–2978.
83. Lee, J. W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T. A., Wang, Z., and Povirk, L. F. (2004) Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts, *J. Biol. Chem.* **279**, 805–811.
84. Ma, Y., Lu, H., Tippin, B., Goodman, M. F., Shimazaki, N., Koiwai, O., Hsieh, C. L., Schwarz, K., and Lieber, M. R. (2004) A biochemically defined system for mammalian nonhomologous DNA end joining, *Mol. Cell* **16**, 701–713.
85. Mahajan, K. N., Nick McElhinny, S. A., Mitchell, B. S., and Ramsden, D. A. (2002) Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair, *Mol. Cell Biol.* **22**, 5194–5202.
86. Daniel, R., Marusich, E., Argyris, E., Zhao, R. Y., Skalka, A. M., and Pomerantz, R. J. (2005) Caffeine inhibits human immunodeficiency virus type 1 transduction of nondividing cells, *J. Virol.* **79**, 2058–2065.
87. Ting, N. S., and Lee, W. H. (2004) The DNA double-strand break response pathway: becoming more BRCAish than ever, *DNA Repair* **3**, 935–944.
88. Bassing, C. H., and Alt, F. W. (2004) H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity, *Cell Cycle* **3**, 149–153.
89. Morrish, T. A., Gilbert, N., Myers, J. S., Vincent, B. J., Stamato, T. D., Taccioli, G. E., Batzer, M. A., and Moran, J. V. (2002) DNA repair mediated by endonuclease-independent LINE-1 retrotransposition, *Nat. Genet.* **31**, 159–165.

Probing the Mg^{2+} Blockade Site of an *N*-Methyl-D-aspartate (NMDA) Receptor with Unnatural Amino Acid Mutagenesis

Kathryn A. McMenimen[†], E. James Petersson[†], Henry A. Lester[‡], and Dennis A. Dougherty^{†,*}

[†]Division of Chemistry and Chemical Engineering and [‡]Division of Biology, California Institute of Technology, Pasadena, California 91125

N-Methyl-D-aspartate (NMDA) receptors play a central role in modulating synaptic changes associated with learning and memory. NMDA receptors are members of a family of ligand-gated ion channels that respond to glutamate, the major excitatory neurotransmitter in the central nervous system (CNS), which also includes the α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and kainate receptors (1). Along with their pivotal biological role, NMDA receptors have fascinating chemical properties. In order to become active, NMDA receptors must bind both glutamate and either glycine or D-serine (2, 3) (Figure 1, panel a). However, at normal resting membrane potentials (more negative than -40 mV), the binding of both ligands is still insufficient for receptor activation. This is because NMDA receptors contain yet another binding site that is occupied by Mg^{2+} , and this blocks the pore. Membrane depolarization (*i.e.*, a change toward a more positive value) induced by the opening of other ion channels in the synapse promotes dissociation of the Mg^{2+} , allowing Na^+ , K^+ , and Ca^{2+} flux through the channel (Figure 1, panel a). As such, NMDA receptors serve as a kind of “coincidence detector”, responding both to the presence of glutamate and glycine and the simultaneous activation of other ion channels. These features make NMDA receptors unique among glutamate receptors and implicate them in the types of synaptic plasticity generally associated with learning and memory (4, 5). In addition, dysfunction of glutamate receptors may be associated with schizophrenia, epilepsy, stroke damage, and altered perception of pain (6–9).

Recent studies indicate that NMDA receptors are heterotetrameric transmembrane proteins, composed of

ABSTRACT The *N*-methyl-D-aspartate (NMDA) receptor plays a central role in learning and memory in the mammalian CNS. At normal neuronal resting membrane potentials, the pore of this glutamate-gated ion channel is blocked by a Mg^{2+} ion. Previous work suggests that the Mg^{2+} binding site is quite novel, involving several asparagine residues and a cation– π interaction between Mg^{2+} and a conserved tryptophan in the pore. Using unnatural amino acid mutagenesis, we show that no such cation– π interaction exists. The implicated tryptophan instead appears to play a structural role that can only be fulfilled by a rigid, flat, hydrophobic residue. This is the first demonstration of unnatural amino acid incorporation in the NMDA receptor, and it opens the way for future investigations of this pivotal neuroreceptor.

*To whom correspondence should be addressed.
E-mail: dadougherty@caltech.edu.

Received for review March 1, 2006
and accepted April 21, 2006

Published online May 12, 2006
10.1021/cb6000944 CCC: \$33.50

© 2006 by American Chemical Society

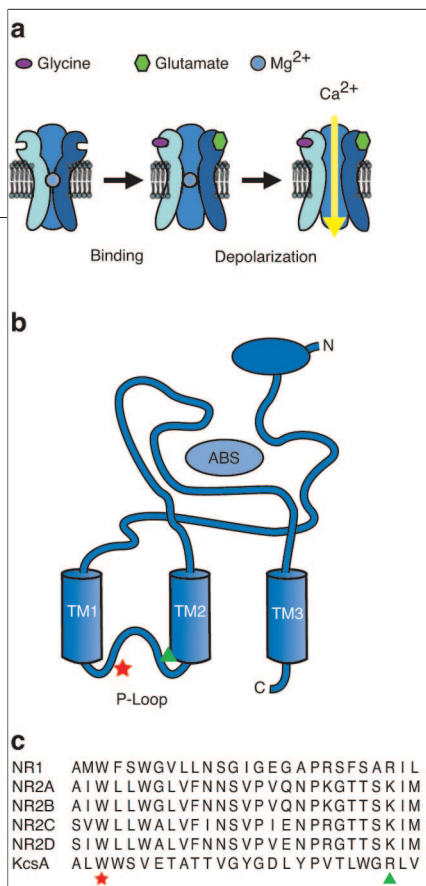


Figure 1. The NMDA receptor. a) Overall layout of the tetramer (one subunit removed). Binding of glycine to NR1 and glutamate to NR2 activates the channel, and membrane depolarization favors dissociation of the blocking Mg²⁺. b) Topology of a single subunit, showing the extracellular N-terminal domain, the location of the agonist (Glu/Gly) binding site (ABS), three transmembrane domains, and the re-entrant P loop. The approximate location of Trp 607 (NR2B) studied in the present work is denoted as a red star. Arg 630 (NR1a) is denoted as a green triangle. c) Sequence alignments. Trp 607 (NR2B) is denoted as a red star, and Arg 630 (NR1a) is denoted as a green triangle.

intriguing from a chemical standpoint (1). Several asparagine residues, N616 in NR1 and corresponding residues in NR2, have been implicated in Mg²⁺ block. Surprisingly, a tryptophan residue, NR2BW607 is also considered to contribute to the Mg²⁺ binding site. Williams *et al.* (16) showed that only an aromatic residue functions at this site; other residues abolish Mg²⁺ block. This observation led to the speculation that a cation- π interaction (17–19) between NR2BW607 and Mg²⁺ is crucial to Mg²⁺ blockade. Importantly, the same residues influence the binding of organic amine-containing blockers such as MK-801, TB-3-4, and the clinically useful memantine, along with the addictive drugs, PCP and ketamine (20, 21).

a combination of two NR1 and two NR2/3 subunits. Several different subunit subtypes have been identified, including NR1, NR2A/B/C/D, and NR3A/B (1). The presently accepted layout of each subunit is shown (Figure 1, panel b) involving a large extracellular N-terminus, three transmembrane segments (TM1, TM2, and TM3), an intracellular C-terminus, and a re-entrant P loop (10) thought to resemble the P loop of K⁺ channels (11) (although inverted relative to the K⁺ channel P loop). The re-entrant P loop contributes to the ion permeability properties of the channel.

While the glutamate and glycine binding sites have been extensively investigated by Gouaux and co-workers (12–14), much less structural information is available for the magnesium binding site. The large voltage dependence of block suggests that the Mg²⁺ binding site lies deep within the pore (15). Further work has led to a proposed NMDA receptor Mg²⁺ binding site that is novel and that is very

Protein binding sites for divalent cations such as Mg²⁺ or Ca²⁺ are typically comprised of negatively charged residues (Asp or Glu) that chelate the metal cation (22). While the side chain carbonyls of Asn residues can be envisioned as strong binders of Mg²⁺, the proposed cation- π interaction is quite unusual. There is no evidence to support a direct cation- π interaction between a divalent ion and an aromatic π system in aqueous media. A few isolated examples of water-mediated cation- π interactions have been proposed (23): a Mg²⁺OH₂-Phe interaction in the bacterial chemotaxis protein CheY and a Ca²⁺OH₂-Tyr interaction in troponin C. A Mg²⁺OH₂-cytosine interaction in a DNA complex has been proposed (24), but disputed (25, 26).

Over the last decade, we have established that unnatural amino acid mutagenesis can decisively test the importance of cation- π interactions in neuroreceptors and ion channels (27–31). The goals of the present work were thus twofold. First, we wished to establish whether the *in vivo* nonsense suppression methodology for unnatural amino acid incorporation which has been used extensively in other receptor and channel systems could be adapted to the NMDA receptor (32). Given the pivotal biological role of these channels and the many intriguing chemical questions associated with them, success would open up a wide range of intriguing experiments. Second, we wished to evaluate the proposed cation- π interaction between NR2BW607 and Mg²⁺. If the proposed interaction is correct, this would constitute a highly novel binding site in a critical receptor.

We find that the nonsense suppression methodology can be implemented in the NMDA receptor. Our results, however, show that incorporation of fluorinated Trp residues at NR2BW607 did not alter Mg²⁺ binding, indicating that Mg²⁺ does not bind to NR2BW607 through a cation- π interaction. Instead, our data suggest that NR2BW607 is best considered a large, flat, hydrophobic residue in the pore of the channel acting as a structural element during receptor block.

RESULTS AND DISCUSSION

Incorporation of Unnatural Amino Acids. The *in vivo* nonsense suppression method was used to incorporate unnatural amino acids into the NMDA receptor (Figure 2), and the mutant receptors were evaluated electrophysiologically (Table 1). All mutations reported are in the NR2B subunit unless otherwise noted and are

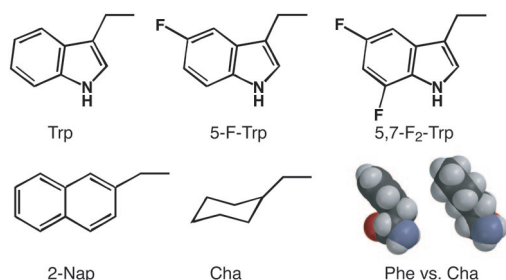


Figure 2. Structures of side chains studied here. Also shown is a representation of Phe and Cha, showing their similarities in size and shape.

co-expressed with NR1a (a splice variant of the NR1 gene lacking exon 5). Given the 2:2 stoichiometry of this tetrameric receptor, a single mutation in NR2B changes two side chains. As these are the first applications of this methodology to the NMDA receptor, several key control experiments were essential. There are reports that NR1 subunits can form homomeric channels in oocytes (33), and if so, this would complicate our analyses. However, we found that, when we expressed mRNA corresponding to NR1a only and applied agonists in the presence or absence of Mg^{2+} , no current was recorded (34), ruling out distortion by NR1a homotetramers.

To establish the viability of nonsense suppression in the NMDA receptor, a “wild-type recovery” experiment was performed. First, conventional wild-type NR1a/NR2B receptors were expressed. They responded to appropriate concentrations of glutamate/glycine (always applied together), and they were inhibited by Mg^{2+} with an IC_{50} of $76 \pm 5.5 \mu M$ at -60 mV (Figure 3, panels a and b, blue). We also made the conventional mutant W607L, and this abolished Mg^{2+} block (Figure 3, panel c), as previously reported (16). We then incorporated the TAG stop codon at W607 and injected this mRNA and NR1a mRNA along with our established nonsense suppression tRNA THG73 (35) ligated to tryptophan. Robust responses to glutamate/glycine were seen (Figure 3, panels a and b, red), and the channels were blocked by Mg^{2+} with an IC_{50} of $82 \pm 22 \mu M$ at -60 mV, equivalent to the conventionally prepared wild-type receptor within experimental error (Table 1). If no tRNA was added, no signals were seen. If tRNA that had not been charged with an amino acid was added, very small currents that could not be blocked by Mg^{2+} were measured. This establishes that we can cleanly incorporate specific residues into the NMDA receptor.

We then incorporated a number of indole analogues at this tryptophan site. Taking advantage of the well-

TABLE 1. IC_{50} data for wild-type and mutant NMDA receptors^a

Subunit/amino acid	Mg^{2+} IC_{50} (μM) -60 mV	Cation- π binding (Na^+) kcal mol ⁻¹	<i>n</i>	<i>I</i> max avg. (nA)	δ	$K_{0.5}$ (0 mV) (mM)
Wild-type NR1a/2B	76 ± 5.5	32.6	7	850	0.74 ± 0.11	3.7 ± 2.9
NR1a/2B607TAG W	82 ± 22	32.6	6	102	0.77 ± 0.07	11 ± 3.2
NR1a/2B607TAG 2-Nap	37 ± 4	28.9	5	102	0.54 ± 0.15	0.5 ± 0.3
NR1a/2B607TAG 5-F-Trp	52 ± 5	27.5	8	170	0.36 ± 0.11	1.6 ± 1.1
NR1a/2B607TAG 5,7-F ₂ -Trp	52 ± 10	23.3	5	35	0.73 ± 0.03	2.9 ± 0.6
NR1a/2B607TAG Cha	190 ± 11	8.4	4	160	0.39 ± 0.06	1.1 ± 0.3
NR1a/2B607L	No Block		4	510		
NR1a/2B607TAG Y	170 ± 28	26.9	4	1100		
NR1a/NR2BW607R	No Block		4	100		
NR1aR630W/2B	NF					
NR1aR630W/2BW607R	NF					
NR1a/2B607TAG 76mer	No Block		5	25		

^a*n* is the number of oocytes measured to obtain each value. IC_{50} values, δ values, and $K_{0.5}$ values are shown as mean \pm SEM. NF indicates no channels detected. Cation- π binding values are computed as in ref 36.

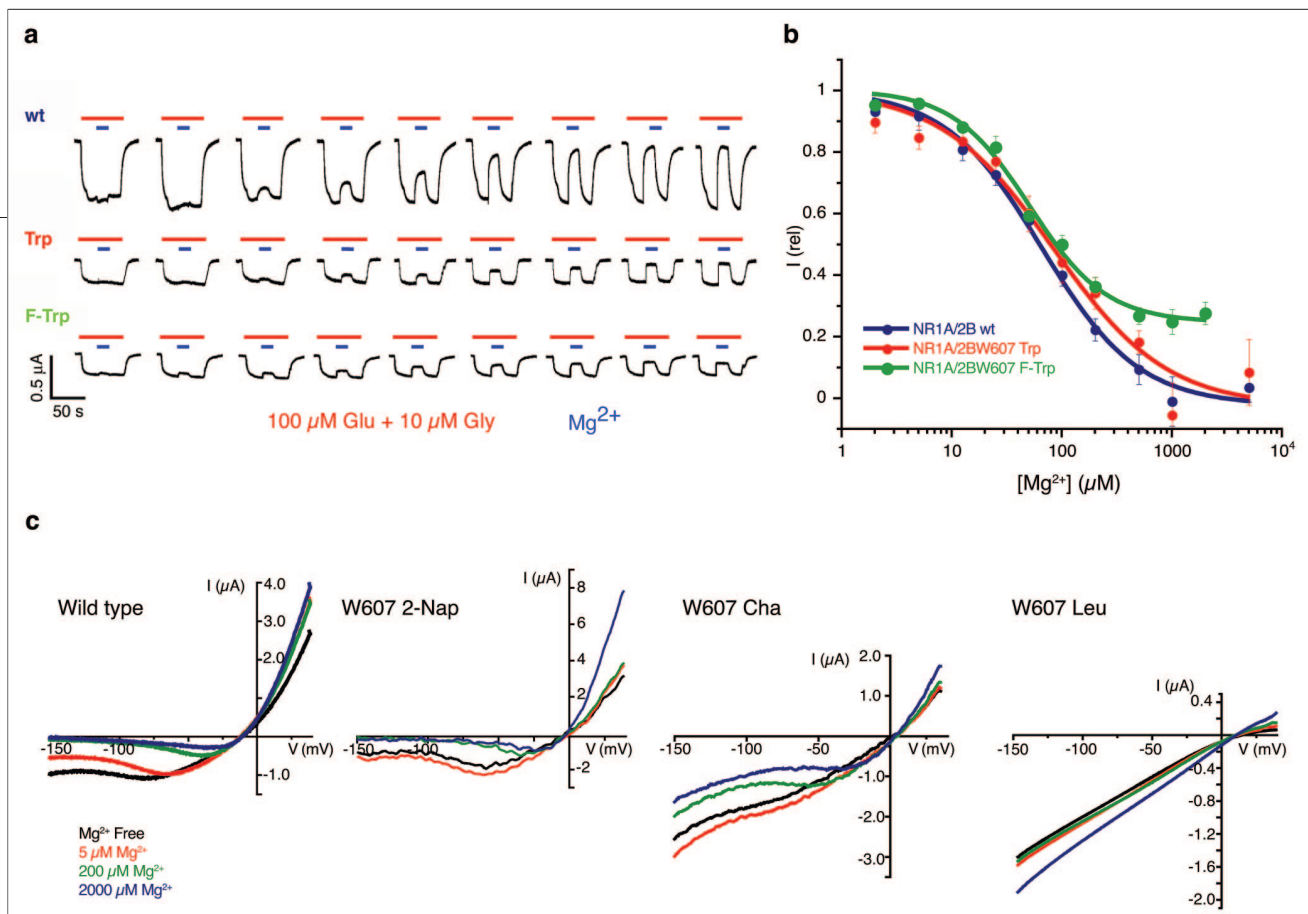


Figure 3. Mg^{2+} block of NMDA receptors. **a**) Sample traces of voltage-clamp currents for conventional wild-type (wt), wt-recovery, and F-Trp electrophysiology experiments. **b**) Dose–response relations for Mg^{2+} block, generated from the data in panel a. NR1a/NR2B (conventional wt/blue), NR1a/NR2BW607 W (wt-recovery/red), and NR1a/NR2BW607 5-F-Trp (F-Trp/green). Residual current seen in the 5-F-Trp case at high $[Mg^{2+}]$ has been seen in other mutants (16) and has been attributed to slight permeability of Mg^{2+} at negative potentials. In previous work (16) and in our own experience, even the wild-type receptor shows small and variable Mg^{2+} permeability at very high $[Mg^{2+}]$. **c**) Current–voltage (I – V) relations generated by voltage ramps; see Methods. NR1a/NR2B, NR1a/NR2BW607 2-Nap, NR1a/NR2BW607 Cha, and NR1a/NR2BW607L.

established effect of fluorination in diminishing the cation– π binding ability of an aromatic ring (27–31), we incorporated 5-F-Trp at W607 (Figure 3, panels a and b). Measurement of the subsequent Mg^{2+} block gave an IC_{50} of $52 \pm 5 \mu$ M. The difference from wild-type is small, but if anything, the affinity of the channel for Mg^{2+} has increased. In the cation– π interaction, 2-naphthylalanine (2-Nap) is similar to 5-F-Trp, and 5,7-F₂-Trp is much weaker, yet all show near wild-type Mg^{2+} blockade (Figure 3, panel c; Table 1).

In previous work, we used the *ab initio* calculated binding energy of a Na^+ ion to an aromatic ring as a measure of the ring’s cation– π binding ability (27, 36). Fluorination of indole decreases the negative electrostatic potential of the indole ring, weakening the cation– π binding ability of the side chain. To verify that nothing unusual is occurring with Mg^{2+} , we performed comparable calculations for Mg^{2+} binding to representative aromatics. In the gas phase, the binding of Mg^{2+} to indole gives a ΔE of $-142.3 \text{ kcal mol}^{-1}$ (vs $-32.6 \text{ kcal mol}^{-1}$ for Na^+). With 5-fluoroindole, the

affinity is lessened by $11.3 \text{ kcal mol}^{-1}$ (vs $5.1 \text{ kcal mol}^{-1}$ when Na^+ is the probe cation). As anticipated for an interaction with a large electrostatic component, the cation– π interaction is much stronger for Mg^{2+} than for Na^+ , due to the larger charge density on the divalent ion. The effect of fluorination is also larger. Comparable results are seen when implicit solvation models for water, ethanol, or THF are added to the calculations. The calculations clearly establish that a substantial change would be seen in the Mg^{2+} IC_{50} experiments if a direct Mg^{2+} -indole cation– π interaction were important. By way of reference, in receptors for serotonin and γ -aminobutyric acid, both ligands of the type RNH_3^+ , difluorination of a binding site aromatic decreased affinities by values ranging from 30- to 180-fold (28, 30). The lack of any effect for the mutants probed here convincingly rules out a cation– π interaction in Mg^{2+} blockade.

Previous work showed that only aromatic residues function well at the W607 site (16), but the present results rule out both a cation– π interaction as well as

any other interaction with a sizable electrostatic component (e.g., π - π interactions). At the same time, just having a hydrophobic side chain is not enough, as the W607L mutant shows no Mg^{2+} block. The lack of block in the W607L mutant also rules out an alternative explanation of our observations, that the cation- π interaction “jumps” to the NR1 subunit (e.g., W608) when deactivated substituents are placed in NR2B. Another possibility is that a relatively rigid, flat, hydrophobic side chain is required. To probe this, we compared tyrosine and cyclohexylalanine (Cha), which are similar in size and shape (Figure 2), but, of course, only the former is aromatic. The IC_{50} for Mg^{2+} blockade is essentially identical for the two residues. It thus appears that a rigid, relatively flat, hydrophobic residue is required at position 607 for potent Mg^{2+} block.

I-V Relationships. The Mg^{2+} block of the NMDA receptor is voltage-dependent. We therefore examined agonist-induced current-voltage (*I-V*) relationships in the presence and absence of Mg^{2+} . To obtain Woodhull parameters (37) the logarithm of the IC_{50} at different potentials was plotted against the holding potentials. A straight line was fit to the data to determine δ and $K_{0.5}$. The wild-type NR1a/NR2B receptors show the expected concentration- and voltage-dependent Mg^{2+} block (Figure 3, panel c). Block of the W607 2-Nap mutant was also strongly concentration- and voltage-dependent (Figure 3, panel c). The side chain, 2-Nap, has a smaller cation- π binding ability than Trp, yet the voltage dependence of the Mg^{2+} block is similar to that of wild-type. The W607Y mutant never reaches full block (data not shown), although its cation- π binding energy is similar to 2-Nap (Table 1). This indicates that, like IC_{50} , the voltage dependence of the channel block is not correlated to the cation- π binding ability at W607.

The W607 Cha mutant does show concentration- and voltage-dependent Mg^{2+} block of the receptor, but the block is incomplete at even the highest Mg^{2+} concentration tested (2 mM). Furthermore, the block may be relieved at high negative potentials (Figure 3, panel c), suggesting that Mg^{2+} can permeate as well as block in this mutant (37). The W607L mutant, as shown in previous studies, displays little or no Mg^{2+} block at any concentration or voltage tested. Structurally, both Cha and Leu are hydrophobic residues with no significant cation- π binding ability. However, there are noticeable differences between the block of Cha and Leu. The receptor containing Cha is partially blocked by

Mg^{2+} , whereas the Leu-containing receptor is not (Figure 3, panel c). These data again suggest that the size and shape of the side chain influences block more than aromaticity.

Similarities between Potassium Channels and Glutamate Receptors. The absence of a cation- π interaction between NR2BW607 and Mg^{2+} led us to investigate other possible roles for this Trp in channel function. Previous studies and sequence comparisons have suggested a structural similarity between the conduction pores of potassium channels and glutamate receptors (11). The pore-forming regions of K^+ channels have high homology with the P loop of glutamate receptors (11, 38), including conservation of the residue that corresponds to NR2BW607 (Figure 1, panel c). Structural observations of the KcsA channel identified an intersubunit cation- π interaction between KcsA W67 (NR2BW607 equivalent) and KcsA R49 (see Supporting Information). The equivalent residue to KcsA R49 is either an Arg or Lys in glutamate receptors (Arg in all subunits that would participate in this interaction). To study the possible importance of an intersubunit cation- π interaction, we performed a “residue swapping” experiment, evaluating the NR1aR630W and NR2BW607R mutants and the double mutant NR1aR630W/NR2BW607R. The NR2BW607R mutant activated upon agonist application, but as expected, block by extracellular Mg^{2+} was abolished. The NR1aR630W single mutant and NR1aR630W/NR2BW607R double mutant did not produce any current in response to agonist application.

To determine whether the apparently nonfunctional mutant receptors were expressing at the plasma membrane, the localization of these mutants was studied using an NR1-specific antibody. HEK293T cells expressing the wild-type or the functional mutant (NR1a/NR2BW607R) showed clear labeling of the membrane in nonpermeabilized cells (see Supporting Information). In contrast, no such labeling was seen for the nonfunctional receptor mutants (NR1aR630W/NR2B and NR1aR630W/NR2BW607R). Permeabilization of the membranes allowed labeling of intracellular NR1a subunits in all cases. While it is possible that expression patterns would be different in HEK293T cells than in *Xenopus* oocytes, we feel the more reasonable explanation is that the receptors that were “nonfunctional” in fact failed to fold and/or transport properly to the cell surface. Although we hypothesized that the double

mutant (NR1aR630W/NR2BW607R) might rescue the receptor by re-establishing a cation- π interaction, the results suggest that the Arg at position 630 in NR1a is necessary for receptor assembly and/or transport.

In all glutamate receptor subunits and several K⁺ channel subunits, a conserved Trp is present, equivalent to NR2B W607 (Figure 1, panel c). Previous studies of this Trp suggested a role in the Mg²⁺ block of the NMDA receptor, perhaps via a cation- π interaction. From the results presented here, however, it is clear that W607 is not involved in a cation- π interaction during the Mg²⁺ blockade. Incorporation of 5-F-Trp and 5,7-F₂-Trp did not alter Mg²⁺ IC₅₀, while incorporation of Cha introduced only an ~2.5-fold change in IC₅₀ from wild-type receptors. In the case of Cha, the cation- π binding ability is negligible, and so the relatively small IC₅₀ change completely eliminates the cation- π interaction from consideration. Nonetheless, results with many different side chains do show that there are some structural requirements necessary to obtain functional block of the receptor. Large, flat side chains, such as Trp, 2-Nap, F-Trp, F₂-Trp, and Cha, all maintain significant levels of Mg²⁺ block.

The data also have subtler features. Williams *et al.* (16) reported a difference in Mg²⁺ block between wild-type and W607Y mutants: the tyrosine residue produces an incomplete Mg²⁺ block, which is relieved at high negative potentials. Our experiments reveal similar results with W607Y (not shown) as well as with W607 Cha. This implies that the size and/or shape of the Tyr or Cha side chains at position 607 does subtly distort the pore, so that Mg²⁺ can permeate more easily. This relatively small permeation of Mg²⁺ under some conditions

also fails to correlate with cation- π binding and may be influenced by other residues.

The structural implications for W607 based on homologous residues in potassium channels suggested that this conserved Trp might be involved in an intersubunit cation- π interaction with a conserved Arg (on NR1a in NMDA receptors). The NR1a/NR2BW607R mutant produced otherwise functional channels that are not blocked by Mg²⁺. However, the mutations NR1aR630W/NR2B and NR1aR630W/NR2BW607R produced receptors that were not trafficked to the cell surface. Additionally, the W607 Cha data would suggest that an aromatic residue is not absolutely necessary for receptor function, since partial block is observed in these channels. It is reasonable to suggest that the important features of the Trp residue are its size, shape, and hydrophobicity, which allow it to fit into the channel architecture to create the desired Mg²⁺ block during NMDA receptor function.

The present results establish that the powerful unnatural amino acid methodology that has been applied to several other channels and receptors can be applied to one of the more complicated neuroreceptors, the NMDA receptor. This lays the foundation for a large number of potentially intriguing studies. In this first study, we provide compelling evidence that there is not a cation- π interaction between a conserved tryptophan residue and a Mg²⁺ ion that has a critical channel blocking function. Instead, it appears that a rigid, relatively flat, hydrophobic side chain is required at this site. In future studies, we will evaluate several asparagine residues that may contribute to the novel Mg²⁺ binding site of this critical CNS receptor.

METHODS

Electrophysiology. Stage V-VI *Xenopus laevis* oocytes were injected with 50 nL cell⁻¹ of mRNA/tRNA mixtures. Oocytes were evaluated in a Mg²⁺ and Ca²⁺ free saline solution (96 mM NaCl, 5 mM HEPES, 2 mM KCl, and 1 mM BaCl₂). The receptors were activated in a Mg²⁺ and Ca²⁺ free solution containing 100 μ M glutamate (Aldrich), 10 μ M glycine (Aldrich), and 100 μ M niflumic acid (to reduce activity of Ca²⁺-activated Cl⁻ channels, Sigma). All oocyte recordings were made 48 h after initial injection in two-electrode voltage clamp mode using the OpusXpress 6000A (Molecular Devices). Solutions were superfused at flow rates of 1 and 4 mL min⁻¹ during Mg²⁺ application and 3 mL min⁻¹ during wash. Eight oocytes were simultaneously voltage-clamped at -60 mV, and dose-response relationships were obtained by delivery of various concentrations of Mg²⁺ in 1 mL aliquots. The data were analyzed using the Clampfit 9.0 software (Axon). The Hill equation was used to fit data:

$I/I_{\max} = 1/(1 + (IC_{50}/[A])^{n_H})$, where I is peak current at drug concentration (A), IC_{50} is the concentration of drug that inhibits 50% of the maximal response, and n_H is the Hill coefficient. Voltage ramps (-150 to +40 mV during 4 s) were used to construct I - V curves for (glutamate and glycine)-evoked conductance in the presence and absence of Mg²⁺. Leak currents measured in the absence of glutamate, glycine, niflumic acid, and Mg²⁺ were subtracted. The voltage dependence of block (δ , fraction of the electric field that the blocker experiences) and the affinity of Mg²⁺ at 0 mV ($K_{0.5}$) were calculated by determining the IC₅₀ at varying potentials (-40 to -100 mV). The logarithm of IC₅₀ was plotted against the holding potentials, and a straight line was fit to the data to determine δ and $K_{0.5}$. The δ and $K_{0.5}$ were calculated according to Woodhull (37): $K_{0.5}(V) = K_{0.5} \exp(z\delta VF/RT)$, V is membrane potential, z is valence, and R , T , and F have their traditional meanings.

Mutagenesis and preparation of cRNA and Unnatural Amino Acid Suppression. The NR1a and NR2B subunits were subcloned from pBluescript into pAMV (32) for increased expression in *X. laevis* oocytes. Mutant NR1a and NR2B subunits were prepared by site-directed mutagenesis. All mutant and wild-type cDNA was linearized with *NotI*, and mRNA was synthesized by *in vitro* runoff transcription using the T7 mMESSAGE mMACHINE kit (Ambion). Conventional mutants and wild-type mRNA were injected in a NR1a:NR2B (1:5) ratio. Synthetic amino acids were ligated to truncated 74-nt tRNA as described previously (32). The aminoacyl tRNA was deprotected by photolysis immediately prior to co-injection with mRNA. Typically, 5 ng of mRNA and 25 ng of tRNA-aa were injected into oocytes in a total volume of 50 nL. To increase expression of suppressed receptors, 1 day after the first injection, a “booster” of 25 ng of tRNA-aa was given, and the cells were incubated 1 additional day before electrophysiological measurements were made (48 h total).

Immunolocalization of Wild-Type and Mutant NMDA Receptors. These experiments were performed by adapting previously reported procedures (39, 40). HEK293T cells were calcium phosphate-transfected with 10 ng of cDNA. Cells were incubated at 37 °C for 2 days. Transfected cells were washed with Tris-buffered saline (3×) and fixed using ice-cold 4% paraformaldehyde in phosphate buffer. The receptors were labeled with anti-NMDAR1 clone 54.1 (Zymed) at a 1:100 dilution in phosphate-buffered saline. Inclusion of 0.3% Triton X-100 permeabilized the membranes for detection of intracellular receptor expression. The cells were incubated with primary antibody for 1 h at room temperature. Cells were washed (3×) with phosphate-buffered saline. Biotinylated anti-mouse IgG (Vector) was added to the cells, incubated for 1 h, and then rinsed (3×) with phosphate-buffered saline. Fluorescein isothiocyanate avidin D (Vector) was then added to the cells and incubated for 1 h. Coverslips were mounted in Vectashield mounting medium (Vector), and immunofluorescence was observed using a confocal microscope.

Acknowledgment: We thank K. Williams for the NMDA receptor subunit clones and C. Waters for use of the Beckman Institute Biological Imaging Center. This work was supported by grants from the National Institutes of Health (NS-34407, and NS-11756).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Dingledine, R., Borges, K., Bowie, D., and Traynelis, S. F. (1999) The glutamate receptor ion channels, *Pharmacol. Rev.* **51**, 7–61.
- Moriyoshi, K., Masu, M., Ishii, T., Shigemoto, R., Mizuno, N., and Nakanishi, S. (1991) Molecular cloning and characterization of the rat NMDA receptor, *Nature* **354**, 31–37.
- Nakanishi, N., Axel, R., and Schneider, N. A. (1992) Alternative splicing generates functionally distinct *N*-methyl-D-aspartate receptors, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8552–8556.
- Kullmann, D. M., Asztely, F., and Walker, M. C. (2000) The role of mammalian ionotropic receptors in synaptic plasticity: LTP, LTD and epilepsy, *Cell. Mol. Life Sci.* **57**, 1551–1561.
- Liu, L., Wong, T. P., Pozza, M. F., Lingenhoehl, K., Wang, Y., Sheng, M., Auberson, Y. P., and Wang, Y. T. (2004) Role of NMDA receptor subtypes in governing the direction of hippocampal synaptic plasticity, *Science* **304**, 1021–1024.
- Hollmann, M., and Heinemann, S. (1994) Cloned glutamate receptors, *Annu. Rev. Neurosci.* **17**, 31–108.
- Malinow, R., Mainen, Z. F., and Hayashi, Y. (2000) LTP mechanisms: From silence to four-lane traffic, *Curr. Opin. Neurobiol.* **10**, 352–357.
- Madden, D. R. (2002) The structure and function of glutamate receptor ion channels, *Nat. Rev. Neurosci.* **3**, 91–101.
- Cull-Candy, S., Brickley, S., and Farrant, M. (2001) NMDA receptor subunits: Diversity, development and disease, *Curr. Opin. Neurobiol.* **11**, 327–335.
- Kuner, T., Wollmuth, L. P., Karlin, A., Seeburg, P. H., and Sakmann, B. (1996) Structure of the NMDA receptor channel M2 segment inferred from the accessibility of substituted cysteines, *Neuron* **17**, 343–352.
- Wood, M. W., VanDongen, H. M., and VanDongen, A. M. (1995) Structural conservation of ion conduction pathways in K channels and glutamate receptors, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4882–4886.
- Furukawa, H., and Gouaux, E. (2003) Mechanisms of activation, inhibition and specificity: Crystal structures of the NMDA receptor NR1 ligand-binding core, *EMBO J.* **22**, 2873–2885.
- Furukawa, H., Singh, S. K., Mancusso, R., and Gouaux, E. (2005) Subunit arrangement and function in NMDA receptors, *Nature* **438**, 185–192.
- Lunn, M. L., Hogner, A., Stensbol, T. B., Gouaux, E., Egebjerg, J., and Kastrop, J. S. (2003) Three-dimensional structure of the ligand-binding core of GluR2 in complex with the agonist (S)-ATPA: Implications for receptor subunit selectivity, *J. Med. Chem.* **46**, 872–875.
- Ascher, P., and Nowak, L. (1988) The role of divalent cations in the *N*-methyl-D-aspartate responses of mouse central neurones in culture, *J. Physiol.* **399**, 247–266.
- Williams, K., Pahk, A. J., Kashiwagi, K., Masuko, T., Nguyen, N. D., and Igarashi, K. (1998) The selectivity filter of the *N*-methyl-D-aspartate receptor: A tryptophan residue controls block and permeation of Mg²⁺, *Mol. Pharmacol.* **53**, 933–941.
- Zacharias, N., and Dougherty, D. A. (2002) Cation- π interactions in ligand recognition and catalysis, *Trends Pharmacol. Sci.* **23**, 281–287.
- Ma, J. C., and Dougherty, D. A. (1997) The cation- π interaction, *Chem. Rev.* **97**, 1303–1324.
- Dougherty, D. A. (1996) Cation- π interactions in chemistry and biology: A new view of benzene, Phe, Tyr, and Trp, *Science* **271**, 163–168.
- Kashiwagi, K., Masuko, T., Nguyen, C. D., Kuno, T., Tanaka, I., Igarashi, K., and Williams, K. (2002) Channel blockers acting at *N*-methyl-D-aspartate receptors: Differential effects of mutations in the vestibule and ion channel pore, *Mol. Pharmacol.* **61**, 533–545.
- Morris, B. J., Cochran, S. M., and Pratt, J. A. (2005) PCP: From pharmacology to modelling schizophrenia, *Curr. Opin. Pharmacol.* **5**, 101–106.
- Dudev, T., and Lim, C. (2003) Principles governing Mg, Ca, and Zn binding and selectivity in proteins, *Chem. Rev.* **103**, 773–788.
- Zaric, S. D., Popovic, D. M., and Knapp, E. W. (2000) Metal ligand aromatic cation- π interactions in metalloproteins: Ligands coordinated to metal interact with aromatic residues, *Chemistry* **6**, 3935–3942.
- McFail-Isom, L., Shui, X., and Williams, L. D. (1998) Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base II systems, *Biochemistry* **37**, 17105–17111.
- Magnuson, E. C., Koehler, J., Lamm, G., and Pack, G. R. (2002) Mg(H₂O)₆²⁺- π (cytosine) interactions in a DNA dodecamer, *Int. J. Quantum Chem.* **88**, 236–243.
- Sponer, J., Sponer, J. E., and Leszczynski, J. (2000) Cation- π and amino-acceptor interactions between hydrated metal cations and DNA bases. A quantum-chemical view, *J. Biomol. Struct. Dyn.* **17**, 1087–1096.
- Zhong, W., Gallivan, J. P., Zhang, Y., Li, L., Lester, H. A., and Dougherty, D. A. (1998) From ab initio quantum mechanics to molecular neurobiology: A cation- π binding site in the nicotinic receptor, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12088–12093.

28. Lummis, S. C., Beene, D. L., Harrison, N. J., Lester, H. A., and Dougherty, D. A. (2005) A cation- π binding interaction with a tyrosine in the binding site of the GABAC receptor, *Chem. Biol.* **12**, 993–997.
29. Cashin, A. L., Petersson, E. J., Lester, H. A., and Dougherty, D. A. (2005) Using physical chemistry to differentiate nicotinic from cholinergic agonists at the nicotinic acetylcholine receptor, *J. Am. Chem. Soc.* **127**, 350–356.
30. Beene, D. L., Brandt, G. S., Zhong, W., Zacharias, N. M., Lester, H. A., and Dougherty, D. A. (2002) Cation- π interactions in ligand recognition by serotonergic (5-HT_{3A}) and nicotinic acetylcholine receptors: The anomalous binding properties of nicotine, *Biochemistry* **41**, 10262–10269.
31. Mu, T. W., Lester, H. A., and Dougherty, D. A. (2003) Different binding orientations for the same agonist at homologous receptors: A lock and key or a simple wedge? *J. Am. Chem. Soc.* **125**, 6850–6851.
32. Nowak, M. W., Gallivan, J. P., Silverman, S. K., Labarca, C. G., Dougherty, D. A., and Lester, H. A. (1998) *In vivo* incorporation of unnatural amino acids into ion channels in *Xenopus* oocyte expression system, *Methods Enzymol.* **293**, 504–529.
33. Green, T., Rogers, C. A., Contractor, A., and Heinemann, S. F. (2002) NMDA receptors formed by NR1 in *Xenopus laevis* oocytes do not contain the endogenous subunit XenU1, *Mol. Pharmacol.* **61**, 326–333.
34. Wenthold, R. J., Prybylowski, K., Standley, S., Sans, N., and Petralia, R. S. (2003) Trafficking of NMDA receptors, *Annu. Rev. Pharmacol. Toxicol.* **43**, 335–358.
35. Saks, M. E., Sampson, J. R., Nowak, M. W., Kearney, P. C., Du, F., Abelson, J. N., Lester, H. A., and Dougherty, D. A. (1996) An engineered *Tetrahymena* tRNA^{Gln} for *in vivo* incorporation of unnatural amino acids into proteins by nonsense suppression, *J. Biol. Chem.* **271**, 23169–23175.
36. Mecozzi, S., West, A. P., Jr., and Dougherty, D. A. (1996) Cation- π interactions in aromatics of biological and medicinal interest: Electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10566–10571.
37. Woodhull, A. M. (1973) Ionic blockage of sodium channels in nerve, *J. Gen. Physiol.* **61**, 687–708.
38. Kuner, T., Seeburg, P. H., and Guy, H. R. (2003) A common architecture for K⁺ channels and ionotropic glutamate receptors? *Trends Neurosci.* **26**, 27–32.
39. Lummis, S. C., Beene, D. L., Lee, L. W., Lester, H. A., Broadhurst, R. W., and Dougherty, D. A. (2005) Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel, *Nature* **438**, 248–252.
40. Spier, A. D., and Lummis, S. C. (2000) The role of tryptophan residues in the 5-hydroxytryptamine(3) receptor ligand binding domain, *J. Biol. Chem.* **275**, 5620–5625.

Chemical Chaperones and Permissive Temperatures Alter the Cellular Localization of Gaucher Disease Associated Glucocerebrosidase Variants

Anu R. Sawkar[†], Martina Schmitz[‡], Klaus-Peter Zimmer[‡], David Reczek[§], Tim Edmunds[§], William E. Balch[¶], and Jeffery W. Kelly^{†,*}

[†]Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, [‡]Universitätskinderklinik, Westfälische Wilhelms-Universität Münster, D-48149 Münster, Germany, [§]Genzyme, 500 Kendall Street, Cambridge, Massachusetts 02142, and [¶]Department of Cell and Molecular Biology and The Institute of Childhood and Neglected Disease, The Scripps Research Institute, La Jolla, California 92037

Gaucher disease is the most common lysosomal storage disorder, with an estimated incidence of 1 in 60,000 in the general population (1) and 1 in 800 among the Ashkenazi Jewish population (2). A functional deficiency of glucocerebrosidase (GC), a lysosomal hydrolase that cleaves glucosylceramide to ceramide and glucose, leads to the accumulation of glucosylceramide in the lysosomes of monocyte-macrophage cells. Glucosylceramide storage results in hepatomegaly, splenomegaly, anemia and thrombocytopenia, bone lesions, and, in severe cases, central nervous system (CNS) involvement (reviewed in 3 and 4).

Over 185 GC mutations give rise to Gaucher disease (<http://www.hgmd.org>) (5) by decreasing GC activity in the lysosome. This decrease may stem from reduced catalytic activity (specific activity), from a reduced GC concentration in the lysosome, or from both. Specific activity can be reduced directly by mutations that impair the catalytic machinery or indirectly by altering the binding of substrate or activators (6). Additionally, point mutations can cause structural destabilization, leading to non-native protein conformations with reduced catalytic activity at lysosomal pH (7). The lysosomal concentration of some GC variants is reduced because mutations compromise folding in the endoplasmic reticulum (ER) resulting in sustained molecular chaperone binding and ER-associated degradation (ERAD) mediated by the proteasome, instead of proper folding and trafficking to the lysosome (8, 9).

We hypothesize that several clinically important GC variants give rise to pathology as a consequence of their inability to exit the ER efficiently, owing to misfolding

ABSTRACT Point mutations in the lysosomal hydrolase, glucocerebrosidase (GC), can cause Gaucher disease, a common lysosomal storage disease. Several clinically important GC mutations impede folding in the endoplasmic reticulum (ER) and target these enzymes for ER-associated degradation (ERAD). The removal of these misfolded proteins decreases the lysosomal concentration of GC, which results in glucosylceramide accumulation. The most common GC variant, N370S, and other clinically relevant variants, G202R and L444P, exhibit different cellular localization patterns in patient-derived fibroblasts. We show that these distributions can be altered by manipulation of the ER folding environment, either by chemical chaperones or by temperature shifts. N370S, L444P, and G202R GC are destabilized in the neutral pH environment of the ER, rendering them prone to ERAD. Fibroblasts harboring the G202R and L444P GC mutations grown at 30 °C localize the mutant proteins to the lysosome, and this increases total GC activity. Both of these temperature-sensitive mutants appear to be stable at 37 °C once they are trafficked to the low pH environment of the lysosome. Chemical chaperones correct the ER instability and significant ER retention of G202R GC. N370S is also destabilized under ER simulating conditions, a deficiency that is corrected by chemical chaperone binding. These data clearly show manipulating the ER environment with chemical chaperones increases the lysosomal concentration of partially active GC variants and suggest that small molecules could be used to treat Gaucher disease.

*To whom correspondence should be addressed.

E-mail: jkelly@scripps.edu.

Received for review March 2, 2006
and accepted May 2, 2006.

Published online May 19, 2006
10.1021/cb600187q CCC: \$33.50

© 2006 by American Chemical Society

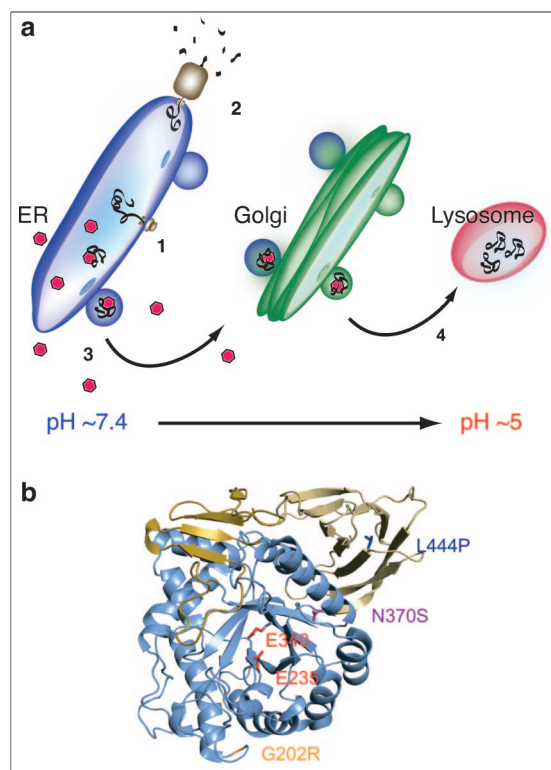


Figure 1. GC structure and function. **a)** Proposed mechanism of chemical chaperoning. Glucocerebrosidase (GC) folding commences when the sequence is inserted into the lumen of the endoplasmic reticulum (ER) at neutral pH (1). Glucocerebrosidase sequences that are deficient in folding in the ER are retained by chaperone binding and targeted for degradation by the proteasome (2). Chemical chaperones (red hexagons) stabilize the native state of GC at neutral pH by binding to it in the ER, leading to increased ER exit (3) and trafficking to the lysosome (4) by way of the Golgi complex. Lysosomal GC enzymes that are unstable in the pH 7 environment of the ER are often stable and catalytically active in the acidic lysosomal environment for which the fold and activity has been optimized. **b)** The structure of GC (22) and the location of the GC mutations investigated in this study. N370S (purple residue) is located in a helix in the active site domain (shaded blue) near the key catalytic residues E235 and E340 (red residues). G202R (orange residue) is located at the end of a helix in the same domain but is farther away from the active site than N370S. L444P (blue residue) is located in a separate Ig-like domain of the protein (shaded yellow).

and ERAD. We (10, 11) have shown that small molecules or chemical chaperones that bind to the active site of GC and template its folding (see below) in the ER can be

used to rescue the misfolding of the most common disease-associated variant N370S, which is usually associated with non-neuropathic disease (Figure 1, panel a) (12). The N370S variant is the most prevalent GC variant in the Ashkenazi Jewish population (1:865 births are homozygous or heterozygous for this mutation) (2) and is also the predominant mutation found in other European populations, albeit at a lower frequency (13, 14). The high prevalence of the N370S mutation makes it an attractive target for pharmaceutical intervention. Aerts and colleagues (15) have reported that the catalytic activity of the N370S variant is severely impaired at pH values above 5.0. Grace and colleagues (6, 16) have shown that this enzyme has a reduced catalytic activity and affinity for various inhibitors and exhibits an increased sensitivity to detergent and phosphatidylserine stimulation relative to WT-GC. These and additional data imply that the N370S mutation affects catalytic activity, conformational stability, and the ability of saposin C binding in the lysosome to activate the enzyme.

Chemical chaperones (for examples see Figure 2) are cell- and ER-permeable small molecules that bind to and stabilize the native fold of a protein in the ER, shifting the folding equilibrium away from the misfolded state (17, 18). This enables the protein–chemical chaperone complex to be trafficked to the proper protein destination environment (Figure 1, panel a). Thus, chemical chaperoning overcomes a loss-of-function phenotype, typically caused by ERAD, as long as the protein can function in its destination environment after dissociation.

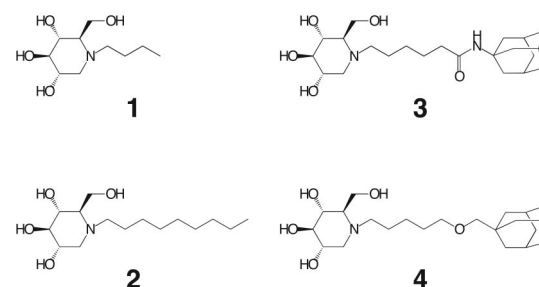


Figure 2. Line drawing depiction of the chemical structures of the GC inhibitors/GC chemical chaperones used in this study: 1, *N*-butyl-deoxyjojirimycin (Miglustat); 2, *N*-nonyl-deoxyjojirimycin; 3, *N*-hexanoic acid adamantyl amide deoxyjojirimycin; 4, *N*-pentyl adamantyl ether deoxyjojirimycin.

tion of the chemical chaperone. Perhaps surprisingly, addition of active-site-directed cell-permeable enzyme inhibitors to cells lacking that activity of that enzyme can restore enzyme function over a given concentration range because the increased folding and trafficking of the mutant enzyme result in an increased concentration that overwhelms the inhibitor, as long as the inhibitor concentration is not too high.

Chemical chaperoning can thus be utilized to increase mutant GC concentration in the lysosome (Figure 1, panel a) to levels that are widely thought to be sufficient to ameliorate Gaucher disease (>5% of WT levels) (4, 10, 11). Incubation of homozygous patient-derived fibroblasts with **2**, **3**, and **4** results in a significant increase in the cellular activity of N370S-GC (11). Chemical-chaperone-assisted GC folding allows GC variants to exit the ER and be transported to the lysosome, where they remain stable after chemical chaperone dissociation due to high substrate concentrations, the binding of protein activators, and the low pH environment for which the protein fold is optimized. Chemical chaperones can thus restore lysosomal variant GC activity by increasing the concentration there and therefore may be useful in ameliorating Gaucher disease.

A recent study suggests that G202R cells are also amenable to chemical chaperoning, whereas L444P cells appear to be resistant (11). However, it is important to realize that cell line specific effects could contribute, and therefore a negative result should be viewed suspiciously until several lines are evaluated (9, 19). Herein, we further evaluate whether the L444P GC variant, which leads to CNS pathology (20) can benefit from chemical chaperoning and/or lowered growth temperature. Importantly, we also address the mechanism of chemical chaperone function in the context of WT, L444P, N370S, and G202R GC utilizing biophysical and cellular trafficking studies. G202R GC is known to be retained in the ER (8, 21), and is therefore a particularly useful mutation to address the question of whether chemical chaperones alter variant GC trafficking. The G202R and L444P mutations are unlikely to directly affect catalytically important residues or ligand binding to GC owing to their remote location relative to the active site (Figure 1, panel b) (22). Moreover, the L444P mutation is located in the immunoglobulin domain of GC, allowing us to probe whether ligand binding to the active site domain (the domain bearing the N370S and G202R mutations) can stabilize GC variants bearing

mutations in the Ig-like domain by thermodynamic linkage of domain stabilities.

We identify point mutations that reduce GC concentration in the lysosome by ER misfolding, mistrafficking, and degradation and show that chemical chaperones and/or reduced growth temperature can be used to increase the lysosomal concentration of these variants by restoring proper folding and trafficking. Moreover, we confirm that chemical chaperoning produces GC enzymes that are stable and functional in their lysosomal destination environment.

RESULTS AND DISCUSSION

The pH-Dependent Enzymatic Activity of WT-GC and N370S-GC.

To understand the factors contributing to reduced N370S-GC activity in the lysosome, we first compared the enzymatic activities of recombinant human WT-GC (rhWT-GC) and N370S-GC (rhN370S-GC), expressed in and purified from insect cells. We also procured recombinant human WT-GC produced in Chinese hamster ovary (CHO) cells from Genzyme Inc. (this protein is also called imiglucerase and is administered to Gaucher disease patients enabling GC replacement therapy). Even though we were unable to prepare N370S-GC in CHO cells, the GC enzymes that are available enable a meaningful comparison of the glycosylated WT enzymes (WT-GC from fibroblast lysates, purified imiglucerase from CHO cells, and purified rhWT-GC

TABLE 1. The activity of recombinant GC proteins^a

Glucocerebrosidase (GC)	pH	Activity (U mg ⁻¹) ^b
Imiglucerase	5.3	9.1 ± 0.4
Imiglucerase	7.0	4.0 ± 0.3
rhWT	5.3	15.5 ± 0.5
rhWT	7.0	8.1 ± 0.4
rhN370S	5.3	3.0 ± 0.3
rhN370S	7.0	1.6 ± 0.2

^aImiglucerase, rhWT-GC, and rhN370S-GC measured in the presence of 0.1% taurodeoxycholate and 0.1% hydrogenated triton X-100 using 4-methylumbelliferyl- β -D-glucopyranoside as a substrate. Data shown are the average \pm standard deviation of three independent experiments. ^bU = μ .mol of 4-methylumbelliferone released per minute.

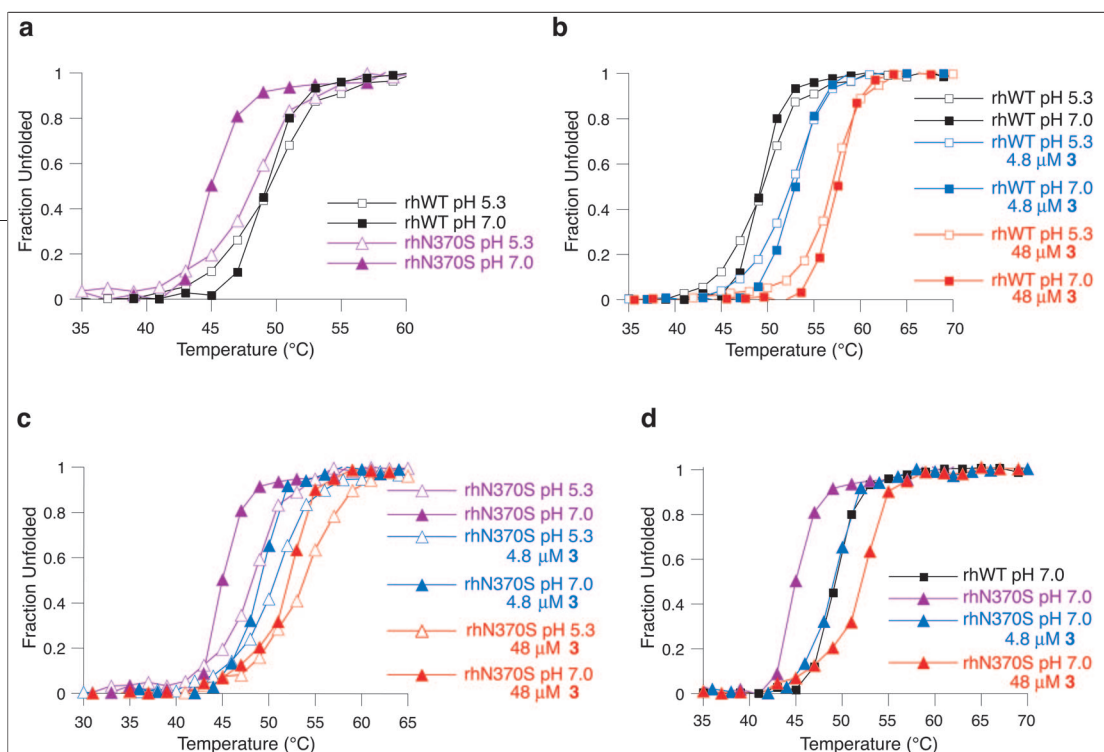


Figure 3. The pH-dependent thermal stability of recombinant WT-GC and N370S-GC in the presence and absence of chemical chaperones. **a)** Thermal denaturation curves of rhWT-GC and rhN370S-GC ($4.8 \mu\text{M}$) at pH 7.0 and 5.3 derived from far-UV circular dichroism spectra as a function of temperature. Data are shown from a representative denaturation experiment. Lines have been added to guide the eye. Thermal denaturation curves of **b)** rhWT-GC and **c)** rhN370S-GC in the presence of chemical chaperone 3 derived from far-UV circular dichroism spectra as a function of temperature. Recombinant enzyme ($4.8 \mu\text{M}$) was denatured at pH 7.0 and 5.3 in the presence of 1 equiv ($4.8 \mu\text{M}$) or 10 molar equiv ($48 \mu\text{M}$) of 3. Data are shown from a representative denaturation experiment; lines have been added to guide the eye. **d)** The data have been replotted to show that stabilization of rhN370S-GC by 3 ($4.8 \mu\text{M}$) affords a 3-N370S complex that is equal in stability to WT-GC at pH 7.0 in the absence of chemical chaperone.

from insect cells) with the glycosylated N370S enzymes (N370S-GC from fibroblast lysates and purified rhN370S-GC from insect cells) to understand the pH-dependence of GC activity and conformational stability. Imiglucerase contains a single C-terminal amino acid substitution (R495H) and a slightly different glycan structure than the recombinant human enzyme produced in insect cells (see Methods). The activities of imiglucerase, rhWT-GC, and rhN370S-GC were assessed at a lysosome-simulating pH (5.3) and at neutral pH (mimicking the ER environment) in the presence of detergents (Table 1). The activity of rhN370S-GC is reduced at lysosomal pH to $\approx 1/3$ of imiglucerase levels and $\approx 1/5$ of rhWT-GC levels. All of the enzymes are optimized for activity at lysosomal pH and are less active at pH 7.0 (Table 1). For this reason, all subsequent enzyme assays were performed under acidic conditions. A previous assessment of activities reveals the Gaucher disease associated variants (N370S, $32 \pm 7\%$; G202R, $10 \pm 5\%$; L444P, $12 \pm 7\%$) all have reduced activity relative to WT-GC (100%) in fibroblast lysates at pH 4 (11).

The pH-Dependent Stability of WT-GC and N370S-GC. To assess relative conformational stabilities, thermal denaturation curves of recombinant N370S-GC and WT-GC were recorded in buffers utilizing far-UV

circular dichroism spectroscopy at pH 7.0 and pH 5.3, simulating ER and lysosome pHs, respectively (Figure 3, panel a). The denaturation curves of rhWT-GC at both pH values are similar (open and filled black squares, Figure 3, panel a) and exhibit the same thermal denaturation midpoint (T_m). rhN370S-GC is destabilized by 4.2°C with respect to rhWT-GC in buffer at pH 7 (Figure 3, panel a, filled purple triangles versus filled black squares), implying that N370S-GC is less stable in the ER, a neutral pH compartment, relative to WT-GC. At pH 5.3, rhN370S-GC is destabilized by less than 1°C with respect to rhWT-GC (Figure 3, panel a, open purple triangles versus open black squares). Thus, trafficking N370S-GC from the neutral pH environment of the ER to the acidic environment of the lysosome should increase its stability. While the relative T_m values recorded in buffer are useful to make comparisons, the absolute values cannot be assumed to be the same in a cell owing to the distinct intracellular environment. The slope of the thermal transition appears to be pH-sensitive, as the enzyme at neutral pH undergoes a sharper melting transition than the enzyme at acidic pH, possibly because unfolding of the domains becomes uncoupled at low pH.

pH-Sensitive GC Variant Tertiary Structural Stability.

Some disease-associated variants of α -galactosidase A, the lysosomal hydrolase associated with Fabry disease, have kinetic properties similar to the WT enzyme (23) but are thermolabile at neutral pH (24). α -Galactosidase A variants that are less stable in the neutral pH environment of the ER, relative to the WT stability, are likely to be subjected to ERAD to an extent that appears to correlate with the extent of ER destabilization. To test the hypothesis that the decreased stability of GC variants at neutral pH could also lead to ERAD and therefore loss of function, we assessed the pH-dependent stability of GC variants. While we demonstrated above the recombinant N370S GC is less stable than WT GC at neutral pH in buffer, the unavailability of the purified, glycosylated recombinant L444P and G202R human GC variants led us to compare the stability of all four sequences in human fibroblast lysates.

WT, N370S, L444P, and G202R patient-derived fibroblast lysates were heat denatured at neutral and acidic pH and assayed for residual GC activity to discern whether these GC variants exhibit pH-sensitive instability. WT-GC lysates are thermostable at both pH 7.0 and 5.3 for 100 min at 40 °C (Figure 4, panel a, black diamonds) based on the complete retention of catalytic activity. L444P GC lysates also exhibit thermal enzymatic stability under these conditions (Figure 4, panel a, orange circles). This may be surprising to some as L444P is a severe mutation. However, the L444P mutation is located in the Ig-like domain of the protein, and it may simply be that the catalytic domain of this protein exhibits WT-like stability and activity while the Ig-like domain is conformationally less stable and is recognized by the ERAD machinery in the cell mediating the degradation of the entire GC enzyme. The catalytic domain of the N370S variant is slightly destabilized with respect to WT at pH 5.3 (Figure 4, panel a, closed purple triangles) and significantly destabilized at neutral pH found in the ER (Figure 4, panel a, pH 7.0; open purple triangles). The stability of the catalytic domain of the G202R variant is also pH-sensitive, with the protein retaining little residual activity at ER pH (Figure 4, panel a, pH 7.0, open green squares) and more residual activity at lysosomal pH (Figure 4, panel a, pH 5.3; closed green squares). The G202R variant is by far the least stable GC variant tested, samples incubated at pH 7.0 have <5% residual activity after 100 min at 40 °C. The stability trends revealed by biophysical

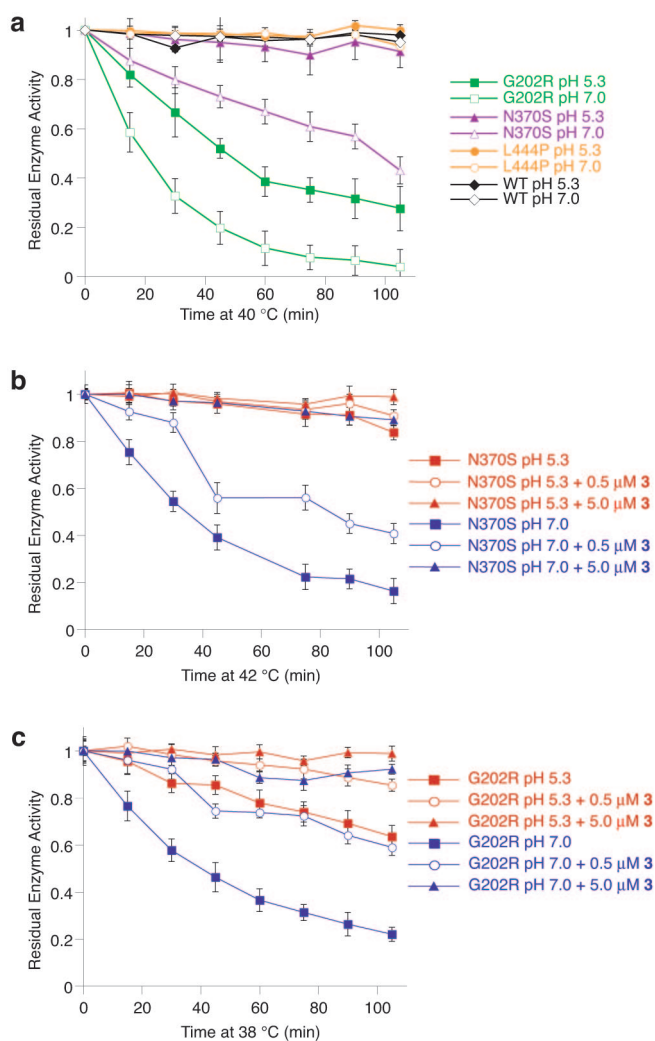


Figure 4. The pH sensitivity of GC variant conformational stability probed by activity. a) Cell lysates were incubated at 40 °C, and the residual activity was measured at various time points and compared to the activity of lysates maintained at 4 °C. An activity of 1 represents activity equal to lysates from the same cell line maintained at 4 °C. b) N370S lysates were heated at 42 °C as a function of the concentration of 3. The residual activity was compared to the activity of lysates maintained at 4 °C containing the same concentration of chemical chaperone. c) G202R lysates were heated at 38 °C as a function of the concentration of 3. The residual activity was compared to the activity of lysates maintained at 4 °C containing the same concentration of 3.

experiments on recombinant proteins are substantiated and extended by the fibroblast lysate experiments, supporting the notion that the GC variants are destabilized in the ER.

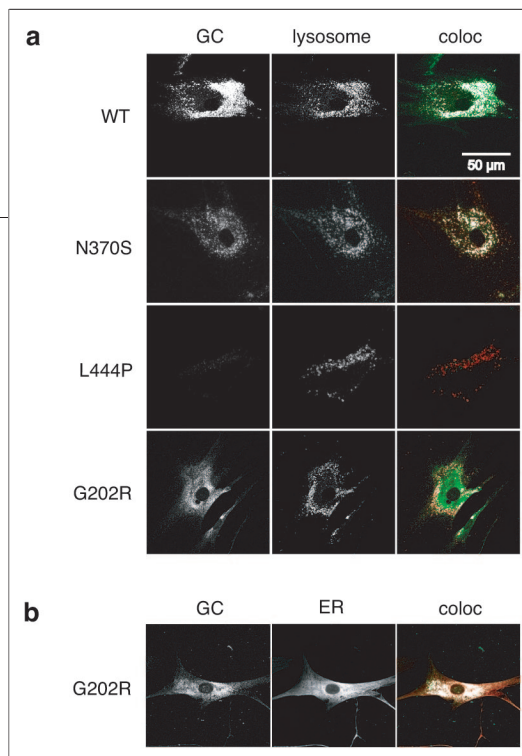


Figure 5. The subcellular localization of GC variants. **a)** Immunofluorescence colocalization was performed using mouse anti-GC (column 1) and rabbit anti-LAMP2 (column 2) as a lysosome marker. In column 3, the colocalization of GC (green) and LAMP2 (red) is shown in white. **b)** Immunofluorescence colocalization was performed using mouse anti-GC (column 1) and rabbit anti-calnexin (column 2) as an ER marker. In column 3, the colocalization of GC (green) and calnexin (red) is shown in white.

Chemical Chaperone Binding Stabilizes pH-Sensitive GC Variants *in Vitro*. A variety of GC active site-binding chemical chaperones increase the cellular activity of GC in patient-derived fibroblasts including **1** (25) and **2–4** (Figure 2) (10, 11). Herein, we evaluate the ability of the established chemical chaperone, *N*-hexanoic acid adamantyl amide deoxynojirimycin (Figure 2, structure **3**) (11), to stabilize the N370S and G202R GC variants exhibiting pH 7 instability in patient-derived fibroblast lysates. Chemical chaperone **3** stabilizes both of these variants in a dose-dependent manner at pH 5.3 and especially at pH 7 (Figure 4, panels b and c). The incubation temperature was optimized for each variant. The relatively stable N370S variant was heated to a higher temperature (42 °C) to increase the amount of denaturation observed in the absence of chemical chaperone (Figure 4, panel b). Since this variant is fairly stable at a lysosome simulating pH (5.3), the added stabilization bestowed on the catalytic domain by the binding of a chemical chaperone is quite small (Figure 4, panel b, red lines). However, the addition of 0.5 μM **3** notably increases the stability of the N370S catalytic domain at an ER simulating

pH (Figure 4, panel b, pH 7.0; *cf.* closed blue squares to open blue circles). Dramatically, the denaturation of N370S GC at pH 7 is negligible in the presence of 5.0 μM **3** (Figure 4, panel b, closed blue triangles). Since the catalytic domain of the G202R variant is less stable than that of N370S, the samples were heated at 38 °C (Figure 4, panel c). The denaturation of G202R GC at pH 5.3 is greatly reduced in the presence of 0.5 μM **3** and is abolished in the presence of 5.0 μM **3** (Figure 4, panel c, red lines). While G202R GC is substantially less stable at pH 7.0, the stability of the catalytic domain is dramatically increased at this pH in the presence of **3** in a dose-dependent fashion (Figure 4, panel c, closed blue triangles and open blue circles versus closed blue squares). In conclusion, the GC binding of **3** compensates for the neutral pH instability of both N370S and G202R GC *in vitro*, which should increase the amount of folded GC in the ER, reducing ERAD, consistent with the increased trafficking/activity demonstrated below.

Chemical Chaperone Binding Increases

Recombinant WT-GC and N370S-GC Stability. Recombinant WT and N370S GC were thermally denatured in the presence of chemical chaperone **3** to quantify GC stabilization. Addition of **3** (Figure 3) increases the T_m of rhWT-GC and rhN370S-GC at pH 5.3 and pH 7.0 in a dose-dependent manner (Figure 3, panels b and c). The T_m increases of rhWT-GC at neutral and acidic pH are quite similar at each chaperone concentration (Figure 3, panel b), whereas the T_m increases of rhN370S-GC are higher at pH 7.0 than at pH 5.3, reflecting the high pH instability of this variant (Figure 4, panel c). These data demonstrate that chemical chaperones stabilize the native state of GC at neutral pH. Significantly, binding of chemical chaperone **3** (4.8 μM) at pH 7 increases the T_m of rhN370S-GC by 4 °C, resulting in a stability that is comparable to that of rhWT-GC (Figure 3, panel d, blue triangles versus black squares). It seems likely that this stabilization would lead to more N370S-GC being able to exit the ER, thus reducing the extent of ER-associated protein degradation, consistent with the observed trafficking and cellular activity increases discussed below.

Cellular Distribution of GC Variants in Patient-Derived Fibroblasts. The subcellular distribution of the aforementioned GC variants in patient-derived fibroblasts was visualized using immunofluorescence microscopy (see Methods). WT-GC distributes in a punc-

tate manner consistent with lysosomal localization. This was verified by the colocalization of WT-GC with the lysosome marker LAMP2 (Figure 5, panel a, overlap shown in white, column 3). The N370S-GC distribution looks similar to that of WT and is predominantly lysosomal (Figure 5, panel a), with the exception that there is significantly less N370S-GC in the lysosome and in the cell in general, presumably as a result of a greater extent of ERAD. Since N370S is the predominant Gaucher disease associated mutation, quantifying both the concentration and the activity of N370S-GC in patient-derived fibroblast lysosomes, with and without the application of chemical chaperones, is a short term goal that requires further technology development. The G202R-GC distribution pattern is more diffuse than that exhibited by WT or N370S (Figure 5, panel a). Importantly, only a small portion of the G202R-GC signal overlaps with the lysosomal marker LAMP2. Instead, G202R shows significant colocalization with the ER marker calnexin (Figure 5, panel b). This observation is consistent with previous reports that this variant has difficulty exiting the ER (8, 11, 21). As a consequence of the marked trafficking defects exhibited by the G202R variant, this variant proves very useful for demonstrating the chaperone restoration of proper trafficking (see below) as well as permissive temperature enhanced folding and trafficking, also demonstrated below. The subcellular distribution of L444P-GC could not be discerned with confidence under these conditions, likely due to very low L444P levels in the cell (Figure 5, panel a), probably as a result of extensive ERAD.

A Subset of GC Variants Are Temperature Sensitive for Folding in Fibroblasts. Many proteins that are deficient in folding and subjected to ERAD display improved folding and trafficking when cells are grown at a reduced “permissive” temperature (26–31). In order to explore the possibility that the intracellular folding and trafficking of GC variants is temperature sensitive (TS), we cultured Gaucher disease patient fibroblasts at 30 °C for 7 days. The GC activity of WT cells grown at this temperature is similar to that of cells maintained at 37 °C (data not shown). The GC activity of the N370S cells is similarly insensitive to culture temperature (Figure 6, panel a, blue versus black bar); looking at other N370S cell lines is required to discern the generality of this result. In contrast, the GC activity of the G202R cells proved to be notably sensitive to culture temperature. Cells grown at 30 °C have 4-fold higher GC activity than

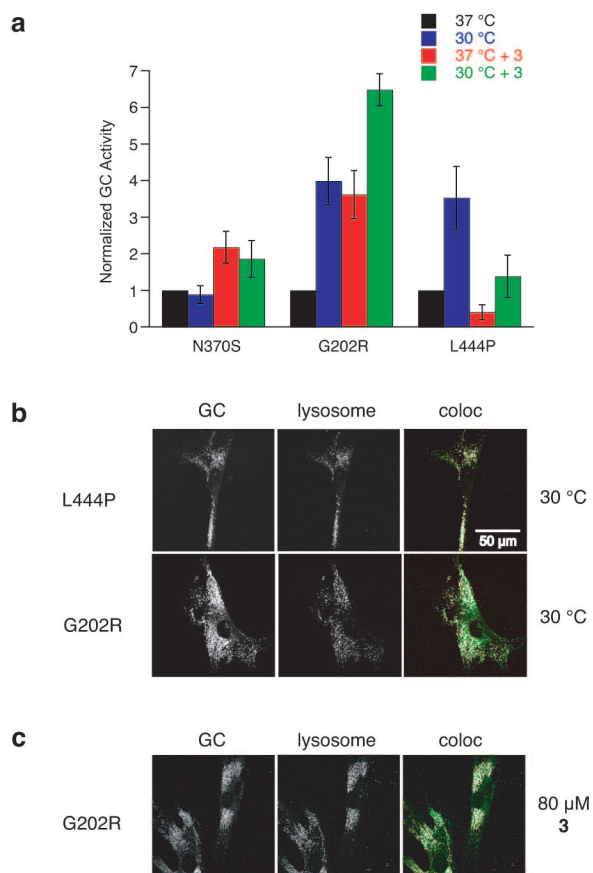


Figure 6. Temperature and chemical chaperone dependent GC folding and trafficking. **a**) The relative activity of GC variant expressing fibroblasts in response to reduced temperature culturing, culturing with chemical chaperone 3, or a combination thereof. The normalized activity of GC variants subjected to permissive temperature growth, chaperone 3, or a combination thereof compared to the activity of untreated cells harboring the same mutation. The activity of untreated cells (black) was normalized to 1. The activity of cells grown for 7 days with 80 μM 3 is shown in red. The activity of cells grown for 7 days at 30 °C (permissive temperature) is shown in blue. The activity of cells grown for 7 days with 80 μM 3 at 30 °C is shown in green. **b**) Immunofluorescence colocalization of G202R and N370S GC in cells grown at 30 °C for 7 days. Protein was visualized using mouse anti-GC antibody (column 1) and rabbit anti-LAMP2 antibody (column 2) as a lysosome marker. In column 3, the colocalization of GC (green) and LAMP2 (red) is shown in white. **c**) Immunofluorescence colocalization of G202R GC in cells grown with 80 μM chemical chaperone 3 for 7 days. Protein was visualized using mouse anti-GC antibody (column 1) and rabbit anti-LAMP2 antibody (column 2) as a lysosome marker. In column 3, the colocalization of GC (green) and LAMP2 (red) is shown in white.

cells grown at 37 °C (Figure 6, panel a, blue versus back bar). Notably, the cellular activity of the L444P-GC variant in fibroblasts is also similarly increased at low temperature (Figure 6, panel a, blue versus black bar), a striking result, given that this variant has not been responsive to treatment by various chemical chaperones (10, 11, 32) and was thought to be severely misfolded, at least with respect to the Ig domain.

Distribution of GC Variants in Cells Grown at the Permissive Temperature. In order to assess whether the TS-GC variants exhibit altered subcellular distribution at the permissive temperature, cells grown at 30 °C were analyzed by immunofluorescence microscopy in comparison to cells grown at 37 °C. The distribution of both G202R- and L444P-GC in cells grown at the permissive temperature (Figure 6, panel b) is strikingly different than the distribution of these variants in cells grown at 37 °C (Figure 5). Whereas G202R-GC immunofluorescence is diffuse in cells grown at 37 °C and is ER localized, the cells grown at the permissive temperature contain punctate GC immunofluorescence that substantially colocalizes with the lysosome marker LAMP2 (Figure 6, panel b, third column). While the L444P variant is not visible in cells grown at 37 °C, it is easily detected, punctate, and is observed to colocalize with lysosomes in cells grown at the permissive temperature. In summary, growing variant fibroblasts at 30 °C (the permissive temperature) leads to increased trafficking of the G202R- and L444P-GC variants to the lysosome, presumably as a result of increased ER folding efficiency.

The Interplay between Growth at the Permissive Temperature and Chemical Chaperoning. We have previously shown that N370S- and G202R-GC expressing fibroblasts are amenable to chemical chaperoning (10, 11). Culturing N370S fibroblasts in the presence of the active-site-directed chemical chaperone **3** leads to a greater than 2-fold increase in the total GC activity of these cells (Figure 6, panel a, red versus black bar). The total GC activity of G202R is also increased (4-fold) in the presence of **3** (Figure 6, panel a, red versus black bar). The total GC activity of L444P is not increased in the presence of **3**, but rather it is inhibited (Figure 6, panel a, red versus black bar). Small molecules that chaperone N370S- and G202R-GC in patient-derived fibroblasts either are inhibitory or have no effect on L444P fibroblasts (Figure 6, panel a) (11),

consistent with the idea that the L444P mutation destabilizes the Ig-like domain, but not the catalytic domain.

We also grew cells at the permissive temperature in the presence of chemical chaperone to investigate whether the TS and chaperone effects are additive. Recall that the N370S cells do not exhibit TS. Moreover, N370S cells grown with **3** at 30 °C have the same GC activity as cells grown with **3** at 37 °C (Figure 6, panel a, green versus red bar). G202R cells grown with chemical chaperone at 30 °C (Figure 6, panel a, green bar) have roughly twice the GC activity of cells grown at the permissive temperature (30 °C) without chemical chaperone or cells grown at 37 °C with chemical chaperone **3** (Figure 6, panel a, green versus red and blue bars). In this case, the effects of the treatments appear to be roughly additive. These data imply that **3** is a suboptimal chaperone for G202R-GC. L444P cells grown at 30 °C with chemical chaperone have lower GC activity than cells grown at 30 °C without chemical chaperone (Figure 6, panel a, green versus blue bar), suggesting that the chemical chaperone is acting as an inhibitor of the L444P-GC that does make it to the lysosome.

Altered Cellular Distribution of G202R-GC in Cells Treated with Chemical Chaperone. While we would like to demonstrate that the increase in N370S lysosomal GC activity in patient-derived fibroblasts in response to chemical chaperone treatment is a result of enhanced folding within and trafficking out of the ER, this is difficult because N370S is found in both the ER and the lysosome in the absence of chemical chaperone treatment. Precise quantification of the amount of N370S-GC in each subcellular location is needed to demonstrate increased folding and trafficking in the presence of a chemical chaperone, a feat we hope to be able to perform soon with the more sensitive methods under development. In the absence of these, immunofluorescent microscopy examination of the subcellular distribution of G202R-GC in fibroblasts treated with **3** for 7 days (Figure 6, panel c) versus chaperone-free control cells is easy to interpret because G202R barely reaches the lysosome in the absence of a chemical chaperone (Figure 5, panels a and b). The chemical chaperone-treated G202R cells contain punctate GC immunofluorescence that colocalizes with the lysosome marker LAMP2, in contrast to the ER localization found in cells grown at 37 °C in the absence of **3** (cf. Figure 6, panel c and Figure 5, panels a and b). The G202R-GC distribution in cells grown in the presence of chemical chap-

erone looks like the G202R-GC distribution in cells grown at the permissive temperature (Figure 6, panel b, 30 °C). Treatment of L444P fibroblasts with **3** did not alter L444P-GC trafficking (data not shown) consistent with the folding defect being in the Ig-like domain. Importantly, these data and the results outlined in the next section provide the first direct evidence that chemical chaperone treatment increases the lysosomal trafficking of a Gaucher disease associated GC variant, leading to increased lysosomal enzyme concentrations that may be sufficient to ameliorate Gaucher disease.

GC Oligosaccharide Processing Changes in Cells Treated with Chemical Chaperone. The glycosylation state of GC is altered as the protein matures through folding and is trafficked, offering an opportunity to monitor GC trafficking. Nascent GC is glycosylated in the ER affording a 62 kDa mannose-rich GC precursor. As the protein matures in the Golgi, reflecting departure from the ER, and trafficking toward the lysosome, the precursor is converted into a complex-glycosylated 66 kDa protein (complex GC). The mannose-rich GC precursor found in the ER is sensitive to complete deglycosylation by the enzyme endo-H, in contrast to the mature complex-glycosylated protein formed in the Golgi which is endo-H resistant (12).

WT GC cells contain a mixture of complex-glycosylated (mature) and mannose-rich GC (ER form), which results in a broad smear when the protein is analyzed by a Western blot (Figure 7, lane 1). Endo-H treatment of a WT-GC lysate reveals that the endo-H sensitive and resistant GC forms of GC are both present (Figure 7, lane 2). When G202R-GC cells are analyzed in a similar manner, the result is a compact band corresponding to the mannose-rich precursor consistent with its ER retention (Figure 7, lane 3), which upon endo-H treatment results in a single, endo-H-sensitive band (Figure 7, lane 4). Collectively, these results suggest that very little mature GC is present in homozygous G202R cells and that the protein that is present is the immature ER precursor form, fully consistent with the immunofluorescence microscopy results presented above. However, the glycosylation pattern of G202R-GC changes when the cells are cultured with the chemical chaperone **3** (80 μ M) for 7 days. The protein is detected by Western blot as a molecular weight distribution (Figure 7, lane 5) and endo-H treatment reveals the presence of both the endo-H sensitive and resistant G202R-GC (Figure 7, lane 6). While chaperoned G202R cells clearly produce

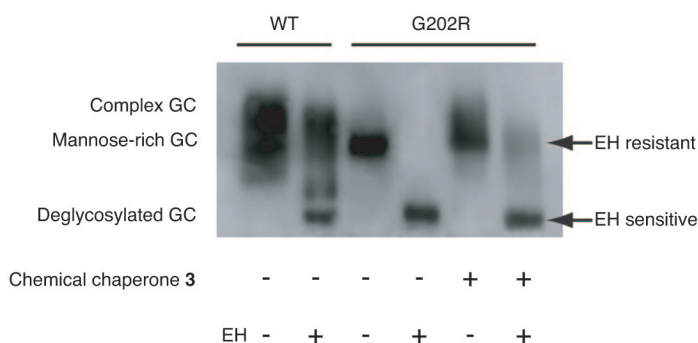


Figure 7. The endo-H (EH) sensitivity of WT and G202R-GC as a means of following GC trafficking. G202R cells were cultured in the presence and absence of 80 μ M chemical chaperone **3** for 7 days. GC was isolated by immunoprecipitation using mouse anti-GC and then subjected to EH treatment, SDS-PAGE, and Western blot analysis. WT cells cultured without chemical chaperone were subjected to the same treatment for comparison. Data from a representative experiment are shown.

less mature GC than WT cells, chemical chaperone treatment increases the trafficking of the G202R-GC from the ER as demonstrated by the presence of endo-H resistant complex glycosylated band.

The Activity of a Chemical Chaperone *in Vitro* Does Not Predict Its Efficacy *in Vivo*. In order to ascertain whether cellular chemical chaperone potency and the extent of native state stabilization *in vitro* are related, a series of deoxynojirimycin analogues (Figure 2) were incubated with homozygous WT-GC and N370S-GC cells for 5 days. The candidate chemical chaperones displayed a range of chaperoning activities as a function of inhibitor concentration, with **2** and **3** exhibiting the best activity (Figure 8, panels a and b). Since the deoxynojirimycin analogues are targeted to the GC active site, the compounds were expected to become inhibitory at sufficiently high concentrations. While the relative chaperoning activity of the compounds was higher in N370S cells than in WT cells (note that even some of the WT-GC is subjected to ERAD as surmised by its increase in activity in the presence of chemical chaperone), the activity trends are similar (*cf.* Figure 8, panel b vs panel a). The butyl analogue **1** (Figure 8, panels a and b, gray circles) has no activity in either cell line, despite reports that it is a chaperone (25). The nonyl analogue **2** (Figure 8, panels a and b, blue squares) is a modest chaperone, increasing cellular WT-GC activity by \approx 20% and N370S-GC activity by \approx 60%. The adamantyl amide analogue **3** (Figure 8,

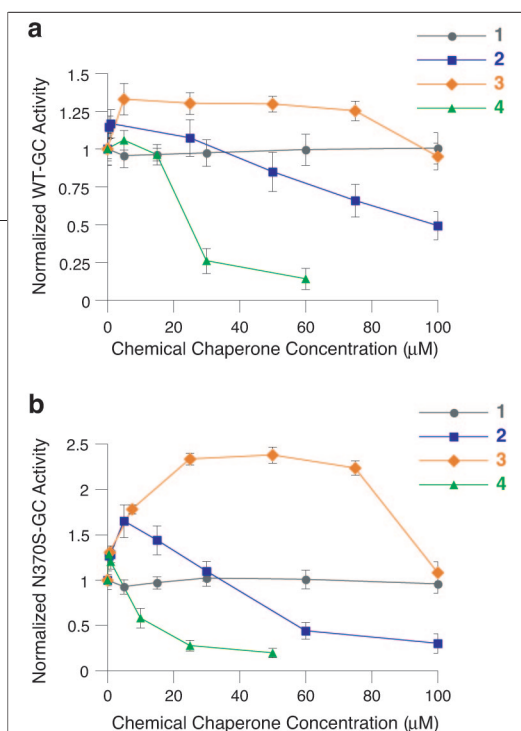


Figure 8. Chemical chaperoning of WT and N370S GC. Chemical chaperoning profiles of deoxynojirimycin-based compounds utilizing a) WT and b) N370S patient-derived fibroblasts. Cells were treated with the indicated compound for 5 days before cellular GC activity levels were determined. Levels above 1 represent elevated activity relative to untreated cells. Data shown are the average of three experiments, and the error bars correspond to the standard deviation. Lines have been added to guide the eye.

at concentrations above 10 μM . The media concentration at which a chemical chaperone becomes inhibitory is generally higher than its IC_{50} value, most likely due to cell permeability issues.

Imiglucerase, rhWT-GC, and rhN370S-GC were subjected to thermal denaturation in the presence of

panels a and b, orange diamonds) is one of the best chemical chaperones tested to date, increasing cellular WT-GC activity by $\approx 30\%$ and N370S-GC activity by $\approx 140\%$ over a broad concentration range. It is not functionally inhibitory up to a concentration of 80 μM . The adamantyl ether analogue **4** (Figure 8, panels a and b, green triangles) is slightly activating at low μM concentrations but becomes inhibitory

these deoxynojirimycin analogues (T_m values are summarized in Table 2). The T_m values of imiglucerase at pH 5.3 and pH 7.0 are slightly higher than those of rhWT-GC, possibly due to differences in glycan structure, but the trends are similar. Compound **1** exhibits no significant influence on the T_m of imiglucerase or rhWT-GC (only tested at pH 7.0, 48 μM in the case of rhWT-GC), consistent with its inability to chaperone in our hands. On the basis of these results, **1** was not tested with rhN370S-GC due to limited quantities of this GC variant. Compounds **2**, **3**, and **4** exhibit similar T_m increases despite their different cellular chemical chaperoning profiles. We did not expect the *in vitro* stability data to correlate with or to predict the cellular chaperone efficacy because the compound's cellular permeability, metabolic half-life, and intracellular distribution, among other features, were expected to differ.

The changes in imiglucerase and rhWT-GC T_m values do not seem to be pH dependent (within the standard deviation of measurement). Compounds **2**, **3**, and **4** increase the stability of rhN370S-GC at both neutral and acidic pH. While the absolute stability is higher at pH 5.3, the net stability increase due to chemical chaperone binding is greater at pH 7.0. The slight differences in T_m (all chemical chaperones excepting **1** confer WT-like stability to rhN370S-GC at neutral pH) and IC_{50} values (Table 3) do not satisfactorily explain the differences in the chemical chaperoning profiles of **2**, **3**, and **4**, (Figure 8, panels a and b). It is more likely that

TABLE 2. T_m values for the thermal denaturation of imiglucerase, rhWT-GC, and rhN370S-GC in the presence of deoxynojirimycin analogues^a

Chemical chaperone	T_m ($^{\circ}\text{C}$) Imiglucerase pH 5.3	T_m ($^{\circ}\text{C}$) Imiglucerase pH 7.0	T_m ($^{\circ}\text{C}$) rhWT pH 5.3	T_m ($^{\circ}\text{C}$) rhWT pH 7.0	T_m ($^{\circ}\text{C}$) rhN370S pH 5.3	T_m ($^{\circ}\text{C}$) rhN370S pH 7.0
None	50.5 \pm 0.4	51.1 \pm 0.4	49.3	49.2	48.5	45.0
4.8 μM 1	51.3	51.4	–	–	–	–
4.8 μM 2	53.3	53.7	–	51.9	49.6	47.8
4.8 μM 3	54.3	54.0	52.6	52.8	50.8	49.0
4.8 μM 4	54.1	54.0	–	–	–	–
48 μM 1	51.7	51.5	–	49.3	–	–
48 μM 2	56.9	57.1	55.8	55.8	53.5	51.7
48 μM 3	58.5	58.2	57.3	57.7	53.8	52.4
48 μM 4	58.2	57.9	56.9	56.5	53.3	50.6

^aExperiments done at pH 7.0 and 5.3. The T_m values were obtained by interpolation based on the best-fitted curves (Microsoft Excel) and standard deviations were less than 0.4 $^{\circ}\text{C}$ for three independent experiments.

TABLE 3. IC₅₀ values of deoxynojirimycin analogues tested as candidate GC chaperones^a

Chemical chaperone	Imiglucerase IC ₅₀ (μM)	rhWT-GC IC ₅₀ (μM)	rhN370S-GC IC ₅₀ (μM)
1	>1000	>1000	>1500
2	1.0–3.0	2.0–5.0	3.0–5.0
3	0.5–0.8	0.9–1.5	0.9–1.5
4	0.7–0.9	0.8–1.5	1.0–2.5

^aIC₅₀ values were determined using the same assay conditions described in Table 1.

intracellular ER and/or lysosomal concentration differences or differences in chaperone cellular metabolism explain the observed range of chaperoning efficacy, and we posit that this will be the typical relationship between *in vitro* and in cell data. Besides high affinity binding, an efficacious inhibitor has to be cell and ER permeable, it has to have a reasonable intracellular half-life, and ideally it would not concentrate in the lysosome.

Lysosomal Stability of Trafficked G202R- and L444P-GC Variants. The stability of the mutant GC enzymes in the lysosome is likely to be greater than those in the other subcellular compartments due to the low pH for which the sequences were evolutionarily optimized and due to the high concentrations of stabilizing molecules such as the substrate and protein-based activators. In order to examine the stability of mutant GCs in the context of the lysosome, the TS-GC variant fibroblasts were equilibrated at 30 °C for 7 days and then shifted to 37 °C before determining GC activity as a function of time. The amount of time that it takes for the elevated GC activity resulting from permissive growth to return to the basal activity of cells grown at 37 °C (normalized here as an activity of 1) is related to the lysosomal stability of the enzyme. An enzyme that is unstable in the lysosome should lead to rapid re-equilibration to basal activity levels (Figure 9, dashed lines), whereas a stable enzyme should retain elevated activity levels for days, as the half-life of WT GC is ~60 h (33). Both of the TS GC variants examined (G202R and L444P) appear to have similar stabilities in the lysosome, comparable to the stability reported for WT GC (33). Importantly, these data show that at least a subset of the GC variants are stable in their desti-

nation environment. It would be undesirable to utilize chemical chaperoning to enable lysosomal trafficking if the GC variants rapidly misfold in the lysosome after chaperone dissociation. Superior recombinant N370S stability at pH 5.3 relative to pH 7 (Figure 3, panel c) also predicts that this variant will likely be stable in the lysosome after chemical chaperone dissociation.

Implications for Treating Gaucher Disease The N370S-, G202R-, and L444P-GC mutations appear to reduce lysosomal GC activity via distinct, but related, misfolding and mistrafficking mechanisms. Manipulating the unique folding environment of the ER can restore variant GC function by increasing the extent of ER folding, ER departure, and trafficking of variant GC to the lysosome, thus increasing the concentration of a partially active GC variant there.

The ER plays a central role in the folding of non-cytoplasmic proteins. The efficiency of folding in the ER is mediated by numerous factors including proteins such as molecular chaperones, folding enzymes including disulfide isomerases, and small molecules such as osmolytes (19, 34, 35). Proteins that are unable to adopt a native conformation in the ER are bound in a sustained fashion by one or more chaperones, triggering ERAD (36, 37). The mechanism by which these molecular chaperones discriminate between folded and misfolded proteins is still under investigation and has important implications for disease. Numerous familial loss of function diseases are associated with mutations that lower the conformational stability of the protein in the ER and to a lesser degree in the destination environment. Many of these variants retain enough activity that if they could reach their destination environment, they would restore function (38). Unfortunately, the quality control machinery retains these variant proteins in the ER and targets them for degradation (39).

The recognition and destruction of variant enzymes that are catalytically inactive, are highly destabilized, and/or are unable to properly oligomerize is clearly beneficial to the cell, as buildup of these nonfunctional proteins would likely be toxic (38). In the case of variant enzymes that retain partial catalytic activity, it remains unclear whether ERAD is benefiting the cell if degradation ultimately causes disease by loss of function. Both G202R- and N370S-GC are destabilized at the neutral pH found in the ER relative to WT-GC but are significantly more stable and wild type-like at the acidic pH of their

destination environment, *i.e.*, in the lysosome. It is generally accepted that the conformational stability of a protein is linked to its ability to exit the ER (40–43). While it is likely that the energetics of the protein fold and the cell-specific environment (including small molecule, protein chaperone, and folding enzyme distribution) all play a role in determining which proteins are exported from the ER, our data demonstrate that quality control decisions made in the ER cannot reflect the behavior of a protein in its functional or destination environment (19). In the case of the pH-sensitive lysosomal hydrolases studied herein, ER “quality control” is a bit of a misnomer, in that ERAD can degrade functional hydrolases, which leads to a loss of function disease.

Since temperature changes influence the folding energy landscape of proteins and the physiology of the host cell, growing cells at reduced temperature has been found to facilitate more efficient folding of variant proteins (26–31). At the cellular level, the response of mammalian cells to reduced temperature is quite complex: transcription, translation, and metabolic processes are reduced and the lipid composition of

membranes is altered. Many of the net cellular physiological effects of cold exposure (reviewed in refs 44 and 45) are similar to those seen in heat-stressed cells; hence, growing cells at a reduced temperature has global effects beyond simply altering protein folding energetics. While many proteins that are deficient in folding at 37 °C show improved trafficking at a reduced “permissive” temperature, G202R- and L444P-GC are the first examples of temperature-sensitive (TS) GC enzymes. TS protein variants may be grouped into two classes: (1) proteins that remain thermolabile in post-ER compartments and (2) proteins that are thermostable after reaching their destination environment. The Δ F508 cystic fibrosis (CF) associated mutation in the CF transmembrane conductance regulator (CFTR) is an example of the former. Organic solutes rescue Δ F508 trafficking to the cell surface by stabilizing the mature form of the protein (46). At temperatures below 30 °C, the Δ F508 CFTR ion channel is also trafficked properly, and chloride channel activity is detected (26). However, rescued Δ F508 CFTR has a short residence time at the cell surface at nonpermissive temperatures, implying that the TS folding defect persists in the mature, trafficked protein at the destination environment and that the native state of Δ F508 CFTR appears to have a lower stability than WT CFTR at the cell surface (47). TS variants of the P22 tailspike protein fail to reach a native state conformation at 37–42 °C but traffic properly at 28 °C. In contrast to Δ F508 CFTR, once the TS tailspike variants reach maturity at 28 °C, they remain native when the cells are transferred to a nonpermissive temperature (31). The G202R and L444P TS GC variants are also members of the second group, in that they are sufficiently stable and active in their destination environment at 37 °C. The biophysical basis for the failure of these GC variants to fold and exit the ER merits careful investigation. These results demonstrate that native-like stability of a variant protein in its destination environment is not sufficient to ensure export from the ER.

Chemical chaperones have been used to increase the cellular activity of variant lysosomal enzymes associated with Fabry (39, 48, 49), Gaucher (10, 11, 32), Tay-Sachs, and Sandhoff diseases (50) as well as G_{M1} -gangliosidosis (51). The different responses of N370S-, G202R-, and L444P-GC patient-derived cell lines to chemical chaperone treatment reflect distinct mutation-dependent GC folding defects in the ER. Both the N370S

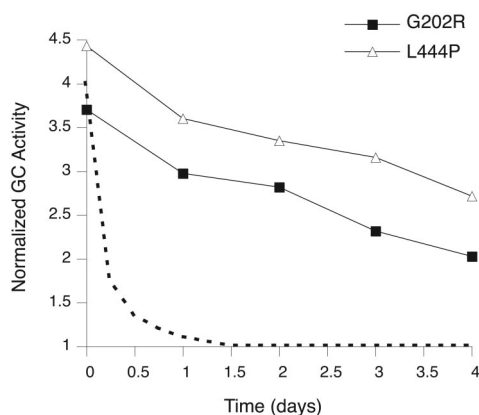


Figure 9. The stability of GC trafficked to the lysosome. The temperature-sensitive GC variant G202R (filled squares) and L444P (open triangles) fibroblasts were cultured at 30 °C for 7 days before returning the cells to 37 °C for the indicated amount of time. The activity of the cells was measured and normalized to the activity of variant cells maintained at 37 °C. An activity of 1 represents the GC activity of cells of a given type grown at 37 °C. The amount of time required to return to an activity of 1 is related to the lysosomal stability of the variant. A hypothetical curve representing an unstable protein is shown for comparison (dashed line).

and G202R mutations compromise the folding of the active site domain of GC in the ER, resulting in proteins that are amenable to chaperoning by active-site-directed small molecules. While culturing N370S fibroblasts at the permissive temperature does not increase total GC activity (at least in this cell line), chemical chaperone **3** is able to correct the neutral pH sensitivity of this variant and increase its folding stability and export from the ER as reflected by the increase in lysosomal activity. In the case of G202R-GC, the effects of chemical chaperone and reduced temperature culture conditions are roughly additive, suggesting that neither of these treatments is optimized and that better chemical chaperones for G202R-GC may further improve trafficking.

Asano and co-workers (49) have previously suggested that the inhibitory potency of a chemical chaperone for α -galactosidase A *in vitro* is directly related to its chemical chaperoning potency *in vivo*, but this trend does not hold in the context of variant GC chaperoning by **2–4** (Figure 8). The slight differences in T_m (all these chemical chaperones confer WT-like stability to rhN370S-GC at neutral pH) and IC_{50} values do not satisfactorily explain the large differences in the chemical chaperoning profiles of **2**, **3**, and **4**. The intracellular ER and/or lysosomal concentration differences or differences in chemical chaperone cellular metabolism are expected to dramatically influence chaperoning efficacy. Thus, chaperone binding to and stabilization of GC are necessary but are not sufficient for intracellular chaperoning efficacy. At a minimum the compound has to be cell and ER permeable, it has to have a reasonable intracellular half-life, and ideally it would not concentrate in the lysosome.

The L444P mutation is unique among the GC variants characterized in this paper in that this variant destabilizes the Ig-like domain of GC rather than the catalytic domain destabilized by the N370S and G202R mutations. Binding of chemical chaperone **3** (Figure 2) to the active site of L444P-GC does not correct the folding defect of this protein, in contrast to the dramatic rescue of this variant at a permissive temperature. In fact, inhibition is observed when L444P expressing cells are cultured with chaperone at the permissive temperature (Figure 6, panel a), strongly suggesting that the stability of the Ig-like and catalytic domains are not thermodynamically linked and that chaperone binding to the cata-

lytic domain inhibits function, probably because there is so little L444P-GC present in cells cultured at 30 or 37 °C. The inhibition of L444P-GC provides strong evidence that **3** is indeed present in the lysosome. In the case of the other GC variants, the efficacy of a chemical chaperone (*e.g.*, **3**) in improving enzyme trafficking must be substantial such that the inhibitory action of the compounds is overwhelmed by quantity of the enzyme that is rescued and is trafficked to the lysosome.

Others have reported that L444P-GC fails to be chaperoned by small molecules that rescue the activity of the F213I-GC variant (32), leading to the suggestion that the proline substitution results in an unsalvageable protein conformation. The characterization of the temperature-sensitive nature of this mutation indicates, however, that this is not the case. It may be possible to develop chemical chaperones that bind specifically to the Ig-like domain of GC, which should substantially stabilize the L444P variant.

Even though the GC variants studied herein are less catalytically active than WT-GC, increasing their concentration in the lysosome by chemical chaperoning and/or a permissive growth temperature may be sufficient to ameliorate Gaucher disease, based on the apparent 5% activity threshold that is thought to result in glucosylceramide storage (4). The data herein reveal that the N370S mutation may cause more problems than lowered activity and ER instability, for example, lysosomal activation by saposin C binding, not evaluated herein, also be compromised. Even though there is more to learn about the basis of GC mutant disease etiology, we demonstrate herein and elsewhere (10, 11) that it is likely that we can increase N370S- and G202R-GC activity levels to more than 5% of that exhibited by WT-GC, providing optimism that we can ameliorate Gaucher disease with this strategy.

In summary, the data demonstrate that the N370S, G202R, and L444P Gaucher disease associated mutations all compromise GC folding in the ER, lowering the amount of GC trafficked to the lysosomes. Fluorescence microscopy and oligosaccharide processing analysis demonstrate for the first time that chemical chaperones increase variant GC trafficking to the lysosome, and temperature shift experiments and/or biophysical data confirm that variant GCs are conformationally stable in this unique environment. While the L444P destabilizing mutation in the Ig domain of GC cannot be chaperoned by the active site-directed small molecules discovered

thus far that stabilize the active site domain, it may ultimately be possible to develop chemical chaperones that stabilize the Ig domain. It is becoming clear that manipulating the environment of the ER with small

molecules can enable proper folding and trafficking of disease-associated GC variants, offering optimism that these data will ultimately translate into new therapeutics for lysosomal storage disorders.

METHODS

Cell Cultures. Primary skin fibroblast cultures were established from patients homozygous for either the N370S (c.1226A>G) mutation or the G202R (c.721G>A) mutation. Type 2 Gaucher disease fibroblasts containing the L444P (c.1448T>C) mutation (GM10915) and apparently normal fibroblast cultures (GM05659, GM00498) were obtained from the Coriell Cell Repositories. Fibroblasts were maintained in minimum essential medium with Earle's salts and nonessential amino acids (Gibco) supplemented with 10% fetal bovine serum and 1% glutamine Pen-Strep (Irvine Scientific) at 37 °C in 5% CO₂. Culture medium was replaced every 3 to 4 days, and monolayers were passaged upon confluency with TrypLE Express (Gibco). All cells used in this study were between the 4th and 18th passages.

Cells that were subjected to temperature shift were grown at 30 °C for 7 days prior to assaying or plating for microscopy. Cells grown on microscope slips were maintained at 30 °C until the cells were fixed and processed.

Cells that were subjected to chemical chaperone treatment were grown in the presence 80 μM of chemical chaperone **3** for 7 days prior to assaying or plating for microscopy. The medium was replaced with fresh drugged medium every 3 days. Cells grown on microscope slips were maintained in drugged medium until the cells were fixed and processed.

Immunofluorescence. Cells grown on cover glass slips were fixed with 3.7% paraformaldehyde in phosphate-buffered saline (PBS, Irvine Scientific). The slips were washed with PBS, quenched with 15 mM glycine in PBS, and permeabilized with 0.2% saponin in PBS. The antibodies were prepared in 0.2% saponin and 5% goat serum in PBS. Cells were incubated with primary antibodies (mouse anti-GC (52) (1:100), rabbit anti-LAMP2 (53) (1:1500), and rabbit anti-calnexin (Stressgen Biotechnologies) for 1 h, washed with 5% goat serum in PBS, and then incubated with secondary antibodies (Alexa Fluor 594 goat anti-mouse IgG and Alexa Fluor 488 goat anti-rabbit IgG) from Molecular Probes for 1 h. Images were captured using a Bio-Rad Radiance 2100 Rainbow laser scanning confocal microscope attached to a Nikon TE2000-U microscope with infinity corrected optics.

The pH Sensitivity of GC Variants. WT, N370S, G202R, and L444P fibroblast lysates were prepared in 0.1 M citrate phosphate (pH 7.0 or 5.3) with 0.1% taurodeoxycholate, 0.1% hydrogenated triton X-100, and protease inhibitor cocktail (Sigma) added. Lysates were maintained at 40 °C for the indicated amount of time and then transferred to 4 °C. The pH of all samples was adjusted to pH 5.3 before the enzyme activity was assayed using 5 mM 4-methylumbelliferyl-β-D-glucopyranoside (Sigma) in the presence of 0.1% taurodeoxycholate, 0.1% hydrogenated triton X-100. Conduritol B epoxide (Toronto Research Chemicals) was used as a control to evaluate the extent of nonspecific activity. The residual activity of the heated samples is reported compared to activity of lysates maintained at 4 °C. An activity of 1 indicates thermal stability under these conditions.

In Vitro Stabilization of GC Variants by **3.** An assessment of the ability of chemical chaperones to stabilize against denaturation was performed using **3**. Cell lysates were prepared as described

above and incubated with 0.5 μM or 5 μM **3** for 30 min prior to heating. The samples were heated at the indicated temperature and assayed as described above. The residual activity of the heated samples is reported compared to the activity of lysates maintained at 4 °C containing the same concentration of chemical chaperone. An activity of 1 indicates thermal stability under these conditions.

Intact Cell Activity Assay. The detergent-free intact cell GC assay has been previously described (10). Cells were grown at 30 °C or in the presence of chemical chaperone for 7 days at 37 °C before assessing GC activity. Quadruplicate samples were prepared for each condition, and two samples were sacrificed to obtain total cell protein (Micro BCA, Pierce) because cell growth at reduced temperature is greatly retarded. The data reported are normalized to the activity of untreated cells grown at 37 °C.

Immunoblotting. The cells were cultured with chemical chaperone **3** (80 μM) for 7 days prior to lyses of the cells in 25 mM Tris-HCl, 50 mM NaCl, pH 8.0, containing 0.5% Na deoxycholate, and 0.5% Triton X-100, and a cocktail of protease inhibitors (antipain, pepstatin, aprotinin, leupeptin, PMSF, and trypsin inhibitor). The cell extracts were immunoprecipitated using the mouse anti-GC antibody 8E4 (54) followed by protein-A-Sepharose at 4 °C (55). Washing of the immunoprecipitates, treatment with endoglycosidase H, and SDS-PAGE (10% acrylamide) were performed essentially as previously described (56). Western Blot analysis was done with antibody 8E4 and anti-mouse-HRP. Detection was done using the ECL system (Amersham Pharmacia).

Temperature Shift Activity Assay. The lysosomal stability of temperature-sensitive GC variants was assessed using a temperature shift assay. G202R and L444P fibroblasts were equilibrated at 30 °C for 7 days. Plates of cells were shifted to 37 °C for 1, 2, 3, or 4 days prior to assaying GC activity using the intact cell assay. Duplicate samples were sacrificed to measure total protein in order to account for temperature-sensitive cell growth. The data reported are normalized to the activity of cells maintained at 37 °C.

Recombinant Enzymes. Imiglucerase. Injection quality rhWT-GC (trade name Cerezyme) was obtained from Genzyme and was reconstituted according to the manufacturers instructions. Imiglucerase is produced in a transformed Chinese hamster ovary cell line and sequentially reacted with neuraminidase, galactosidase, and acetylglucosamidase to expose more mannose residues at the nonreducing ends of the oligosaccharide chains of the glycoprotein to improve lysosomal targeting (57). The amino acid sequence of imiglucerase differs from human placental GC by one amino acid at position 495, where histidine is substituted for arginine.

WT-GC and N370S-GC. Since isolating Chinese hamster ovary cell pools or lines expressing even low levels of the N370S-GC proved to be difficult, both the WT-GC and N370S-GC were expressed in insect cells. The sequence encoding WT-GC (identical to the placental GC sequence) was amplified by PCR from cDNA and subcloned into the pFastBac-1 expression vector (Invitrogen). The N370S mutant GC was subsequently derived from the WT construct using the Quickchange PCR mutagenesis kit (Stratagene) according to the manufacturer's protocols and was used to generate recombinant baculovirus by employing the

Invitrogen Bac-to-Bac system. All recombinant sequences were determined to be free of PCR errors by nucleotide sequence analysis (Sequagen Inc.). rhWT-GC and rhN370S-GC were expressed in the baculovirus expression vector system by infection of Tn-5 cells (Expression Systems) with recombinant virus at an MOI equal to 1. Conditioned medium was harvested 48 h postinfection by centrifugation at 500g and filtered (0.22 μ m). For purification, conditioned medium was loaded onto a butyl Toyopearl 65 °C column (Tosohaas) pre-equilibrated in 20 mM sodium acetate, 150 mM sodium chloride, pH 5.0. The protein was eluted in 50% ethylene glycol in 20 mM sodium acetate, 150 mM sodium chloride, pH 5.0. Fractions enriched in GC were pooled, diluted with distilled water, and then loaded onto a heparin-sepharose fast flow column (AP Biotech) pre-equilibrated in 20 mM sodium acetate, 50 mM sodium chloride, 20% ethylene glycol, pH 5.0. The protein was eluted in a linear gradient from 0.05 to 1.0 M sodium chloride. Glucocerebrosidase-enriched fractions were pooled, diluted with 20% ethylene glycol, and then loaded onto a CM-sepharose column (AP Biotech) pre-equilibrated with 30 mM sodium citrate, 0.01% Tween-80, pH 5.7. Bound enzyme was eluted in 55 mM sodium citrate, 0.01% Tween-80, pH 6.3. rhWT-GC and rhN370S-GC primarily have an oligomannose glycan structure as purified, and no subsequent carbohydrate modification was performed.

Inhibitors/Chemical Chaperones. *N*-(*n*-Butyl) deoxynojirimycin (**1**), *N*-(*n*-nonyl) deoxynojirimycin (**2**), and *N*-(5-adamantane-1-yl-methoxy-pentyl) deoxynojirimycin (**4**) were obtained from Toronto Research Chemicals. The synthesis of deoxynojirimycin analogue **3** has been described (**11**).

Determination of T_m Values for Imiglucerase, rhWT-GC, and rhN370S-GC. Circular dichroism spectra were recorded on an AVIV model 202SF spectrometer using a 0.1 cm path length quartz cell. The cuvette was placed in a jacketed cell holder maintained at the desired temperature. Temperature scans were performed at a heating rate of 2 °C min⁻¹, with a 10 min pause at each temperature prior to measurement for thermal equilibration. Ellipticity at 225 nm was measured as a function of temperature between 25 and 95 °C. The data are plotted as fraction unfolded vs temperature. The heat denaturation of the protein was irreversible, as the protein precipitated under these conditions and did not redissolve upon cooling. Recombinant enzyme (4.8 μ M protein) was dialyzed into 0.1 M citrate-phosphate (pH 7.0 or 5.3) with 0.1% taurodeoxycholate (Calbiochem) and 0.1% hydrogenated triton X-100 (Calbiochem) prior to thermal denaturation studies. The effect of deoxynojirimycin analogues on native state stability was tested at 4.8 and 48 μ M.

Enzyme Activity for Imiglucerase, rhWT-GC, and rhN370S-GC. The activity of imiglucerase, rhWT-GC, and rhN370S-GC (0.01 μ M) was assayed in 0.1 M citrate-phosphate (pH 7.0 or 5.3) with 0.1% taurodeoxycholate and 0.1% hydrogenated triton X-100. 4-Methylumbelliferyl- β -D-glucopyranoside (5 mM) was added, and the samples were incubated for 10, 20, or 30 min prior to quenching with 0.2 M glycine buffer (pH 10.6). The enzyme activity was found to be linear on this time scale. The activity is reported as the number of mol of 4-methylumbelliferone released per minute per mg of purified protein.

Chemical Chaperone IC₅₀ Values for Imiglucerase, rhWT-GC, and rhN370S-GC. IC₅₀ values were determined by preincubating imiglucerase, rhWT-GC, and rhN370S-GC with inhibitors for 15 min in 0.1 M citrate-phosphate buffer (pH 5.3) in the presence of 0.1% taurodeoxycholate and 0.1% hydrogenated triton X-100 at 37 °C before assaying for enzyme activity. Inhibitor concentrations ranged from 10.0 nM to 1.0 mM.

Chemical Chaperone Activity Assay. The intact cell GC assay previously described was used to monitor chemical chaperoning activity (**10**). Briefly, cells were incubated in media treated with inhibitors for 5 days before GC activity was assayed using 5 mM 4-methylumbelliferyl- β -D-glucopyranoside (Sigma) in 0.2 M acetate buffer (pH 4.0). Chemical chaperones were evaluated in triplicate at each concentration and each compound was assayed at least three times. Condurotol B epoxide (Toronto Research Chemicals) was used as a control to evaluate the extent of nonspecific activity. Data are reported normalized to the activity of untreated cells. Total cell protein was measured using the Micro BCA assay reagent (Pierce).

Acknowledgments: This work was supported by grant DK075295 (J.W.K.) of the National Institutes of Health, the National Gaucher Foundation, Gaucher Disease Divot Classic, Grant No. 70, the Skaggs Institute for Chemical Biology, the Lita Annenberg Hazen Foundation, and a National Science Foundation Predoctoral Fellowship (to A.R.S.).

REFERENCES

- Grabowski, G. A. (1993) Gaucher disease. Enzymology, genetics, and treatment. *Adv. Hum. Genet.* **21**, 377–441.
- Beutler, E., Nguyen, N. J., Henneberger, M. W., Smolec, J. M., McPherson, R. A., West, C., and Gelbart, T. (1993) Gaucher disease: Gene frequencies in the Ashkenazi Jewish population. *Am. J. Hum. Genet.* **52**, 85–88.
- Zhao, H., and Grabowski, G. A. (2002) Gaucher disease: Perspectives on a prototype lysosomal disease. *Cell. Mol. Life Sci.* **59**, 694–707.
- Sawkar, A. R., D'Haese, W. D., and Kelly, J. W. (2006) Therapeutic strategies for Gaucher disease, a prototypical lysosomal storage disorder. *Cell. Mol. Life Sci.*, <http://dx.doi.org/10.1007/s00018-005-5437-0>.
- Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S. T., Abeyasinghe, S., Krawczak, M., and Cooper, D. N. (2003) Human gene mutation database (HGMD): 2003 update. *Hum. Mutat.* **21**, 577–581.
- Grace, M. E., Newman, K. M., Scheinker, V., Berg-Fussman, A., and Grabowski, G. A. (1994) Analysis of human acid β -glucosidase by site-directed mutagenesis and heterologous expression. *J. Biol. Chem.* **269**, 2283–2291.
- Nagy, J. K., and Sanders, C. R. (2004) Destabilizing mutations promote membrane protein misfolding. *Biochemistry* **43**, 19–25.
- Schmitz, M., Alfalah, M., Aerts, J. M. F. G., Naim, H. Y., and Zimmer, K.-P. (2005) Impaired trafficking of mutants of lysosomal glucocerebrosidase in Gaucher's disease. *Int. J. Biochem. Cell Biol.* **37**, 2310–2320.
- Ron, I., and Horowitz, M. (2005) ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity. *Hum. Mol. Genet.* **14**, 2387–2398.
- Sawkar, A. R., Cheng, W.-C., Beutler, E., Wong, C.-H., Balch, W. E., and Kelly, J. W. (2002) Chemical chaperones increase the cellular activity of N370S β -glucosidase: a therapeutic strategy for Gaucher disease. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15428–15433.
- Sawkar, A. R., Adamski-Werner, S. L., Cheng, W. C., Wong, C. H., Beutler, E., Zimmer, K. P., and Kelly, J. W. (2005) Gaucher disease-associated glucocerebrosidases show mutation-dependent chemical chaperoning profiles. *Chem. Biol.* **12**, 1235–1244.
- Jonsson, L. M. V., Murray, G. J., Sorrell, S. H., Strijland, A., Aerts, J. F. G. M., Ginns, E. I., Barranger, J. A., Tager, J. M., and Schram, A. W. (1987) Biosynthesis and maturation of glucocerebrosidase in Gaucher fibroblasts. *Eur. J. Biochem.* **164**, 171–179.
- Cormand, B., Vilageliu, L., Burguera, J. M., Balcells, S., Gonzalez-Duarte, R., Grinberg, D., and Chabas, A. (1995) Gaucher disease in Spanish patients: Analysis of eight mutations. *Hum. Mutat.* **5**, 303–309.

14. Lacerda, L., Amaral, O., Pinto, R., Aerts, J., and Sá Miranda, M. C. (1994) The N370S mutation in the glucocerebrosidase gene of Portuguese type 1 Gaucher patients: Linkage to the Pvull polymorphism, *J. Inherited Metab. Dis.* **17**, 85–88.
15. Van Weely, S., Van den Berg, M., Barranger, J. A., Miranda, M. C. S., Tager, J. M., and Aerts, J. M. F. G. (1993) Role of pH in determining the cell-type-specific residual activity of glucocerebrosidase in type 1 Gaucher disease, *J. Clin. Invest.* **91**, 1167–1175.
16. Grace, M. E., Desnick, R. J., and Pastores, G. M. (1997) Identification and expression of acid beta-glucosidase mutations causing severe type 1 and neurologic type 2 Gaucher disease in non-Jewish patients, *J. Clin. Invest.* **99**, 2530–2537.
17. Bemier, V., Bichet, D. G., and Bouvier, M. (2004) Pharmacological chaperone action on G-protein-coupled receptors, *Curr. Opin. Pharm.* **4**, 528–533.
18. Ulloa-Aguirre, A., Janovick, J. A., Brothers, S. P., and Conn, P. M. (2004) Pharmacologic rescue of conformationally-defective proteins: Implications for the treatment of human disease, *Traffic (Oxford, U.K.)* **5**, 821–837.
19. Sekijima, Y., Wiseman, R. L., Matteson, J., Hammarstrom, P., Miller, S. R., Sawkar, A. R., Balch, W. E., and Kelly, J. W. (2005) The biological and chemical basis for tissue-selective amyloid disease, *Cell* **121**, 73–85.
20. Stone, D. L., Tayebi, N., Orvisky, E., Stubblefield, B., Madike, V., and Sidransky, E. (2000) Glucocerebrosidase gene mutations in patients with type 2 Gaucher disease, *Hum. Mutat.* **15**, 181–188.
21. Zimmer, K.-P., Le Coutre, P., Aerts, H. M. F. G., Harzer, K., Fukuda, M., O'Brien, J. S., and Naim, H. Y. (1999) Intracellular transport of acid β -glucosidase and lysosome-associated membrane proteins is affected in Gaucher's disease (G202R mutation), *J. Pathol.* **188**, 407–414.
22. Dvir, H., Harel, M., McCarthy, A. A., Tokar, L., Silman, I., Futerman, A. H., and Sussman, J. L. (2003) X-ray structure of human acid- β -glucosidase, the defective enzyme in Gaucher disease, *EMBO Rep.* **4**, 704–709.
23. Kase, R., Bierfreund, U., Klein, A., Kolter, T., Utsumi, K., Itoh, K., Sandhoff, K., and Sakuraba, H. (2000) Characterization of two α -galactosidase mutants (Q279E and R301Q) found in an atypical variant of Fabry disease, *Biochim. Biophys. Acta* **1501**, 227–235.
24. Ishii, S., Kase, R., Sakuraba, H., and Suzuki, Y. (1993) Characterization of a mutant α -galactosidase gene product for the late-onset cardiac form of Fabry disease, *Biochem. Biophys. Res. Commun.* **197**, 1585–1589.
25. Alfonso, P., Pampin, S., Estrada, J., Rodriguez-Rey, J.-C., Giraldo, P., Sancho, J., and Pocovi, M. (2005) Miglustat (NB-DNJ) works as a chaperone for mutated acid β -glucosidase in cells transfected with several Gaucher disease mutations, *Blood Cells, Mol., Dis.* **35**, 268–276.
26. Denning, G. M., Anderson, M. P., Amara, J. F., Marshall, J., Smith, A. E., and Welsh, M. J. (1992) Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive, *Nature* **358**, 761–764.
27. Zhou, Z., Gong, Q., and January, C. T. (1999) Correction of defective protein trafficking of a mutant HERG potassium channel in human long QT syndrome. Pharmacological and temperature effects, *J. Biol. Chem.* **274**, 31123–31129.
28. Michalovitz, D., Halevy, O., and Oren, M. (1990) Conditional inhibition of transformation and of cell proliferation by a temperature-sensitive mutant of p53, *Cell* **62**, 671–680.
29. Maroney, A. C., Qureshi, S. A., Foster, D. A., and Brugge, J. S. (1992) Cloning and characterization of a thermolabile v-src gene for use in reversible transformation of mammalian cells, *Oncogene* **7**, 1207–1214.
30. Kulka, R. G., Raboy, B., Schuster, R., Parag, H. A., Diamond, G., Ciechanover, A., and Marcus, M. (1988) A Chinese hamster cell cycle mutant arrested at G2 phase has a temperature-sensitive ubiquitin-activating enzyme, E1, *J. Biol. Chem.* **263**, 15726–15731.
31. Sturtevant, J. M., Yu, M. H., Haase-Pettingell, C., and King, J. (1989) Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein, *J. Biol. Chem.* **264**, 10693–10698.
32. Lin, H., Sugimoto, Y., Ohsaki, Y., Ninomiya, H., Oka, A., Taniguchi, M., Ida, H., Eto, Y., Ogawa, S., Matsuzaki, Y., Sawa, M., Inoue, T., Higaki, K., Nanba, E., Ohno, K., and Suzuki, Y. (2004) N-Octyl- β -valienamine up-regulates activity of F2131 mutant β -glucosidase in cultured cells: a potential chemical chaperone therapy for Gaucher disease, *Biochim. Biophys. Acta* **1689**, 219–228.
33. Qi, X., and Grabowski, G. A. (2001) Molecular and cell biology of acid beta-glucosidase and prosaposin, *Prog. Nucleic Acid Res. Mol. Biol.* **66**, 203–239.
34. Ellgaard, L., and Helenius, A. (2003) Quality control in the endoplasmic reticulum, *Nat. Rev. Mol. Cell Biol.* **4**, 181–191.
35. Kim, P. S., and Arvan, P. (1998) Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones, *Endocr. Rev.* **19**, 173–202.
36. Kostova, Z., and Wolf, D. H. (2003) For whom the bell tolls: Protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection, *EMBO J.* **22**, 2309–2317.
37. McCracken, A. A., and Brodsky, J. L. (2003) Evolving questions and paradigm shifts in endoplasmic-reticulum-associated degradation (ERAD), *Bioessays* **25**, 868–877.
38. Welch, W. J. (2004) Role of quality control pathways in human diseases involving protein misfolding, *Semin. Cell Dev. Biol.* **15**, 31–38.
39. Fan, J.-Q., Ishii, S., Asano, N., and Suzuki, Y. (1999) Accelerated transport and maturation of lysosomal α -galactosidase A in Fabry lymphoblasts by an enzyme inhibitor, *Nat. Med.* **5**, 112–115.
40. Kowalski, J. M., Parekh, R. N., Mao, J., and Wittrup, K. D. (1998) Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control, *J. Biol. Chem.* **273**, 19453–19458.
41. Kowalski, J. M., Parekh, R. N., and Wittrup, K. D. (1998) Secretion efficiency in *Saccharomyces cerevisiae* of bovine pancreatic trypsin inhibitor mutants lacking disulfide bonds is correlated with thermodynamic stability, *Biochemistry* **37**, 1264–1273.
42. Kjeldsen, T., Ludvigsen, S., Diers, I., Balschmidt, P., Sorensen, A. R., and Kaarsholm, N. C. (2002) Engineering-enhanced protein secretory expression in yeast with application to insulin, *J. Biol. Chem.* **277**, 18245–18248.
43. Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M., and Wittrup, K. D. (1999) Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency, *J. Mol. Biol.* **292**, 949–956.
44. Fujita, J. (1999) Cold shock response in mammalian cells, *J. Mol. Microbiol. Biotechnol.* **1**, 243–255.
45. Sanna, L. A., Fujita, J., Gaffin, S. L., and Lilly, C. M. (2002) Invited review: Effects of heat and cold stress on mammalian gene expression, *J. Appl. Physiol.* **92**, 1725–1742.
46. Zhang, X., Wang, X., Yue, H., Leung, S., Thibodeau, P., Thomas, P., and Guggino, S. (2003) Organic solutes rescue the functional defect in delta F508 cystic fibrosis transmembrane conductance regulator, *J. Biol. Chem.* **278**, 51232–51242.
47. Sharma, M., Benharouga, M., Hu, W., and Lukacs, G. L. (2001) Conformational and temperature-sensitive stability defects of the delta F508 cystic fibrosis transmembrane conductance regulator in post-endoplasmic reticulum compartments, *J. Biol. Chem.* **276**, 8942–8950.
48. Yam, G. H., Zuber, C., and Roth, J. (2005) A synthetic chaperone corrects the trafficking defect and disease phenotype in a protein misfolding disorder, *FASEB J.* **19**, 12–18.

49. Asano, N., Ishii, S., Kizu, H., Ikeda, K., Yasuda, K., Kato, A., Martin, O. R., and Fan, J.-Q. (2000) In vitro inhibition and intracellular enhancement of lysosomal α -galactosidase A activity in Fabry lymphoblasts by 1-deoxygalactonojirimycin and its derivatives, *Eur. J. Biochem.* **267**, 4179–4186.
50. Tropak, M. B., Reid, S. P., Guiral, M., Withers, S. G., and Mahuran, D. (2004) Pharmacological enhancement of β -hexosaminidase activity in fibroblasts from adult Tay-Sachs and Sandhoff patients, *J. Biol. Chem.* **279**, 13478–13487.
51. Matsuda, J., Suzuki, O., Oshima, A., Yamamoto, Y., Noguchi, A., Takimoto, K., Itoh, M., Matsuzaki, Y., Yasuda, Y., Ogawa, S., Sakata, Y., Nanba, E., Higaki, K., Ogawa, Y., Tominaga, L., Ohno, K., Iwasaki, H., Watanabe, H., Brady, R. O., and Suzuki, Y. (2003) Chemical chaperone therapy for brain pathology in GM1-gangliosidosis, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15912–15917.
52. Beutler, E., Kuhl, W., and Sorge, J. (1984) Cross-reacting material in Gaucher disease fibroblasts, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6506–6510.
53. Carlsson, S. R., Roth, J., Piller, F., and Fukuda, M. (1988) Isolation and characterization of human lysosomal membrane glycoproteins, h-lamp-1 and h-lamp-2. Major sialoglycoproteins carrying polylactosaminoglycan, *J. Biol. Chem.* **263**, 18911–18919.
54. Ginns, E. I., Tegelaers, F. P. W., Barneveld, R., Galjaard, H., Reuser, A. J. J., Brady, R. O., Tager, J. M., and Barranger, J. A. (1983) Determination of Gaucher's disease phenotypes with monoclonal antibody, *Clin. Chim. Acta* **131**, 283–287.
55. Aerts, J. M. F. G., Donker-Koopman, W. E., Murray, G. J., Barranger, J. A., Tager, J. M., and Schram, A. W. (1986) A procedure for the rapid purification in high yield of human glucocerebrosidase using immunoaffinity chromatography with monoclonal antibodies, *Anal. Biochem.* **154**, 655–663.
56. Naim, H. Y., Lacey, S. W., Sambrook, J. F., and Gething, M. J. (1991) Expression of a full-length cDNA coding for human intestinal lactase-phlorizin hydrolase reveals an uncleaved, enzymatically active, and transport-competent protein, *J. Biol. Chem.* **266**, 12313–12320.
57. Furbish, F. S., Steer, C. J., Krett, N. L., and Barranger, J. A. (1981) Uptake and distribution of placental glucocerebrosidase in rat hepatic cells and effects of sequential deglycosylation, *Biochim. Biophys. Acta* **673**, 425–434.

Fluorogenic Label for Biomolecular Imaging

Luke D. Lavis[†], Tzu-Yuan Chao[‡], and Ronald T. Raines^{†,*,*}

Departments of [†]Chemistry and [‡]Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706

ABSTRACT Traditional small-molecule fluorophores are always fluorescent. This attribute can obscure valuable information in biological experiments. Here, we report on a versatile “latent” fluorophore that overcomes this limitation. At the core of the latent fluorophore is a derivative of rhodamine in which one nitrogen is modified as a urea. That modification enables rhodamine to retain half of its fluorescence while facilitating conjugation to a target molecule. The other nitrogen of rhodamine is modified with a “trimethyl lock”, which enables fluorescence to be unmasked fully by a single user-designated chemical reaction. An esterase-reactive latent fluorophore was synthesized in high yield and attached covalently to a cationic protein. The resulting conjugate was not fluorescent in the absence of esterases. The enzymatic activity of esterases in endocytic vesicles and the cytosol induced fluorescence, enabling the time-lapse imaging of endocytosis into live human cells and thus providing unprecedented spatiotemporal resolution of this process. The modular design of this “fluorogenic label” enables the facile synthesis of an ensemble of small-molecule probes for the illumination of numerous biochemical and cell biological processes.

Fluorescent molecules are critical tools in the study of biochemical and cell biological processes (1). In many studies, however, only few of the fluorescent molecules experience a phenomenon of interest. Because traditional fluorophores, such as rhodamine and fluorescein, are always fluorescent, bulk fluorescence can obscure valuable information. To overcome this limitation, molecules can be designed such that a chemical reaction elicits a change in their fluorescence. Such “latent” fluorophores are at the core of common methods, including the enzyme-linked immunosorbent assay (ELISA), high-throughput screening of enzyme inhibitors, detection of reporter genes, and evaluation of cell viability (1). We reasoned that the use of a latent fluorophore as a “fluorogenic label” could overcome limitations of traditional fluorescent labels and thereby improve the spatial and temporal resolution of bioimaging.

Recently, our laboratory reported on a new class of latent fluorophores based on the “trimethyl lock” (2, 3). The rapid lactonization (4, 5) of the trimethyl lock had been exploited previously to prepare stable pro-drugs that were unmasked by an enzyme-catalyzed reaction (6, 7). We first used the trimethyl lock to shroud the fluorescence of a xanthene dye, rhodamine 110 (Rh₁₁₀) (2), and then an oxazine dye, cresyl violet (3). This approach afforded highly stable bis(trimethyl lock) “pro-fluorophores” that were labile to esterase catalysis *in vitro* and *in cellulo*.

Our bis(trimethyl lock) pro-fluorophores had two problematic attributes. First, two chemical reactions were necessary to unveil the vast majority of their fluorescence, decreasing the rate of fluorescence manifestation and limiting the linear range of assays (8, 9). Analogous fluorogenic protease substrates based on a rhodamine diamide display complex hydrolysis kinetics (10–12), as we observed with our bis(trimethyl lock) pro-fluorophores (2, 3). The second problematic attribute was the absence of a

*To whom correspondence should be addressed.
E-mail: raines@biochem.wisc.edu.

Received for review March 24, 2006
and accepted May 1, 2006.

Published online May 19, 2006

10.1021/cb600132m CCC: \$33.50

© 2006 by American Chemical Society

handle for target–molecule conjugation. Such a handle is available in derivatives, such as 5- or 6-carboxyrhodamine, that are accessible only from low-yielding synthetic routes.

We suspected that we could solve both problems by capping one of the amino groups of Rh₁₁₀. The capping of rhodamine dyes with an amide (13–15), carbamate (16), or urea (17) can preserve much of their fluorescence. We were especially intrigued by the attributes of urea–rhodamine, which according to recent reports in the scientific (17) and patent (18, 19) literature appears to retain significant fluorescence intensity relative to Rh₁₁₀.

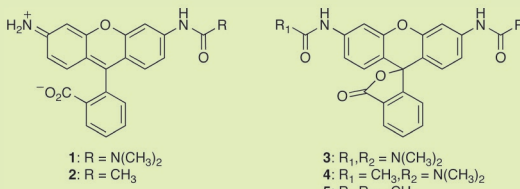
Here, we report on a versatile fluorogenic label for biomolecular imaging. First, we describe the synthesis of a complete set of ureated and amidated derivatives of Rh₁₁₀, as well as a characterization of their fluorescent properties. Then, we show that imposing our trimethyl lock strategy upon a urea–rhodamine yields a stable latent fluorophore with a high rate of enzymatic hydrolysis. Finally, we demonstrate the power of our modular approach by using the urea moiety as a handle for protein conjugation and subsequent continuous imaging of endocytosis by live human cells.

RESULTS AND DISCUSSION

Synthesis of Model Compounds. To gain a comprehensive understanding of the urea and amide derivatives of rhodamine, we undertook the synthesis of compounds 1–5 (Table 1). Rhodamine itself and these five derivatives encompass the ensemble of possible ureated and amidated derivatives. We were especially interested in those properties of 1–5 with biological implications, such as the extinction coefficient and quantum yield in aqueous solution. Previous reports (10, 11, 17–20) of similar derivatives did not provide a complete listing of relevant fluorescent characteristics.

Installation of the urea moiety to produce urea 1 proved to be surprisingly difficult. In our hands, the reported conditions (17) involving the reaction of Rh₁₁₀ with a carbamoyl chloride using Hünig's base gave an intractable mixture of products. In contrast, we found that Rh₁₁₀ was deprotonated effectively with NaH and that the resulting anion reacted with dimethylcarbamyl chloride to yield the desired urea 1. This deprotonation strategy also proved useful for the synthesis of amide 2 and diurea 3. The additional acetamide group in urea–

TABLE 1. Spectroscopic Properties of Rh₁₁₀ and Its Derivatives



Dye	λ_{\max} (nm)	ϵ (M ⁻¹ cm ⁻¹)	λ_{em} (nm)	Φ	$\epsilon \times \Phi$ (rel)
Rh ₁₁₀	496	74,000	517	0.92	100%
1	492	48,600	518	0.49	35%
2	489	30,200	522	0.28	12%
3	482	3300	517	0.01	0.05%
4	475	400	—	—	—
5	~469	≤200	—	—	—

amide 4 and diamide 5 were installed by reaction with acetyl chloride in the presence of a base.

Fluorescence Properties. The absorbance and fluorescence spectra of Rh₁₁₀ and each derivative are shown (Figure 1). The corresponding values of λ_{\max} , extinction coefficient at λ_{\max} (ϵ), λ_{em} , and quantum yield (Φ) are listed (Table 1). We determined the relative fluorescence intensity of these compounds by calculating the product of extinction coefficient and quantum yield and then normalizing these values to those of Rh₁₁₀. In our measurement, urea 1 retained 35% of the fluorescence intensity of Rh₁₁₀ with a quantum yield value of 0.49. Amide 2 is only 12% as fluorescent as Rh₁₁₀, which is consistent with earlier reports (10, 11). The fluorescence of the bis-substituted dyes was largely quenched in aqueous solution. Diurea 3 did, however, possess significant absorbance and fluorescence compared to the urea–amide 4 or diamide 5. These latter two rhodamine derivatives are essentially nonfluorescent.

We also determined the pH dependence of the fluorescence of urea 1 and amide 2. The fluorescence of Rh₁₁₀ is relatively insensitive to pH values between 4 and 10 (1). This property is beneficial in biological assays, where unknown variations in pH can hamper quantitative measurements. Like Rh₁₁₀, urea 1 and amide 2 show no significant spectral change between pH values of 4 and 10; details may be seen in Supplementary Figure 1.

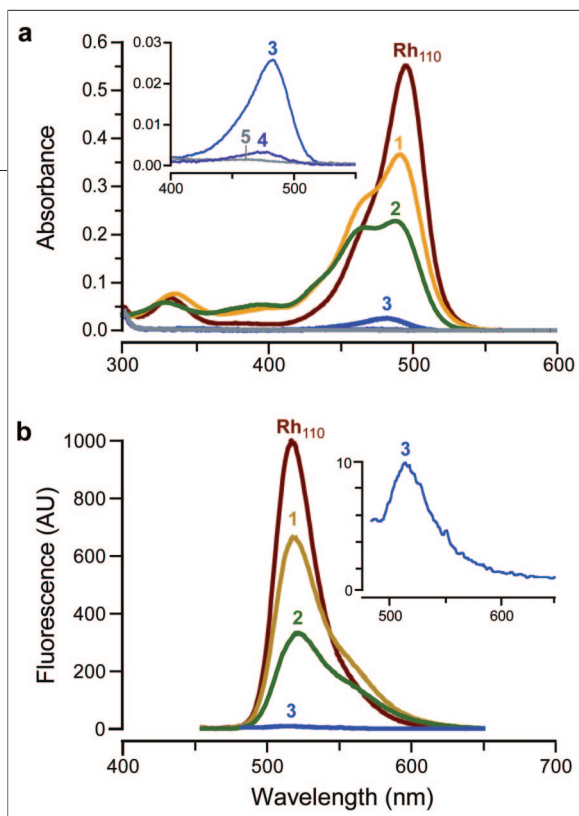


Figure 1. Spectra of Rh₁₁₀ and its derivatives. **a)** Absorption spectra of Rh₁₁₀ and derivatives 1–5 (7.5 μM). **b)** Fluorescent emission spectra of Rh₁₁₀ and 1–3 ($\lambda_{\text{ex}} = 450 \text{ nm}$, not to scale).

cent and a lactone that is colorless and nonfluorescent. Substitution on nitrogen can affect both this open–closed equilibrium and the spectral characteristics of the fluorescent zwitterions (22, 23). We suspected that the differences in optical properties seen in compounds 1–5 could be rationalized, in part, through examination of the electron-donation capability of the different nitrogen substituents. In agreement with this reasoning, weakly donating substituents would favor the colorless lactone as well as decrease the intrinsic absorptivity of the zwitterions and, hence, the extinction coefficient. Weakly donating substituents could also reduce the quantum yield by decreasing the C–N bond order and thereby enhancing nonradiative decay of the excited state through vibrational relaxation processes (24, 25).

We explored the relationship between the values of extinction coefficient and quantum yield and the Hammett σ_p substituent constants (26). An unprotonated amino group is a good electron donor ($\sigma_p = -0.66$), whereas an amide group is a relatively poor donor ($\sigma_p = 0.00$), due to amidic resonance. A urea group is peculiar; its carbonyl group is cross-conjugated and both of its nitrogens participate in amidic resonance. This cross-conjugation attenuates its electron-donating ability, as reflected in an intermediate Hammett constant ($\sigma_p = -0.26$). A plot of both extinc-

tion coefficient and quantum yield versus σ_p substituent constant for Rh₁₁₀ and monosubstituted rhodamines 1 and 2 is shown (Figure 2). The correlation indicates that both spectral properties are affected by electron donation from the nitrogens. A similar trend in quantum yields has been observed in substituted phenoxazine dyes (27).

The moderate electron-donating character of the urea moiety provides an explanation for the advantageous properties of urea 1. Substitution with the cross-conjugated urea suppresses the fluorescence intensity of urea 1 relative to Rh₁₁₀. This decrease is not, however, as severe as seen in amide 2, due to the greater electron-donating properties of the urea moiety. Still, the attenuated electron-donation allows complete suppression of fluorescence upon amidation of the remaining nitrogen in urea–amide 4. Finally, the effect of electron-rich substituents on the rhodamine system is apparent again in the fluorescence of diurea 3, as it is greater than that of diamide 5.

Synthesis of Urea–Rhodamine Trimethyl Lock.

Having affirmed the desirable properties of urea–rhodamine, we next sought to apply our trimethyl lock strategy to this dye. The synthetic route to the fluorogenic substrate, which employs rhodamine morpholino-urea 6 (17), is shown (Scheme 1). Again, we found that the use of Hünig's base in the synthesis afforded a mixture of products. Deprotonation of Rh₁₁₀ with NaH followed by dropwise addition of 4-morpholinecarbonyl chloride furnished rhodamine morpholino-urea 6. This compound exhibited similar fluorescent characteristics to urea 1 (Table 1), and it had an extinction coefficient of

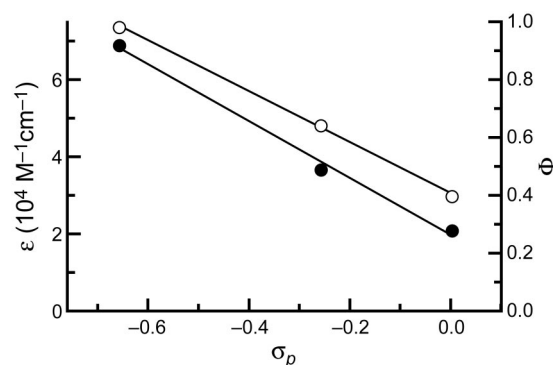
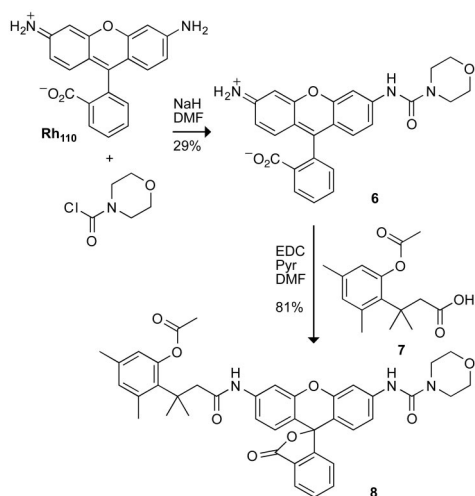


Figure 2. Hammett plot of extinction coefficient (\circ) and quantum yield (\bullet) versus σ_p for Rh₁₁₀, urea 1, and amide 2.



Scheme 1. Synthetic Route to Pro-Fluorophore 8

$51\,700\text{ M}^{-1}\text{ cm}^{-1}$ and quantum yield of 0.44. Carbodiimide coupling of rhodamine morpholino-urea **6** with acid **7** (**28**) afforded the desired pro-fluorophore **8**.

Chemical Stability. Pro-fluorophore **8** must be stable in aqueous solution to be useful in biological assays. Such stability can be problematic for hydrolase substrates, as spontaneous hydrolysis can compete effectively with enzymatic activity and raise background levels. As shown (Figure 3), pro-fluorophore **8** showed remarkable stability in both phosphate-buffered saline (PBS) and Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS). In contrast, fluorescein diacetate, which is a widely used esterase substrate (**29**), suffered relatively rapid hydrolysis in both solutions. This dramatic increase in stability arises from the large difference in $\text{p}K_{\text{a}}$ values between the conjugate acids of the two leaving groups. Specifically, fluorescein ($\text{p}K_{\text{a}}\ 6.32$ (**30**)) is a much better leaving group than is the electron-rich trimethyl-lock phenol (*o*-methylphenol has a $\text{p}K_{\text{a}}$ of 10.28 (**31**)).

Enzymatic Reactivity. An objective in the design of pro-fluorophore **8** was to improve its reactivity as an esterase substrate relative to the original bis(trimethyl lock) rhodamine substrate. The appearance of fluorescence upon reaction of porcine liver esterase (PLE) with pro-fluorophore **8** was indicative of single-hit kinetics (Figure 4). The kinetic constants were calculated to be $k_{\text{cat}}/K_{\text{M}} = 8.2 \times 10^5\text{ M}^{-1}\text{ s}^{-1}$ and $K_{\text{M}} = 0.10\ \mu\text{M}$. Comparison with the apparent kinetic constants from

the original bis(trimethyl lock) rhodamine substrate (**2**) ($k_{\text{cat}}/K_{\text{M}} = 1.9 \times 10^3\text{ M}^{-1}\text{ s}^{-1}$ and $K_{\text{M}} = 0.47\ \mu\text{M}$) shows a 430-fold increase in the $k_{\text{cat}}/K_{\text{M}}$ value. A more appropriate comparison takes into account the expected 65% decrease in fluorescence of urea **6** (Table 1), which is the hydrolysis product of pro-fluorophore **8**, relative to Rh₁₁₀. After this adjustment, latent fluorophore performance is still enhanced by 150-fold.

The substantial increase in catalytic efficiency is likely due to the change from the double-hit kinetics observed for the bis-substituted substrate to the single-hit kinetics of pro-fluorophore **8**. Hydrolysis of the bis-substituted fluorogenic substrate progresses from diamide to free Rh₁₁₀ via a monoamide intermediate, with the unmasking of the second amino group producing the majority (~90%) of the fluorescence (**10**, **11**). In contrast, the urea–rhodamine substrate requires only a single cleavage event for the complete manifestation of fluorescence.

Cellular Imaging. Once the high chemical stability and enzymatic reactivity of pro-fluorophore **8** were established, we next evaluated the behavior of this compound in live human cells. Pro-fluorophore **8** was incubated with HeLa cells and imaged using confocal fluorescence microscopy. The substrate was activated *in cellulo*

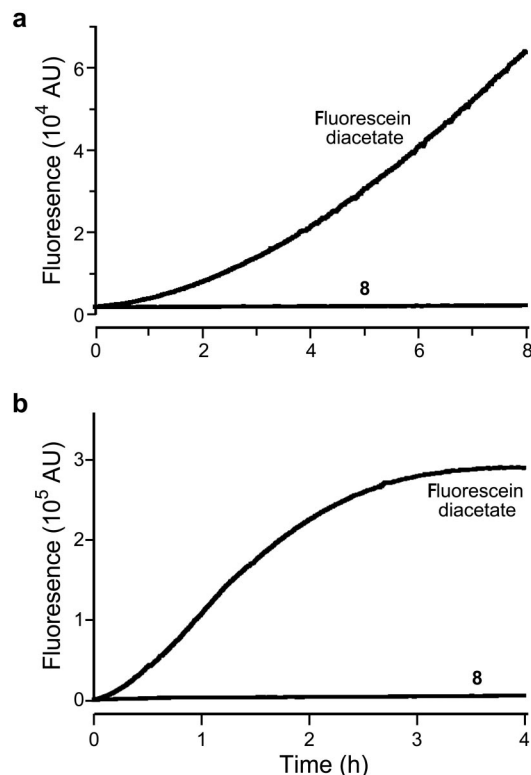


Figure 3. Stability of pro-fluorophore 8 and fluorescein diacetate in aqueous solution. a) Time course for the generation of fluorescence ($\lambda_{\text{ex}} = 496\text{ nm}$, $\lambda_{\text{em}} = 520\text{ nm}$) of pro-fluorophore **8** (25 nM) and fluorescein diacetate (25 nM) in PBS. **b)** Time course for the generation of fluorescence ($\lambda_{\text{ex}} = 496\text{ nm}$, $\lambda_{\text{em}} = 520\text{ nm}$) of pro-fluorophore **8** (25 nM) and fluorescein diacetate (25 nM) in DMEM containing FBS (10%, v/v).

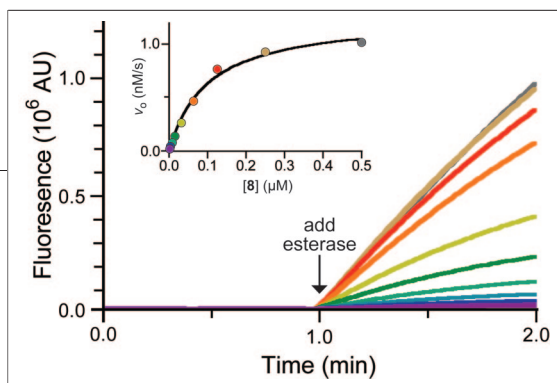


Figure 4. Kinetic traces ($\lambda_{\text{ex}} = 496 \text{ nm}$, $\lambda_{\text{em}} = 520 \text{ nm}$) and Michaelis–Menten plot (inset) for a serial dilution of pro-fluorophore **8** ($0.5 \mu\text{M} \rightarrow 2 \text{ nM}$) with PLE ($2.5 \mu\text{g mL}^{-1}$).

of the fluorogenic probe permitted its imaging in the cytosol without an intermediate washing step. Counterstaining with LysoTracker Red showed significant but incomplete colocalization, suggesting that, after hydrolysis, a portion of the free urea–rhodamine localized in acidic vesicles (yellow color in Figure 5, panel b). To ensure that the fluorescence increase was due to trimethyl-lock activation and not hydrolysis of the urea moiety, we incubated HeLa cells with the relatively nonfluorescent diurea–rhodamine **5**. In these experiments, we observed virtually no intracellular fluorescence; details may be seen in Supplementary Figure 2.

Fluorogenic Label. The high chemical stability and rapid *in cellulo* unmasking of pro-fluorophore **8** prompted us to develop a derivative for bioconjugation. We reasoned that such a fluorogenic label would be stable enough to survive conjugation and purification protocols while still providing a strong signal for continuous biological experiments. It is noteworthy that simple fluorescein diesters have found only limited use as fluorogenic labels (32–35), as fluorescein diesters suffer from low chemical stability in aqueous solution (Figure 3).

Developing pro-fluorophore **8** into a fluorogenic label requires the installation of a functional group with selective reactivity. We chose to install the maleimide functionality (36, 37), which reacts rapidly with thiol groups (38). The resulting conjugates are stable (39), even after the slow hydrolysis of the nascent sulfosuccinimidyl ring (40).

Traditionally, reactive groups are attached to the pendant carboxyphenyl ring of rhodamine and fluorescein dyes (1). Synthesis of these compounds requires difficult chromatographic steps to obtain isomerically pure compounds (41). We envisioned a facile and economical alternative involving the attachment of a maleimide derivative via the desirable urea functionality. Although uncommon, bioconjugation *via* the amino groups of rhodamines has been used previously

by endogenous esterases to produce diffuse green cytosolic staining (Figure 5, panel a). Importantly, the high chemical stability

(14, 42, 43). This strategy allows the use of commercially available (and relatively inexpensive) Rh₁₁₀ as the starting material for the synthesis of maleimidourea–rhodamine trimethyl lock **13**, as shown (Scheme 2). Desymmetrization of Rh₁₁₀ was accomplished by its deprotonation with NaH and reaction with Boc₂O to give *t*-Boc–rhodamine **9**. An isocyanate was generated *in situ* from maleimide **10** by a Curtius rearrangement (44, 45), and this isocyanate was reacted with *t*-Boc–rhodamine **9** to generate a urea (46, 47). Deprotection of maleimidourea–rhodamine–*t*-Boc **11** with TFA afforded fluorescent urea–rhodamine **12**. Condensation with **7** using EDC gave thiol-reactive fluorogenic label **13**.

Bioconjugation. To test the utility of fluorogenic label **13** in a biological experiment, we attached it to a thiol-containing variant of bovine pancreatic ribonuclease (RNase A (48)). RNase A is a cationic protein that is internalized by mammalian cells via endocytosis (49). This internalization is critical to the action of cytotoxic RNase A variants and homologues (50). Fluorogenic label **13** reacted cleanly with the A19C variant of RNase A to give a mono-substituted conjugate as determined by MALDI mass spectrometry. This protein conjugate was stable to purification by cation-exchange chromatography at pH 5.0 and showed a 1200-fold increase in fluorescence upon incubation with PLE (data not shown).

At physiological pH, the protein conjugate was less stable than unconjugated pro-

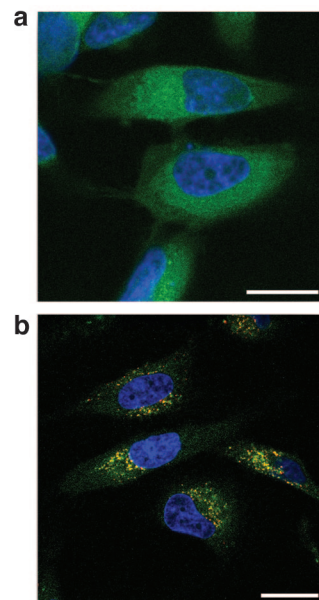
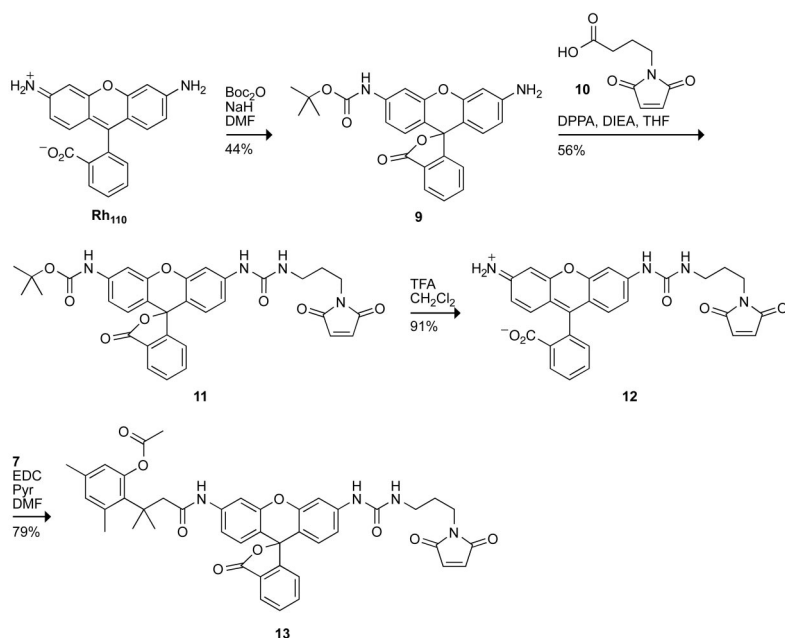


Figure 5. Unmasking of pro-fluorophore **8** in live human cells. a) Unwashed HeLa cells incubated for 1 h with pro-fluorophore **8** ($10 \mu\text{M}$) at $37 \text{ }^\circ\text{C}$ in DMEM and counterstained with Hoechst 33342. b) Washed HeLa cells incubated for 1 h with pro-fluorophore **8** ($10 \mu\text{M}$) at $37 \text{ }^\circ\text{C}$ in DMEM and counterstained with Hoechst 33342 and LysoTracker Red (5%, v/v $\text{CO}_2(\text{g})$, 100% humidity). Scale bar: $20 \mu\text{m}$.



Scheme 2. Synthetic Route to Fluorogenic Label 13

fluorophore **8**. Spontaneous hydrolysis of the acetate ester was slow but significant in PBS, presumably because conjugation to the protein places the probe in close proximity to nucleophilic functional groups of the protein. Storage at pH 5.0 did, however, extend the

stability of the conjugate, allowing multiple experiments to be performed with one preparation.

Cellular Imaging with a Bioconjugate. Fluorescently labeled biomolecules have been used to image endocytotic events (51). We sought to determine the efficacy of our fluorogenic label approach by comparing endocytosis of HeLa cells incubated with Oregon Green-labeled RNase A (49) to that of cells incubated with the protein conjugated with fluorogenic label **13**. The Oregon Green conjugate showed intense extracellular background signal that obscures the fluorescence from endocytosed material

(Figure 6, panel a). This background could be eliminated only with many vigorous washing steps (Figure 6, panel b). In contrast, the pro-fluorophore conjugate allowed imaging without intermediate washing steps. Unwashed HeLa cells incubated with the RNase A conju-

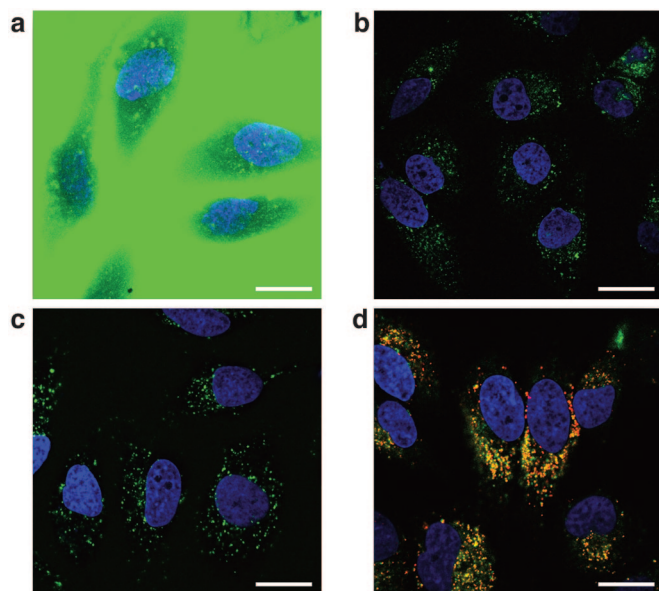
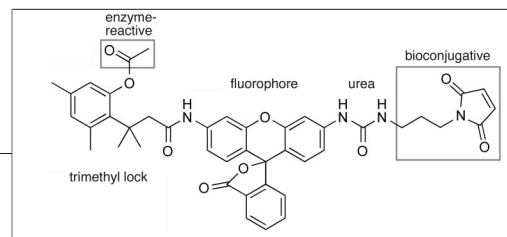


Figure 6. Live-cell imaging experiments with protein conjugates. a) Unwashed HeLa cells incubated for 1 h with Oregon Green–RNase A conjugate ($10\ \mu\text{M}$) at $37\ ^\circ\text{C}$ in DPBS and counter-stained with Hoechst 33342. b) Washed HeLa cells incubated for 1 h with Oregon Green–RNase A conjugate ($10\ \mu\text{M}$) at $37\ ^\circ\text{C}$ in DPBS and counter-stained with Hoechst 33342. c) Unwashed HeLa cells incubated for 1 h with fluorogenic label **13**–RNase A conjugate ($10\ \mu\text{M}$) at $37\ ^\circ\text{C}$ in DPBS and counter-stained with Hoechst 33342. d) Washed HeLa cells incubated for 1 h with fluorogenic label **13**–RNase A conjugate ($10\ \mu\text{M}$) at $37\ ^\circ\text{C}$ in DPBS and counter-stained with Hoechst 33342 and LysoTracker Red (5%, v/v $\text{CO}_2(\text{g})$, 100% humidity). Scale bar: $20\ \mu\text{m}$.



Scheme 3. Modules in Fluorogenic Label 13

gate have bright, punctate staining (Figure 6, panel c), indicative of the conjugate being localized in small vesicles. Counterstaining with LysoTracker Red shows a large degree of colocalization (Figure 6, panel d), suggesting that the latent conjugate is internalized via endocytosis and activated by endosomal or lysosomal esterases (52–54). Images with the protein conjugate (Figure 6, panel c) are less diffuse and more punctate than are images with free pro-fluorophore **8** (Figure 5), which has much more ready access to the cytosol. To ensure that the signal is due to unmasked fluorophore attached to RNase A (Figure 6, panel c), we fixed cells incubated with our latent conjugate and counterstained them with a primary antibody to RNase A and a secondary antibody labeled with AlexaFluor 594. In a fluorescence microscopy image, we observed a significant overlap of the green and red fluorescent signals to produce a yellow signal, indicating that the unmasked RNase A conjugate is largely intact; details may be seen in Supplementary Figure 3.

The high chemical stability and low background fluorescence of the fluorogenic label conjugate allowed for the time-lapse imaging of its endocytosis. Cells were incubated with the fluorogenic label **13**–RNase A conjugate at room temperature, and images were recorded without washing during the next 90 min. The compilation of these images into a movie revealed that internalization of the conjugate occurred continuously and that vesicular fluorescence increased monotonically; details may be seen in the Supplemental Movie.

Envoi. We have demonstrated how a common fluorophore, Rh₁₁₀, can be elaborated into a powerful new

tool for biochemistry and cell biology. The use of a trimethyl lock provides a latent fluorophore with high chemical stability while maintaining enzymatic reactivity (2, 3). The use of a urea group (rather than a second trimethyl lock) improves enzymatic reactivity markedly while preserving desirable fluorescence properties, as in pro-fluorophore **8**. The elaboration of the urea to include an electrophile outfits the latent fluorophore for conjugation, as in fluorogenic label **13**. Conjugation of this fluorogenic label to a target molecule enables, for example, the continuous imaging of the endocytosis of a target molecule by live human cells.

We note that the urea–rhodamine–trimethyl lock probe is modular and, hence, can be tailored to suit a variety of applications (Scheme 3). For example, alteration of the bioconjugative group on the urea moiety could be used to change conjugation chemistry, enhance cellular internalization, or target a conjugate to a specific subcellular location. Modification of the enzyme-reactive group on the trimethyl lock could enable the detection of a conjugate in a particular organelle. The use of fluorogenic labels could even transcend cultured cells, allowing for continuous imaging in tissues or *in vivo*. These applications would be facilitated by extant comprehensive inventories of the enzymes in numerous organs and organelles (55, 56). Accordingly, the fluorogenic label strategy could enable the development of specific probes for biological experiments of ever-increasing sophistication.

METHODS

General Spectroscopic Methods. HEPES (2[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid) was from Research Products International. Fluorescein (reference standard grade) was from Molecular Probes. Other reagents were from Sigma-Aldrich or Fisher Scientific. Phosphate-buffered saline, pH 7.4 (PBS) contained (in 1.00 L) KCl (0.20 g), KH₂PO₄ (0.20 g), NaCl (8.0 g), and Na₂HPO₄·7H₂O (2.16 g). All measurements were recorded at ambient temperature (23 ± 2 °C), and buffers were not degassed prior to measurements. Compounds were prepared as stock solutions in DMSO and diluted such that the DMSO concentration did not exceed 1% (v/v). Porcine liver esterase (PLE, MW = 163 kDa (57)) was obtained from Sigma Chemical (product number E2884) as a suspension in 3.2 M (NH₄)₂SO₄ and was diluted to appropriate concentrations in PBS before use. In pH dependency studies, the pH of PBS was adjusted by addition of 1.0 M HCl or 1.0 M NaOH and measured using a

Beckmann glass electrode that was calibrated prior to each use. Graphs were manipulated and parameters were calculated with Microsoft Excel 2003 and GraphPad Prism 4.

Ultraviolet–Visible and Fluorescence Spectroscopy. Absorption spectra were recorded in 1-cm path length cuvettes having a volume of 1.0 or 3.5 mL on a Cary model 50 spectrometer from Varian. The extinction coefficients were measured in 10 mM HEPES–NaOH buffer, pH 7.5. Fluorometric measurements were made using fluorescence grade quartz or glass cuvettes from Starna Cells and a QuantaMaster1 photon-counting spectrofluorometer from Photon Technology International equipped with sample stirring. The quantum yields of Rh₁₁₀ and compounds **1–5** were measured with dilute samples (*A* ≤ 0.1) in 10 mM HEPES–NaOH buffer, pH 7.5. These values were obtained by the comparison of the integrated area of the emission spectrum of the samples with that of fluorescein in 0.1 M NaOH, which has a quantum efficiency of 0.95 (58). The concentration of the fluo-

rescein reference was adjusted to match the absorbance of the test sample at the excitation wavelength. Under these conditions, quantum yields were calculated using eq 1.

$$\Phi_{\text{sample}} = \Phi_{\text{standard}} \left(\frac{I F_{\text{em, sample}}}{I F_{\text{em, standard}}} \right) \quad (1)$$

Protein Purification and Labeling. The TNB-protected A19C variant of RNase A and the Oregon Green-labeled RNase A conjugate were prepared as described previously (49). The TNB-protected protein was deprotected with a 3-fold molar excess of dithiothreitol (DTT) and desalted by chromatography using a HiTrap Desalting column (Amersham). The protein conjugate then was prepared by reaction with a 10-fold molar excess of thiol-reactive maleimide **13** for 16 h at 4 °C. Purification by chromatography using a HiTrap HP SP column (Amersham) afforded the desired conjugate (MS (MALDI): m/z 14 468 (expected, 14 475)). Protein concentration was determined by using a bicinchoninic acid (BCA) assay kit from Pierce with wild-type RNase A as a standard.

Cell Preparation. HeLa cells were plated on Nunc Lab-Tek II 8-well Chamber Coverglass (Fisher Scientific) and grown to 70–80% confluence at 37 °C in DMEM (Invitrogen) containing FBS (10% v/v). For static imaging, cells were first washed with Dulbecco's phosphate-buffered saline (DPBS, Invitrogen). Cells were then incubated with pro-fluorophore **8** (10 μM), RNase A conjugated to maleimide **13** (10 μM), or Oregon Green-labeled RNase A (10 μM) for 1 h at 37 °C prior to imaging. Nuclear staining was accomplished by addition of Hoechst 33342 (2 $\mu\text{g mL}^{-1}$) for the final 5 min of incubation. Lysosomal staining involved washing the cells with DPBS followed by incubation with 100 nM LysoTracker Red (Molecular Probes) in DPBS for 1 min at ambient temperature. For dynamic imaging, cells were incubated with Hoechst 33342 (2 $\mu\text{g mL}^{-1}$) for 5 min at 37 °C, and then washed twice with DPBS. Pro-fluorophore **13**–RNase A conjugate (10 μM) was added to the cells at ambient temperature (23 \pm 2 °C). Imaging of endocytosis started within 1 min after the addition of the conjugate.

Cell Imaging. Cells were imaged on a Nikon Eclipse TE2000-U confocal microscope equipped with a Zeiss AxioCam digital camera, unless indicated otherwise. Excitation at 408 nm was provided by a blue-diode laser, and emission light was passed through a filter centered at 450 nm with a 35-nm band-pass. Excitation at 488 nm was provided by an argon-ion laser, and emission light was passed through a filter centered at 515 nm with a 40-nm band-pass. Excitation at 543 nm was provided by a HeNe laser, and emission light was passed through a filter centered at 605 nm with a 75-nm band-pass. For time-lapse imaging, one image per minute was recorded during the first 30 min of incubation, two images per min were recorded during the next 10 min, and five images per min were recorded during the last 50 min. The resulting movie condenses these 300 images recorded over 90 min into 40 s. Brightfield images indicated that the cells were alive and appeared to have normal physiology, both before and after the time-lapse imaging.

Acknowledgment: We are grateful to K. A. Dickson for preliminary bioimaging experiments and S. S. Chandran, Z. Diwu, and M. B. Soellner for contributive discussions. L.D.L was supported by Biotechnology Training Grant 08349 (NIH). This work was supported by Grant CA73808 (NIH). The University of Wisconsin–Madison Biophysics Instrumentation Facility was established with Grants BIR-9512577 (NSF) and RR13790 (NIH). NMRFAM was supported by Grant P41RR02301 (NIH).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Haugland, R. P., Spence, M. T. Z., Johnson, I. D., and Basey, A. (2005) *The Handbook: A Guide to Fluorescent Probes and Labeling Technologies, 10th ed.*, Molecular Probes, Eugene, OR.
- Chandran, S. S., Dickson, K. A., and Raines, R. T. (2005) Latent fluorophore based on the trimethyl lock. *J. Am. Chem. Soc.* **127**, 1652–1653.
- Lavis, L. D., Chao, T.-Y., and Raines, R. T. (2006) Latent blue and red fluorophores based on the trimethyl lock. *ChemBioChem*, in press.
- Milstein, S., and Cohen, L. A. (1972) Stereopopulation control. I. Rate enhancement in the lactonizations of *o*-hydroxyhydrocinnamic acids. *J. Am. Chem. Soc.* **94**, 9158–9165.
- Borchardt, R. T., and Cohen, L. A. (1972) Stereopopulation control. II. Rate enhancement of intramolecular nucleophilic displacement. *J. Am. Chem. Soc.* **94**, 9166–9174.
- Shan, D., Nicolaou, M. G., Borchardt, R. T., and Wang, B. (1997) Prodrug strategies based on intramolecular cyclization reactions. *J. Pharm. Sci.* **86**, 765–767.
- Testa, B., and Mayer, J. M. (2003) *Hydrolysis in Drug and Prodrug Metabolism: Chemistry, Biochemistry, and Enzymology*, Verlag Helvetica Chimica Acta, Zürich, Switzerland.
- Fiering, S. N., Roederer, M., Nolan, G. P., Micklem, D. R., Parks, D. R., and Herzenberg, L. A. (1991) Improved FACS-Gal: Flow cytometric analysis and sorting of viable eukaryotic cells expressing reporter gene constructs. *Cytometry* **12**, 291–301.
- Urano, Y., Kamiya, M., Kanda, K., Ueno, T., Hirose, K., and Nagano, T. (2005) Evolution of fluorescein as a platform for finely tunable fluorescence probes. *J. Am. Chem. Soc.* **127**, 4888–4894.
- Leytus, S. P., Melhado, L. L., and Mangel, W. F. (1983) Rhodamine-based compounds as fluorogenic substrates for serine proteinases. *Biochem. J.* **209**, 299–307.
- Leytus, S. P., Patterson, W. L., and Mangel, W. F. (1983) New class of sensitive and selective fluorogenic substrates for serine proteinases. Amino acid and dipeptide derivatives of rhodamine. *Biochem. J.* **215**, 253–260.
- Liu, J., Bhalgat, M., Zhang, C., Diwu, Z., Hoyland, B., and Klaubert, D. H. (1999) Fluorescent molecular probes V: A sensitive caspase-3 substrate for fluorometric assays. *Bioorg. Med. Chem. Lett.* **9**, 3231–3236.
- Guzikowski, A. P., Naleway, J. J., Shipp, C. T., and Schutte, R. C. (2000) Synthesis of a macrocyclic rhodamine 110 enzyme substrate as an intracellular probe for caspase 3 activity. *Tetrahedron Lett.* **41**, 4733–4735.
- Lorey, S., Faust, J., Mrestani-Klaus, C., Kaehne, T., Ansonge, S., Neubert, K., and Buehling, F. (2002) Transcellular proteolysis demonstrated by novel cell surface-associated substrates of dipeptidyl peptidase IV (CD26). *J. Biol. Chem.* **277**, 33170–33177.
- Zhang, H. Z., Kasibhatla, S., Guastella, J., Tseng, B., Drewe, J., and Cai, S. X. (2003) *N*-Ac-DEVD-*N'*-(polyfluorobenzoyl)-R110: Novel cell-permeable fluorogenic caspase substrates for the detection of caspase activity and apoptosis. *Bioconjugate Chem.* **14**, 458–463.
- Cai, S. X., Zhang, H.-Z., Guastella, J., Drewe, J., Yang, W., and Weber, E. (2001) Design and synthesis of rhodamine 110 derivative and caspase-3 substrate for enzyme and cell-based fluorescent assay. *Bioorg. Med. Chem. Lett.* **11**, 39–42.
- Wang, Z.-Q., Liao, J., and Diwu, Z. (2005) *N*-DEVD-*N'*-morpholinecarbonyl-rhodamine 110: Novel caspase-3 fluorogenic substrates for cell-based apoptosis assay. *Bioorg. Med. Chem. Lett.* **15**, 2335–2338.
- Zhang, H.-Z., Cai, S. X., Drewe, J. A., and Yang, W. (2000) Novel fluorescence dyes and their applications for whole cell fluorescence screening assays for caspases, peptidases, proteases and other enzymes and the use thereof. WO Patent Application 200004914.
- Diwu, Z., Liao, J., and Wang, Z. (2003) Preparation of rhodamine peptide derivatives as luminogenic protease substrates. WO Patent Application 2003099780.

20. Ioffe, I. S., and Otten, V. F. (1965) Rhodamine dyes and related compounds. XII. Diacetyl derivatives of rhodamine and rhodol; structure of colorless forms of fluoran dyes. *Zh. Obshch. Khim.* **1**, 336–339.
21. López Arbeloa, F., López Arbeloa, T., Tapia Estevez, M. J., and López Arbeloa, I. (1991) Photophysics of rhodamines: Molecular structure and solvent effects. *J. Phys. Chem.* **95**, 2203–2208.
22. Ioffe, I. S., and Otten, V. F. (1965) Rhodamine dyes and related compounds. XIV. Mutual conversions of colorless and colored forms of rhodamine and rhodol. *Zh. Obshch. Khim.* **1**, 343–346.
23. López Arbeloa, F., Urrecha Aguirresacona, I., and López Arbeloa, I. (1989) Influence of the molecular structure and the nature of the solvent on the absorption and fluorescence characteristics of rhodamines. *Chem. Phys.* **130**, 371–378.
24. Vogel, M., Rettig, W., Sens, R., and Drexhage, K. H. (1988) Structural relaxation of rhodamine dyes with different N-substitution patterns—A study of fluorescence decay times and quantum yields. *Chem. Phys. Lett.* **147**, 452–460.
25. López Arbeloa, T., López Arbeloa, F., Hernández Bartolomé, P., and López Arbeloa, I. (1992) On the mechanism of radiationless deactivation of rhodamines. *Chem. Phys.* **160**, 123–130.
26. Hansch, C., Leo, A., and Taft, R. W. (1991) A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **91**, 165–195.
27. Descalzo, A. B., Rurack, K., Weisshoff, H., Martinex-Manez, R., Marcos, M. D., Amoros, P., Hoffmann, K., and Soto, J. (2005) Rational design of a chromo- and fluorogenic hybrid chemosensor material for the detection of long-chain carboxylates. *J. Am. Chem. Soc.* **127**, 184–200.
28. Amsberry, K. L., Gerstenberger, A. E., and Borchardt, R. T. (1991) Amine prodrugs which utilize hydroxy amide lactonization. II. A potential esterase-sensitive amide prodrug. *Pharm. Res.* **8**, 455–461.
29. Rotman, B., and Papermaster, B. W. (1966) Membrane properties of living mammalian cells as studied by enzymatic hydrolysis of fluorogenic esters. *Proc. Natl. Acad. Sci. U.S.A.* **55**, 134–141.
30. Goldberg, J. M., and Baldwin, R. L. (1998) Kinetic mechanism of a partial folding reaction. 1. Properties of the reaction and effects of denaturants. *Biochemistry* **37**, 2546–2555.
31. Fickling, M. M., Fischer, A., Mann, B. R., Packer, J., and Vaughan, J. (1959) Hammett substituent constants for electron-withdrawing substituents: Dissociation of phenols, anilinium ions and dimethyl-anilinium ions. *J. Am. Chem. Soc.* **81**, 4226–4230.
32. Laurent, A., Debart, F., Lamb, N., and Rayner, B. (1997) Esterase-triggered fluorescence of fluorogenic oligonucleotides. *Bioconjugate Chem.* **8**, 856–861.
33. Bergsdorf, C., Beyer, C., Umansky, V., Werr, M., and Sapp, M. (2003) Highly efficient transport of carboxyfluorescein diacetate succinimidyl ester into COS7 cells using human papillomavirus-like particles. *FEBS Lett.* **536**, 120–124.
34. Drobni, P., Mistry, N., McMillan, N., and Evander, M. (2003) Carboxy-fluorescein diacetate, succinimidyl ester labeled papillomavirus virus-like particles fluoresce after internalization and interact with heparan sulfate for binding and entry. *Virology* **310**, 163–172.
35. Kamal, A., Ramulu, P., Srinivas, O., Ramesh, G., and Kumar, P. P. (2004) Synthesis of C8-linked pyrrolo[2,1-c][1,4]benzodiazepine-benzimidazole conjugates with remarkable DNA-binding affinity. *Bioorg. Med. Chem. Lett.* **14**, 4791–4794.
36. Ji, T. H. (1983) Bifunctional reagents. *Methods Enzymol.* **91**, 580–609.
37. Aslam, M., and Dent, A. (1998) *Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences*, Macmillan Reference, London.
38. Bednar, R. A. (1990) Reactivity and pH dependence of thiol conjugation to *N*-ethylmaleimide: Detection of a conformational change in chalcone isomerase. *Biochemistry* **29**, 3684–3690.
39. Yoshitake, S., Yamada, Y., Ishikawa, E., and Masseyeff, R. (1979) Conjugation of glucose-oxidase from *Aspergillus niger* and rabbit antibodies using *N*-hydroxysuccinimide ester of *N*-(4-carboxycyclohexylmethyl)-maleimide. *Eur. J. Biochem.* **101**, 395–399.
40. Ishii, Y., and Lehrer, S. S. (1986) Effects of the state of succinimido-dering on the fluorescence and structural properties of pyrene maleimide-labeled alpha-alpha-tropomyosin. *Biophys. J.* **50**, 75–80.
41. Jiao, G. S., Han, J. W., and Burgess, K. (2003) Syntheses of regioisomerically pure 5- and 6-halogenated fluoresceins. *J. Org. Chem.* **68**, 8264–8267.
42. Corrie, J. E. T., Craik, J. S., and Munasinghe, V. R. N. (1998) A homo-bifunctional rhodamine for labeling proteins with defined orientations of a fluorophore. *Bioconjugate Chem.* **9**, 160–167.
43. Meunier, F., and Wilkinson, K. J. (2002) Nonperturbing fluorescent labeling of polysaccharides. *Biomacromolecules* **3**, 857–864.
44. Curtius, T. (1890) Ueber Stickstoffwasserstoffsäure (Azoimid) N₃H. *Ber. Dtsch. Chem. Ges.* **23**, 3023–3033.
45. Curtius, T. (1894) Hydrazide und Azide organischer Säuren. *J. Prakt. Chem. (Weinheim, Ger.)* **50**, 275–294.
46. Scriven, E. F. V., and Turnbull, K. (1988) Azides: Their preparation and synthetic uses. *Chem. Rev.* **88**, 297–368.
47. Bräse, S., Gil, C., Knepper, K., and Zimmermann, V. (2005) Organic azides: An exploding diversity of a unique class of compounds. *Angew. Chem. Int. Ed.* **44**, 5188–5240.
48. Raines, R. T. (1998) Ribonuclease A. *Chem. Rev.* **98**, 1045–1065.
49. Haigis, M. C., and Raines, R. T. (2003) Secretory ribonucleases are internalized by a dynamin-independent endocytic pathway. *J. Cell Sci.* **116**, 313–324.
50. Haigis, M. C., Kurten, E. L., and Raines, R. T. (2003) Ribonuclease inhibitor as an intracellular sentry. *Nucleic Acids Res.* **31**, 1024–1032.
51. Watson, P., Jones, A. T., and Stephens, D. J. (2005) Intracellular trafficking pathways and drug delivery: Fluorescence imaging of living and fixed cells. *Adv. Drug Delivery Rev.* **57**, 1024–1032.
52. Leinweber, F.-J. (1987) Possible physiological roles of carboxylic ester hydrolases. *Drug Metab. Rev.* **18**, 379–439.
53. Runquist, E. A., and Havel, R. J. (1991) Acid hydrolases in early and late endosome fractions from rat liver. *J. Biol. Chem.* **266**, 22557–22563.
54. Homick, C. A., Thouron, C., Delamatre, J. G., and Huang, J. (1992) Triacylglycerol hydrolysis in isolated hepatic endosomes. *J. Biol. Chem.* **267**, 3396–3401.
55. Kislinger, T., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M. S., Gramolini, A. O., Morris, Q., Hallett, M. T., Rossant, J., Hughes, T. R., Frey, B., and Emili, A. (2006) Global survey of organ and organelle protein expression in mouse: Combined proteomic and transcriptomic profiling. *Cell* **125**, 173–186.
56. Foster, L. J., de Hoog, C. L., Zhang, Y., Zhang, Y., Xie, X., Mootha, V. K., and Mann, M. (2006) A mammalian organelle map by protein correlation profiling. *Cell* **125**, 187–199.
57. Horgan, D. J., Dunstone, J. R., Stoops, J. K., Webb, E. C., and Zemer, B. (1969) Carboxylesterases (EC 3.1.1). The molecular weight and equivalent weight of pig liver carboxylesterase. *Biochemistry* **8**, 2006–2013.
58. Lakowicz, J. R. (1999) *Principles of Fluorescence Spectroscopy*, 2nd ed., Kluwer Academic/Plenum, New York.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California-Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

The Forefront of Chemical Biology

Scientists in Japan have a strong publishing relationship with the American Chemical Society (ACS). This is because Japan has a vibrant community of researchers exploring synthetic mechanisms, natural-products chemistry, structural biology, nanotechnology, materials science, and chemical engineering. The strength of their science was particularly evident at the inaugural meeting of the Japanese Society for Chemical Biology.

The 2-day meeting took place in Tokyo and attracted participants from across the country and a few from the U.S., Europe, and China. The research spanned diverse fields, from synthetic organic chemistry, to pharmacology, to RNA biochemistry. Many of the talks were presented in Japanese, but there were ample opportunities during the poster sessions and reception to discuss the science in English and at a more leisurely pace. This editor was particularly impressed with the poster presenters, who did an exceptional job of professionally communicating their work in both Japanese and English.

Why establish a Japanese Society for Chemical Biology? The organizers of the meeting, T. Nagano (University of Tokyo), M. Hagiwara (Tokyo Medical and Dental University), and H. Osada (RIKEN) saw the need to develop a formal means by which chemical biologists, including junior scientists, could interact and exchange data (for details, see ref 1). Until now, these scientists have had few opportunities to come together and discuss their work; researchers in Japan are dispersed in many professional societies which hold their own annual meetings with a separate focus on chemistry. This system required chemical biologists in Japan to separate the different aspects of their work and present them at two, or possibly more, venues. Seeing the need for a new way to discuss chemistry, biology, physics, and engineering, at one venue, the organizers established the new society for discussing and exchanging information on chemical biology.

Should the U.S. and other countries each have their own chemical biology societies, or should an international forum for the chemical biology community be established? One advantage of the former is that each can serve the unique needs of its constituents. A disadvantage of this strategy is that the community of chemical biologists may become too dispersed. Furthermore, ACS offers a well-established home for researchers at the interface of chemistry and biology: the Division of Biological Chemistry (www.biochemdivision.org). Now with the establishment of the Japanese Society for Chemical Biology, perhaps it's time to consider how we can bring all chemical biologists together to strengthen the field and help it grow.

With a successful launch of the Japanese Society for Chemical Biology, Japanese scientists are poised to make great strides in chemical biology. This journal expects to bring you many of these exciting papers and to continue to expand the publishing relationship between ACS and scientists in Japan.



Evelyn Jabri
Executive Editor

1. Osada, H. (2006) The Japanese Society for Chemical Biology, *ACS Chem. Biol.* 1, 8.

In this ISSUE

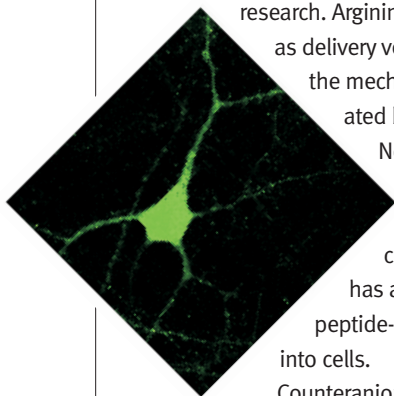
Negatively Affecting Translocation

Effective transport of biomolecules across cell membranes has important implications both therapeutically and in basic research. Arginine-rich peptides have recently found utility as delivery vectors for membrane translocation, but the mechanism of transport, which is likely mediated by endocytosis, is not well-understood.

Now, Takeuchi *et al.* (p 299) demonstrate that the addition of a hydrophobic molecule containing a negatively charged counteranion, such as pyrenebutyrate, has a dramatic, favorable effect on arginine-peptide-mediated translocation of biomolecules into cells.

Counteranion-mediated translocation is based on the concept that the negatively charged hydrophobic molecule interacts with the positively charged arginine peptide, further increasing its net hydrophobicity and facilitating direct

transport through the lipid bilayer. Fluorescence microscopy was used to monitor the translocation of several biomolecules differing in size and structure into a variety of cell types, including the difficult-to-transfect neuronal primary cells. Rapid, diffuse cytosolic and nuclear labeling resulted when cells were incubated with pyrenebutyrate and either alexa 488-labeled octaarginine (R8), enhanced GFP containing R8 at its N-terminus, or the apoptosis-inducing peptide (pro-apoptotic domain peptide, PAD) tagged with R8. Notably, delivery of R8-PAD provoked the expected biological response of mitochondrial-membrane disruption and eventual apoptosis. Very few endosome-like punctate signals formed during the transport process, which also proceeded at 4 °C; this suggests that direct membrane translocation, and not endocytosis, is the mechanism at play. The versatility of this counteranion-mediated translocation method should enable its widespread application for membrane transport.

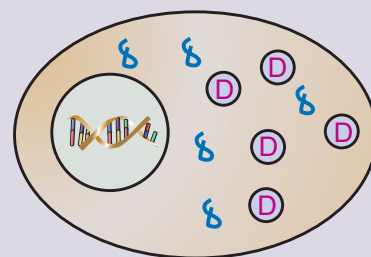


Back to the Basics for Cancer

One strategy in the fight against cancer is to design drugs that will selectively target cancer cells over normal cells. In an innovative approach, Duvvuri *et al.* (p 309 and Point of View p 271) describe a line of attack that exploits the reduced acidification of lysosomes in cancer cells relative to that of normal cells.

The authors rationalized that compounds with weakly basic pKa's would be trapped in the acidic (pH ~ 5) lysosomes of normal cells but be about equally distributed between the cytoplasm and the lysosome of cancer cells; thus, selectivity is conferred for cytosolic targets in cancer cells over normal cells. The validity of this approach was demonstrated with inhibitors of

the molecular chaperone heat-shock protein 90, an important cancer target. Five inhibitors with varying pKa's were tested in two cell lines: HL60 cells, which represent typical cancer cells possessing near-neutral lysosomes, and a multidrug-resistant (MDR) HL60 variant, whose lysosomes have an acidic pH similar to that of normal cells. Indeed, inhibitors with slightly basic pKa's exhibited preferential activity in the HL60 cells, where the compounds are found in higher concentrations in the cytoplasm than in the MDR HL60 cells. Neutral inhibitors did not exhibit any selectivity between the HL60 and the MDR HL60 cells and had IC₅₀ values comparable to the slightly basic inhibitors in the



MDR HL60 cells. Examination of the intracellular distribution of the compounds confirmed that the weakly basic inhibitors were present at much higher levels in the lysosomes of the MDR HL60 cells than in the neutral lysosomes of the HL60 cells. These results point to this strategy as a compelling approach for crafting selectivity into potential anticancer agents.

Published online June 16, 2006.

10.1021/cb600235s CCC: \$33.50

© 2006 by American Chemical Society

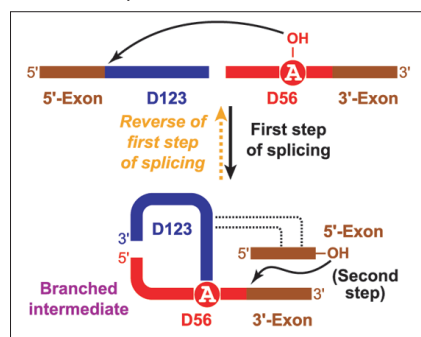
The Proof Is in the Second Step

Group II intron RNAs catalyze a two-step self-splicing reaction in which the intron is excised as a lariat and the exons are ligated. In the first step, the 2'-hydroxyl from an internal adenosine attacks the 5'-splice

site phosphodiester bond, forming a looped intron intermediate and releasing the 5'-exon. The 5'-exon then attacks the 3'-splice site, creating the ligated exon and liberating the intron. The first reaction is

reversible, and it has been hypothesized that the reversibility serves as a "proof-reading" mechanism in the event that an incorrect 5'-splice site is selected. Now, Wang and Silverman (p 316) use synthetic branched RNAs to provide evidence that a "trimming" process in the second step of the reaction, not the reversibility of the first step, functions as a proofreading process for mis-spliced 5'-exons.

The researchers use artificial DNA enzymes termed deoxyribozymes to create several 2',5'-branched RNA variants derived from group II self-splicing reactions mis-spliced at the 5'-splice site. The researchers observed that, while the natural splice-site intermediate underwent reverse splicing in the first step, none of the incorrect variants did. Moreover, when the second step of the reaction was examined with a branched RNA analogous to a splice variant with one extra base, the superfluous nucleotide was trimmed to provide the 5'-exon of correct length, and the ligation reaction proceeded normally. The authors propose that, given the mechanistic similarities between group II intron splicing and the spliceosome, the proofreading mechanism of 5'-exon hydrolysis followed by exon ligation may function in the spliceosome as well.

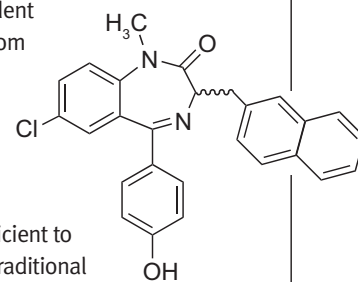


Uncompetitive Inhibition Wins Out

Mitochondrial F_1F_0 -ATPase, the enzyme responsible for generating most of the ATP in eukaryotic cells, could be an important therapeutic target for immune disorders; however, toxic side effects from inhibitors, such as oligomycin, have hindered clinical development of compounds that bind this target. Interestingly, Bz-423, a 1,4-benzodiazepine that binds to the oligomycin-sensitivity-conferring protein subunit of the enzyme, induces apoptosis and slows disease progression without toxicity. Johnson *et al.* (p 304) investigate the mechanisms of oligomycin and Bz-423 to determine the molecular basis for the differing biological effects of the two inhibitors.

Inhibition of mitochondrial F_1F_0 -ATPase can result in a decrease in cellular ATP levels, which can be toxic, and mitochondrial transition from active to resting respiration, which can induce apoptosis. The authors determined that oligomycin is a high-affinity, noncompetitive, time-dependent inhibitor with a slow rate of dissociation from

the enzyme. The combination of these properties results in potent inhibition of the enzyme and drastically reduced cellular ATP levels, which lead to toxic effects. In contrast, Bz-423 is a lower-affinity, uncompetitive inhibitor with a fast dissociation rate. The uncompetitive nature of inhibition of Bz-423 has little effect on ATP synthesis, yet it is sufficient to trigger the respiratory transition that signals for apoptosis. Traditional drug-discovery efforts focus on high-affinity, competitive inhibitors of drug targets. However, the results from this study emphasize how allosteric, lower-affinity inhibitors can have beneficial physiological consequences and identify an important strategy to consider in the design of potential drugs.



Translating Evolution to Inhibition

Aminoacyl-tRNA synthetases (aaRSs) are an essential component of the protein synthesis machinery; they catalyze transfer of specific amino acids to appropriate tRNAs prior to delivery to the ribosome for translation. Their critical role in cell survival, coupled with their diversity and specificity, makes these enzymes intriguing antimicrobial targets. Ataide and Ibbá (p 285) review the mechanisms behind the discriminatory ability of aaRSs and suggest various strategies for the design of effective small-molecule inhibitors.

AaRSs utilize several mechanisms to ensure that the correct amino acids are selected during protein synthesis. Some aaRSs have exquisitely evolved active sites, and others use secondary sites or additional proteins that can recognize and remove incorrect amino acids. In addition, some organisms have alternate biosynthetic pathways for amino acid tRNAs, and others possess duplicate aaRSs that can circumvent the action of potential inhibitors. Investigations with known inhibitors have shed light on the discrimination mechanisms, but much-needed additional structural and functional characterization of these fascinating enzymes will facilitate the design of effective new inhibitors and potential antimicrobial agents.

Spotlight

Elusive Couple Finally Captured

In the eukaryotic nucleus, RNA polymerase II (Pol II) and a vast collection of accessory proteins recognize DNA elements and then initiate and transcribe pre-messenger RNAs (pre-mRNAs). Before an RNA is ripe for translation, it must be processed in a number of ways, including 5' capping, splicing, 3' cleavage, and polyadenylation. A wealth of recent studies indicates that these events are probably not as linear as once envisioned. Rather, a web of partnerships exists between the RNA synthesis machinery and various processing machines. Studying these connections has proved difficult for the biochemist because of the



Courtesy of Getty images

lack of a reliable system where multiple activities can be observed. Now, two groups have optimized an efficient and reliable coupled-transcription/splicing system in a nuclear extract derived from a human cell line. Both groups show that the addition of a DNA template containing the proper sequence elements results in transcription as well as splicing.

Studies of splicing in a test tube have typically involved the addition of a radiolabeled RNA synthesized beforehand by a bacteriophage RNA polymerase. Hicks *et al.* (*PLoS Biol.* 2006, 4, e147) assemble reactions with template DNA containing either a promoter specific for the bacteriophage polymerase, T7, or the eukaryotic polymerase, Pol II. Adding exogenous T7 RNA polymerase to the nuclear extract made both promoters competent for transcription, but a striking difference was observed in splicing. Transcripts synthesized by Pol II entered the splicing pathway efficiently, whereas T7 transcripts accumulated but displayed little splicing. The Pol II transcripts also displayed greater resistance to nucleases. This resistance was dependent upon splicing signals in the pre-mRNA and on the integrity of U2 small nuclear RNA, a spliceosomal RNA employed early in the splicing pathway.

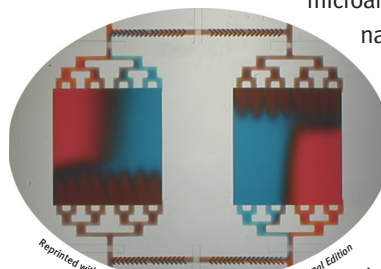
(continued on page 265)

Chaos Helps

Hybridization of DNA is a routine step in a wide variety of methods used to analyze gene sequences and expression levels. The development of DNA microarrays, where thousands of nucleic acid strands can be analyzed on the surface of a single glass microscope slide, has enabled rapid, large-scale detection of DNA. However, the hybridization step in the process is performed in a diffusion-limited manner, and the distance that DNA can diffuse in the time allotted for hybridization is an order of magnitude shorter than the typical length of most microarrays. Now, Liu *et al.* (*Angew. Chem., Int. Ed.* 2006, 45, 3618–3623) present a method to accelerate the DNA hybridization process by using microfluidic chaotic mixing.

A microfluidic device was designed and constructed to facilitate effective circulation and mixing of the hybridization solution while maintaining compatibility with the microarray format. The silicone rubber device, which can be sealed onto a microarray slide, contains two symmetric hybridization chambers integrated with peristaltic pumps that circulate the fluid between the chambers. The chambers are connected by bridge channels, where herringbone patterns on the ceiling continuously introduce the chaotic mixing of the solution. Bifurcating channels equalize the mixed solution distributed into the chambers.

Assessment of the efficacy of the device demonstrated that hybridizations subjected to chaotic mixing resulted in microarrays with stronger signals, enhanced sensitivity, reduced spot-to-spot variability, improved signal specificity, and an increase in molar-hybridization events over static hybridizations. Notably, hybridizations in which fluid



Reprinted with permission from *Angewandte Chemie, International Edition*

was circulated but not subjected to chaotic mixing revealed that chaotic mixing was a significant contributor to the signal enhancement. The superior data quality, higher efficiency, and compatibility with current microarray protocols provided by the device champion its incorporation into microarray protocols. **EG**

Tackling Tuberculosis

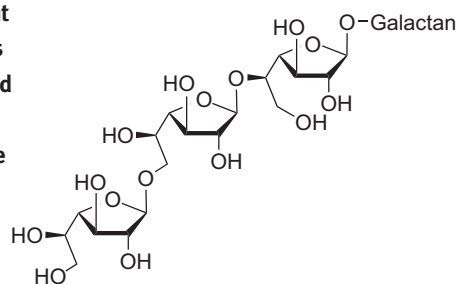
More than one-third of the world's population is infected with *Mycobacterium tuberculosis*, the bacteria that causes tuberculosis (TB), and the emergence of strains that are resistant to antibiotic treatment has fueled interest in finding new drugs to combat the disease. The cell wall of *M. tuberculosis* is a particularly attractive drug target because the organism depends on the integrity of its cell wall for survival. Moreover, many of the enzymes used in cell-wall biosynthesis are not found in humans, so potential drug side effects are minimized. Rose *et al.* (*JACS* 2006, 128, 6721–6729) report the expression, purification, and characterization of a galactosyl-transferase (GT) involved in *M. tuberculosis* cell-wall biosynthesis.

The major structural component of the *M. tuberculosis* cell wall is a lipidated

polysaccharide, the mycoyl–arabinogalactan–peptidoglycan (mAGP) complex. The carbohydrate-based core of mAGP is composed of ~30 D-galactofuranose residues attached via alternating β -(1 \rightarrow 5) and β -(1 \rightarrow 6) linkages. A GT responsible for constructing these linkages, termed *glfT*, was expressed in *Escherichia coli*, and several potential di- and trisaccharide glycosyl acceptors were used to examine the substrate specificity of the enzyme. Characterization of the products by NMR and MS revealed that the recombinant enzyme, like the wild-type enzyme, is capable of creating both β -(1 \rightarrow 5) and β -(1 \rightarrow 6) linkages. In addition, both disaccharides and trisaccharides were substrates for the enzyme, but kinetic analysis revealed that the trisaccharides' acceptors

were better substrates than those of the disaccharides.

The insight into the biosynthesis of the mAGP complex provided by this study enabled the authors to propose that ≤ 3 additional enzymes are required to generate the substrate for *glfT* *in vivo*. Further characterization of *glfT* and other enzymes in the biosynthetic pathway will facilitate discovery of inhibitors of *M. tuberculosis* cell-wall biosynthesis, paving the way for new treatments for TB. EG



Reprinted with permission from the *Journal of the American Chemical Society*

Elusive Couple Finally Captured, *continued*

Finally, the researchers tested an alternative splicing pre-mRNA containing one set of 5' splicing signals, but two sets of 3' splicing signals. Interestingly, splice site choice was rather different when Pol II synthesis was feeding the pre-mRNA to the spliceosome.

In a similar effort, Das *et al.* (*Genes Dev.* 2006, 20, 1100–1109) optimize a coupled reaction and compare Pol II with T7-synthesized transcripts in their splicing fates. Again, the efficiency of splicing was far better with the eukaryotic polymerase driving pre-mRNA production. This study looked closely at spliceosomal complex formation on the pre-mRNA and at how the polymerase identity affected these events. Remarkably, the pre-mRNAs produced by Pol II were nearly quantitatively shuttled into spliceosomes as they were synthesized, whereas T7 transcripts remained in heterogeneous RNA–protein

complexes, which hinder splicing. This transfer from Pol II to the spliceosome depended upon intact splice site signals. Additionally, this study indicated that splicing kinetics are far more rapid with Pol II pre-mRNAs.

Both of these studies demonstrate biochemically that a functional handshake occurs between two nuclear machines, the Pol II complex and the spliceosome. With these reliable *in vitro* systems, the hands that do the shaking can finally be sought and characterized. Of particular interest will be the phosphorylated carboxy-terminus of Pol II, a protein domain that seems to mediate crosstalk with the processing machines. The stark difference between Pol II and T7 polymerase lends itself nicely to various domain-swapping experiments. Finally, this powerful methodology will offer scientists new perspectives on the mysterious realm of alternative splice site choice. JU

Microengraving and Multiplexing Monoclonals

Monoclonal antibodies have found powerful applications in therapeutics and as tools for detection of biomolecules. The process of identification and retrieval of cells that produce antigen-specific antibodies, however, is notoriously labor-intensive and time-consuming. Now, Love *et al.* (*Nat. Biotechnol.* 2006, 24, 703–707) describe a microengraving method that enables rapid identification, recovery, and clonal expansion of monoclonal antibody-generating cell lines.

The microengraving method is designed so that the secreted antibodies of individual cells arrayed in high-density microwells are deposited on a microarray surface. Subsequent analysis of the microarray pinpoints the location of the corresponding antibody-secreting cell, which can then be extracted from the microwell and expanded in culture. The efficacy of the microengraving approach was demonstrated when cells were identified that produce antibodies to major histocompatibility complex type I (H-2K^b).

Hybridoma cells were generated from mice immunized with peptide-loaded H-2K^b/streptavidin tetramers and arrayed into microwells fabricated on slabs of poly(dimethylsiloxane). The slabs were sealed onto a microarray surface coated with appropriate secondary antibody, and secreted antibodies were captured and detected with fluorescently labeled H-2K^b. Out of 200,000 cells screened, 4300 positive spots were generated on the microarray, and of the 50 corresponding cells that were arbitrarily selected for expansion, 17 active hybridoma supernatants were identified.

This approach dramatically improves several aspects of monoclonal antibody generation. First, early screening of the cells eliminates the tedious serial dilution process for achieving monoclonality and the need to maintain several potential cell lines during antibody

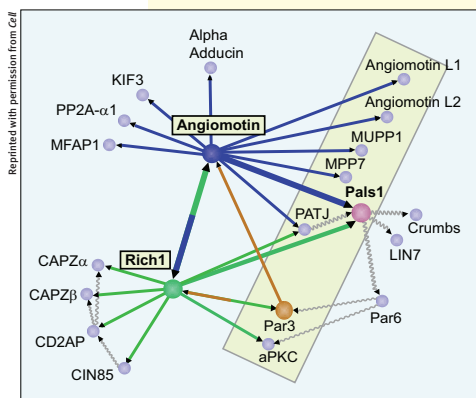
Polarity Networks Get Rich

Cell polarity is a complex process that requires the intricate coordination of many different biomolecular interactions. The Rho-GTPase Cdc42 is involved in maintaining polarity in epithelial cells, but the variety of processes that require Cdc42 has precluded a more precise understanding of its role in cell polarity. It is known that Cdc42 function is regulated by numerous guanine exchange factors (GEFs) and GTPase-activating proteins (GAPs), and now Wells *et al.* (*Cell* 2006, 125, 535–548) describe a protein interaction network surrounding the Cdc42 GAP Rich1 that links Cdc42 to cell polarity.

In order to find proteins that interact with Cdc42 in the context of cell polarity, the authors

used FRET and tandem MS to screen 50 GEFs and 50 GAPs for their ability to alter the GTPase cycle of Cdc42 and to bind to known polarity proteins. The GAP Rich1 was found to associate with polarity proteins at tight junctions (TJs) (complexes that form a seal between epithelial cells), and it was determined that the GAP activity of Rich1 is required for proper TJ maintenance, perhaps by preventing accumulation of Cdc42-GTP at the TJs.

Tandem MS was used to further probe the polarity protein interaction network to reveal additional Rich1 interactions. Rich1 associates with the scaffolding protein angiomin (Amot) through the coiled-coil domain of Amot. In addition, Amot associates with Patj, a protein involved in establishing apical junctions, through the PDZ domain binding motif of Amot. This interaction targets Amot to TJs and links Rich1 to Patj, thereby targeting Rich1 to Cdc42 at TJs. The authors propose that Rich1 and Amot help maintain TJ integrity by regulating Cdc42 activity at TJs and by integrating Cdc42 function with the trafficking of intracellular polarity proteins at the TJ. **EG**



characterization. In addition, the method facilitates isolation of slow-growing and rare clones that are not easily accessed via traditional methods. Finally, clones with unique specificities can be rapidly identified when they are multiplexed with differentially labeled antigens in the microarray analysis. This innovative method could also have widespread applications in the screening and characterization of cells that secrete other types of molecules, such as cytokines. **EG**

Signaling Brain-Energy Balance

Maintaining an organism's energy balance and food intake is a critical function of the central nervous system, the circuitry of which has been shown previously to sense glucose and fatty acid levels directly.

Emerging data also suggest that specific populations of neurons sense fuel availability through

hormones such as leptin, a small protein known to influence hunger, food consumption, and energy expenditure. Now Cota *et al.* (*Science* 2006, 312, 927–930) show that the hypothalamus senses L-leucine and responds to changes in leptin levels. These responses are

governed by the serine–threonine kinase mammalian target of rapamycin (mTOR).

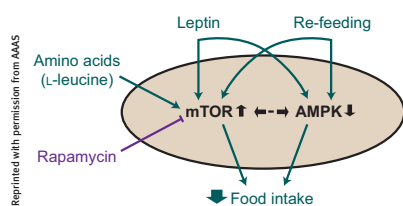
In the presence of mitogens and nutrients, mTOR stimulates cell growth, and aberrant over-activity of this protein has been associated with cancer, diabetes, and obesity.

The authors observed that fasted rats exhibited decreased levels of phosphorylated S6 kinase, a target of mTOR. They then showed that treatment of fasted rats with L-leucine decreased the amount of food these rats ate and contributed to weight reduction. Furthermore, co-treatment of rats with L-leucine and the mTOR inhibitor rapamycin eliminated the L-leucine-

induced anorexia; this suggests that mTOR responds directly to amino acid levels.

The researchers further showed mTOR to be a nutrient sensor by examining the effect of the cytokine leptin on the hypothalamus. Treatment with leptin increased the levels of phosphorylated S6 kinase and induced anorexia; this suggests that mTOR is involved in this process. mTOR's participation was confirmed by co-administration of leptin and rapamycin, which reversed the anorexia.

Further studies of this and other fuel-sensing pathways may provide insight into the relationships between obesity and type 2 diabetes. ST

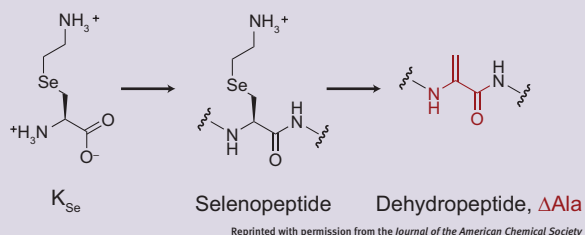


Nonribosomal Peptides from the Ribosome

Nonribosomal peptides (NRPs) are chemically diverse linear or cyclic molecules used extensively in therapeutics. NRPs are synthesized by unicellular organisms, plants, and fungi using a large enzyme complex with a common core structure and many different modules to chemically change the evolving product. This synthetic mechanism has forced researchers to rely on complex solid-phase synthetic schemes

to incorporate the variety of L-, D-, β-, N-methyl, and α, β-unsaturated amino acids into NRPs. Seebeck and Szostak (*JACS* 2006, 128, 7150–7151) now use a simple and versatile *E. coli* translation system to synthesize dehydroalanine-containing linear and cyclic peptides from messenger RNA templates.

Ribosomal peptide synthesis needs an appropriately charged transfer RNA (tRNA), which requires tRNA synthetases to use non-natural amino acids as substrates. The authors show that selenalysine is a substrate for lysine aminoacyl-tRNA-synthetase and can be incorporated into



peptides by a reconstituted translation system from *E. coli*. Oxidative elimination converts the linear peptide into dehydropeptide, generating a scaffold with which many different nucleophiles can react. The authors show that such model dehydropeptides can also be synthesized as a stable cyclic structure. This simple ribosome-based system to make a reactive scaffold provides a mechanism to generate highly decorated peptides without the use of solid-phase synthesis. Furthermore, this synthetic scheme is amenable to high-throughput screening techniques and opens the door to the selection of bioactive peptide compounds. EJ

Metabolizing Metabolites

Thiamin pyrophosphate is an essential coenzyme involved in the metabolism of carbohydrates and amino acids. The chemical structure of thiamin consists of a thiazole portion and a pyrimidine portion. Although it is known that thiamin biosynthesis in bacteria employs five different enzymes, only thiazole synthase (Thi4) has been found in eukaryotes thus far. Chatterjee *et al.* (*JACS* 2006, 128, 7158–7159) identify and characterize a metabolite of Thi4 from *Saccharomyces cerevisiae*, providing clues to the biosynthetic pathway of this important molecule in eukaryotes.

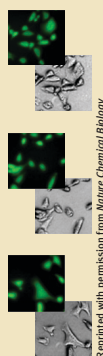
The role of Thi4 in thiamin biosynthesis is not known, but it was observed that denaturation of Thi4 resulted in the release of four major enzyme metabo-

lites. One of the metabolites was stable enough to be characterized, and several analytical tools were used to piece together its structure. Sequence analysis of the enzyme and UV absorption properties of the metabolite suggested that the compound was adenylated. HPLC, NMR, and MS experiments enabled the researchers to propose a structure that contained an adenosyl and a thiazole moiety. To confirm the presence of the thiazole, the authors subjected the metabolite to a series of reactions that yielded an intensely fluorescent thiochrome phosphate, a compound that would be generated only if a thiazole was originally present. The elucidation of the metabolite's structure allowed the authors to propose a mechanism for its biosynthesis wherein it is derived from NAD, and chemistry similar to that involved in ADP ribosylation contributes to the creation of the thiazole functionality. These results demonstrate the feasibility of examining enzyme-bound metabolites to help decipher the function of enzymes with unknown activity. **EG**

Yes to NO Detection

Nitric oxide (NO) is an important signaling molecule involved in many biological processes, including immune response, neurotransmission, and blood-pressure regulation. NO is one of the few known gaseous signaling molecules, and its direct detection *in vivo* has been notoriously difficult because of its rapid diffusion and reactivity in the cell. Current NO detection methods are available, but they suffer from indirect measurement of NO or low spatial resolution, or they require complex instrumentation. Now, Lim *et al.* (*Nat. Chem. Biol.*, published online May 28, 2006, doi: 10.1038/nchembio794) describe the synthesis, characterization, and biological application of a fluorescent probe that enables rapid, direct, and specific detection of NO in live cells.

The NO probe was designed so that exposure of a relatively nonfluorescent, cell-permeable reagent to NO results in the generation of a fluorescent compound that could be visualized in real time by fluorescence microscopy. To this end, a fluorescein derivative was reacted with CuCl_2 to generate a Cu(II) fluorescein (CuFL) species with reduced fluorescence properties. Reaction of CuFL with NO results in a nitrosated fluorescein (FL-NO) derivative. Characterization by EPR, UV-vis, NMR, fluorescence spectroscopy, and LC/MS revealed that exposure of NO to CuFL induces the reduction of Cu(II) to Cu(I) and the concomitant



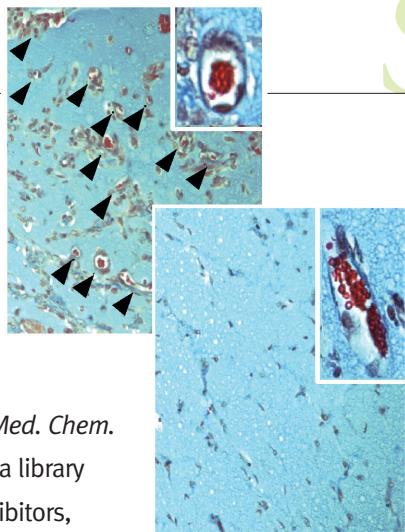
formation of NO^+ , which rapidly reacts to form FL-NO. Dissociation of FL-NO from the Cu(I) species results in the fluorescence increase.

The utility of CuFL to detect NO was tested in two cell lines. In human neuroblastoma cells, estrogen activates constitutive nitric oxide synthases (NOSs) to produce NO. Simultaneous addition of estrogen and CuFL to these cells resulted in an NO-dependent fluorescence response. Similarly, inducible NOSs are activated in macrophages treated with lipopolysaccharide and interferon- γ , and addition of CuFL to these cells also resulted in an NO-dependent increase in fluorescence. These results demonstrate the power of CuFL as a direct NO-detection reagent for a variety of applications. **EG**

Teaching Old Drugs New Tricks

Bringing a new drug to market takes nearly 15 years and almost \$1 billion. Efforts to speed up the process and decrease the costs have led researchers to search for new applications for known drugs. Now, Chong *et al.* (*J. Med. Chem.* 2006, 49, 2677–2680) report the screening of a library of 2450 known drugs for new angiogenesis inhibitors, identify the immunosuppressant mycophenolic acid (MPA) as an inhibitor of endothelial-cell growth, and determine its mechanism of action and efficacy as an antiangiogenic agent.

MPA is commonly used to prevent organ-transplant rejection, and its activity as an inhibitor of inosine monophosphate dehydrogenase (IMPDH), an enzyme involved in *de novo* guanine biosynthesis, is well-established. However, the mechanism of action and molecular target of MPA as an angiogenesis inhibitor needed to be determined. Examination of the effects of MPA on cell proliferation and cell-cycle progression in human umbilical vein endothelial cells (HUVECs) indicated that, as in T and B cells, MPA inhibits HUVEC growth in a guanosine-dependent manner and causes cell-cycle arrest in G1. Moreover, examination of the effects of selectively



knocking down both isoforms of IMPDH (1 and 2) in HUVEC by RNA interference revealed that loss of IMPDH-1 function is sufficient to cause cell-cycle arrest in G1. In contrast, knocking down IMPDH-1 has no effect on T-cell proliferation. Further, an *in vivo*

angiogenesis assay demonstrated that mice treated with MPA had significantly less new blood vessel formation than control mice. Finally, a murine renal cell carcinoma model was used to determine that MPA prevents tumor-induced angiogenesis. Taken together, the data validate IMPDH-1 as the MPA target in endothelial cells, and given that IMPDH-2 appears to be the more essential protein for T-cell proliferation and development, the drug-target potential of IMPDH-1 is substantiated. These results highlight the power of this drug-discovery strategy and point to specific inhibitors of IMPDH-1 as possible antiangiogenic drugs. The library of existing drugs reported in this paper will be made available to the scientific community for screening on other drug targets through collaborative agreements or other mutually agreeable arrangements. **EG**

UPCOMING CONFERENCES

Bioorganic Chemistry, GRC

July 30–August 4, 2006
Oxford, U.K.

Annual Meeting of the Protein Society

August 5–9, 2006
San Diego, CA

Post-Transcriptional Gene Regulation, GRC

August 13–18, 2006
Oxford, U.K.

Protein Phosphorylation and Cell Signaling

August 18–22, 2006
La Jolla, CA

Multiprotein Complexes Involved in Cell Regulation, Keystone

August 18–23, 2006
St. John's College, U.K.

Translational Control, CSHL

September 6–10, 2006
Cold Spring Harbor, NY

Spotlights written by Eva Gordon, Evelyn Jabri, Sarah Tegen, and Jason Underwood

The Great Multidrug-Resistance Paradox

Vivien Y. Chen and Gus R. Rosania*

Department of Pharmaceutical Sciences, University of Michigan College of Pharmacy, 428 Church Street, Ann Arbor, Michigan 48109

Heat-shock proteins (Hsp) are a family of proteins that facilitate cells' survival when they encounter stressful conditions, ranging from heat to toxic chemicals. One of these proteins, Hsp90, is a molecular chaperone that facilitates the folding of aggregated or otherwise misfolded proteins in the cytosol. Geldanamycin (GM; Figure 1) is a small-molecule natural product that binds to and inhibits Hsp90 in mammalian cells. GM is selectively toxic to certain cancer cells. This suggests that Hsp could serve as a molecular target for candidate anticancer drugs. To elucidate the molecular mechanisms rendering certain cancer cells more (or less) sensitive to Hsp90 inhibition, medicinal chemists are synthesizing and studying even more selective GM derivatives. These compounds are interesting because they could be good chemotherapeutic agents. They are also noteworthy because their chemical structure and mechanism of action may be keys to understanding the reason some cell lines respond differently to some small molecules than to others.

The article by Duvvuri *et al.* (1) on page 309 of this issue of *ACS Chemical Biology* highlights a relationship between the subcellular transport properties of a family of GM derivatives and their cell-growth inhibitory activity against multidrug-resistant (MDR) cancer cells. These particular MDR cells lack significant expression of drug efflux pumps. They possess highly acidic lysosomes but were derived from a cell line with neutral lysosomes. GM derivatives with weakly basic functional groups are proto-

nated and ionized at low pH. The neutral form of the molecule is membrane-permeant, but the charged ionic state of the molecule is membrane-impermeant and thus becomes trapped in acidic organelles, such as lysosomes (Figure 2). By identifying MDR clones with acidic lysosomes derived from a drug-sensitive parent cell line with neutral lysosomes, Duvvuri *et al.* (1) were able to address how pH-dependent ion trapping affects differential sensitivity to GM by relating physicochemical properties of various GM derivatives to their lysosomal sequestration and to their growth-inhibitory activity.

A cancer cell can become resistant to anticancer drugs in many ways, and one is to overexpress drug transporters at the plasma membrane to remove drugs from the cell (2). But, do transporters confer drug resistance by preventing drug molecules from accumulating inside the cell? As observed by Duvvuri *et al.* (1), comparing the total intracellular drug mass in MDR cells often reveals a decrease in intracellular drug content that is not sufficient to explain the resistant phenotype. In other studies looking at the cellular pharmacokinetics of drug-resistant cancer cells, intracellular drug concentrations have been measured and found to be much greater than extracellular concentrations (3, 4). This is the great multidrug-resistance paradox. So, if the amount of drug present in the cell cannot account for how drug efflux pumps affect drug resistance, then what does? As shown by Duvvuri *et al.* (1), drug-resistant cells that sequester drug in intracellular compartments

ABSTRACT Much of the attention devoted to the elucidation of multidrug-resistance mechanisms in tumor cells has focused on transmembrane drug transporters and their ability to pump drug molecules from the cytosol to the extracellular medium. However, intracellular drug concentrations often remain high in drug-resistant cells and therefore do not explain how drug pumping at the plasma membrane confers multidrug resistance. Recent work indicates how drug sequestration in cytoplasmic organelles can account for these paradoxical results and how cellular pharmacokinetics may be exploited to target the activity of small molecules to specific cell types.

*To whom correspondence should be addressed.
E-mail: grosania@umich.edu.

Published online June 16, 2006
10.1021/cb600215q CCC: \$33.50
© 2006 by American Chemical Society

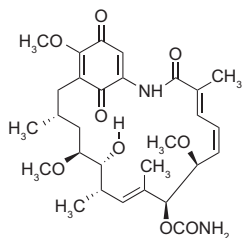


Figure 1. Structure of geldanamycin (GM). GM inhibits mammalian Hsp90 and is selectively toxic to certain cancer cells.

are killed by the drug if the sequestration mechanism is perturbed by making lysosomes more alkaline. Intracellular drug sequestration is therefore a plausible explanation for the great multidrug-resistance paradox and could account for the selective cytotoxic activity of certain GM derivatives against specific cancer cell lines (Figure 2).

Beyond the observed structure–localization–activity relationship, the pH-dependent partitioning mechanism may be used prospectively to optimize the selectivity of anticancer agents against cancer cells. In the case of GM, the experimental results argue that localization of small molecules can be targeted to the intracellular site of action (the cytosol) by optimizing the pH-dependent octanol/water partition property of the molecules (1). Lipophilicity differences between different ionic states of a drug are good surrogates for the relative rates of transport of those ionic states across cell membranes, with hydrophilic ions being less membrane-permeant than more hydrophobic ions. For GM, a strong case can be made for why neutral molecules are preferable to weakly basic molecules when the goal is to avoid drug sequestration in lysosomes (1). This is also a good case for why molecules with a high octanol/water partition coefficient in their charged state may be less prone to sequestration and resistance than molecules that are more lipophilic when protonated and charged in an acidic microenvironment.

For understanding subcellular transport as a mechanism conferring differential sensitivity to anticancer agents, the role of transmembrane diffusive pathways in relation to active transport mechanisms governing microscopic drug distribution across membranes should be considered. Many studies have focused exclusively on the interaction between molecules and drug efflux pumps to look at how chemical structure links to drug resistance (5). However, in the case of hydrophobic molecules, drug–transporter interactions are less relevant, because hydrophobic molecules can short-circuit active transport mechanisms by diffusing directly across phospholipid bilayers. Nevertheless, when multidrug resistance cannot be explained by differences in the intracellular concentration of drug, then differences in the molecular interaction between the small molecule and its target, or differences in the biochemical pathways leading to cell death, are invoked.

However, when the focus is on the role of intracellular diffusive transport mechanisms in the differential activity of small molecules, attention shifts from how specific functional groups on a molecule affect interaction with specific proteins to the effect of local pH gradients on drug biodistribution. In tumor cells, pH gradients may be present between lysosomes and the cytosol, as well as between the cytosol and the extracellular tumor microenvironment (6, 7). At the lysosomal membrane, ATP-coupled proton transporters maintain the acidity of the lysosomal lumen. Tumors are often anaerobic, so glycolysis contributes directly to the acidification of the extracellular microenvironment (8). If pH is a major determinant of intracellular, microscopic drug distribution between lysosomes and cytosol, then pH could also affect the partitioning of drugs between cytosol and the extracellular microenvironment. For weakly basic, hydrophobic molecules, an acidic extracellular medium would also facilitate drug efflux from tumor cells. Hence, a

molecule’s site of action can be as much of a determinant of drug activity as its mechanism of action.

As for the broader relevance of ion trapping and other passive transport mecha-

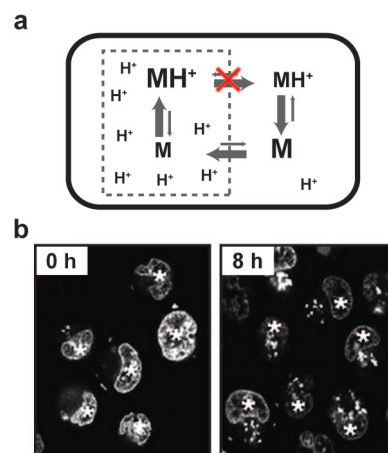


Figure 2. Mechanism and illustration of drug sequestration in the cytosol. a) The ion-trapping mechanism. A molecule contains a functional group that becomes ionized and charged at low pH. The neutral form of the molecule (M) is membrane-permeant, whereas the protonated charged form (MH⁺) is impermeant. At cytosolic pH, the molecule exists mostly in the M form, which is driven into the lysosomes by its concentration gradient across the lysosomal membrane. In the low-pH environment of the lysosome, the equilibrium distribution of the molecule is shifted to the MH⁺ form, which is unable to diffuse down its concentration gradient to the outside of the lysosome because it is membrane-impermeant. Thus, the trapped ion ends up accumulating at high concentrations in the lysosomes. b) The sequestration of doxorubicin in cytosolic compartments in K562 human erythroleukemic cells was directly imaged using confocal microscopy. Cells are pulsed with doxorubicin (whose intrinsic fluorescence allows direct observation) and then incubated in drug-free media to allow for efflux. Sequestration in cytoplasmic vesicles can be seen by comparing the images taken after 0 and 8 h following the doxorubicin pulse. Asterisks indicate cell nuclei. Cytoplasmic vesicles where doxorubicin is sequestered can be seen as bright dots in the nuclear periphery.

nisms, physicochemical properties influencing subcellular drug transport are attracting attention as a way to rationally optimize the activity and specificity of small, bioactive molecules in living cells. In the context of subcellular transport theory, cellular pharmacokineticists are developing computational models for calculating the intracellular distribution properties of small molecules as a function of chemical structure and physicochemical features (9–11). To the extent that many bioactive small molecules are hydrophobic and therefore distribute intracellularly through passive diffusion, the ion-trapping mechanism should be an important determinant of small-molecule activity and specificity for not only GM derivatives but also a wide variety of anti-cancer agents.

To conclude, the great multidrug-resistance paradox offers a potentially useful mechanism for targeting small molecules to specific subcellular locations. In the past few years, statistical methods for determining the significance of quantitative structure localization relationships between small molecules and subcellular distribution have been developed (12–14). In addition, it is now possible to measure intracellular drug concentrations in different subcellular compartments after organellar isolation and biochemical analysis (15, 16). With fluorescent molecules, high-throughput, microscopy-based screening instruments allow direct visualization of the relationship between chemical structure and subcellular distribution, both in fixed-endpoint and kinetic experiments, across large collections of compounds. As a result, a truly original and unique conceptual and experimental framework is emerging from these studies, with the site of action being increasingly recognized as a determinant of the activity and specificity of small molecules in living cells.

REFERENCES

1. Duvvuri, M., Konkar, S., Hong, K. H., Blagg, B. S. J., and Krise, J. P. (2006) A new approach for enhancing differential selectivity of drugs to cancer cells, *ACS Chem. Biol.* **1**, 309–315.
2. Longley, D. B. and Johnston, P. G. (2005) Molecular mechanisms of drug resistance, *J. Pathol.* **205**, 275–292.
3. Li, D., Jang, S. H., Kim, J., Wientjes, M. G., and Au, J. L. (2003) Enhanced drug-induced apoptosis associated with P-glycoprotein overexpression is specific to antimicrotubule agents, *Pharmacol. Res.* **20**, 45–50.
4. Tapiero, H., Nguyen-Ba, G., and Lampidis, T. J. (1994) Cross resistance relevance of the chemical structure of different anthracyclines in multidrug resistant cells, *Pathol. Biol.* **42**, 328–337.
5. Szakacs, G., Paterson, J. K., Ludwig, J. A., Booth-Genthe, C., and Gottesman, M. M. (2006) Targeting multidrug resistance in cancer, *Nat. Rev. Drug Discovery* **5**, 219–234.
6. Belhoussine, R., Morjani, H., Sharonov, S., Ploton, D., and Manfait, M. (1999) Characterization of intracellular pH gradients in human multidrug-resistant tumor cells by means of scanning microspectrofluorometry and dual-emission-ratio probes, *Int. J. Cancer* **81**, 81–89.
7. Vaupel, P., Kallinowski, F., and Okunieff, P. (1989) Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: A review, *Cancer Res.* **49**, 6449–6465.
8. Yamagata, M., Hasuda, K., Stamato, T., and Tannock, I. F. (1998) The contribution of lactic acid to acidification of tumours: Studies of variant cells lacking lactate dehydrogenase, *Br. J. Cancer* **77**, 1726–1731.
9. Jang, S. H., Wientjes, M. G., and Au, J. L. (2003) Interdependent effect of P-glycoprotein-mediated drug efflux and intracellular drug binding on intracellular paclitaxel pharmacokinetics: Application of computational modeling, *J. Pharmacol. Exp. Ther.* **304**, 773–780.
10. Trapp, S., and Horobin, R. W. (2005) A predictive model for the selective accumulation of chemicals in tumor cells, *Eur. Biophys. J.* **34**, 959–966.
11. Zhang, X., Shedden, K., and Rosania, G. R. (2006) A chemoinformatic definition of chemical space occupied by ideal drug candidates with extracellular site of action, *Mol. Pharmacol.*, submitted for publication.
12. Rosania, G. R. (2003) Supertargeted chemistry: Identifying relationships between molecular structures and their sub-cellular distribution, *Curr. Top. Med. Chem.* **3**, 659–685.
13. Chen, V. Y., Khersonsky, S. M., Shedden, K., Chang, Y. T., and Rosania, G. R. (2004) System dynamics of subcellular transport, *Mol. Pharmacol.* **1**, 414–425.
14. Shedden, K., Brumer, J., Chang, Y. T., and Rosania, G. R. (2003) Chemoinformatic analysis of a supertargeted combinatorial library of styryl molecules, *J. Chem. Inf. Comput. Sci.* **43**, 2068–2080.
15. Duvvuri, M., Feng, W., Mathis, A., and Krise, J. P. (2004) A cell fractionation approach for the quantitative analysis of subcellular drug disposition, *Pharmacol. Res.* **21**, 26–32.
16. Chen, V. Y., Posada, M. M., Blazer, L. L., Zhao, T., and Rosania, G. R. (2006) The role of the VPS4a-exosome pathway in the intrinsic egress route of a DNA-binding anticancer drug, *Pharmacol. Res.*, in press.

Chemical Modifications Rescue Off-Target Effects of RNAi

Ola Snøve, Jr., and John J. Rossi*

Division of Molecular Biology, Graduate School of Biological Sciences, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, California 91010

ABSTRACT RNAi's specificity has been questioned for some time. Three recent papers show that off-target effects should be considered normal, but one paper also provides insight on how chemical modifications of siRNAs may overcome the problem.

RNA interference (RNAi) has quickly been accepted as the standard tool for sequence-specific gene silencing in molecular biology. Its perceived advantages over other antisense-based techniques were thought to be an exquisite efficacy and specificity. And while RNAi's efficacy remains impressive, its specificity does not hold up to initial expectations. Three years ago, Jackson *et al.* (1) provided the first evidence that small interfering RNAs (siRNAs) were capable of significant "off-targeting", which could result in sequence-specific and extensive silencing of nontargeted transcripts. The same group, along with investigators from Dharmacon Research (Lafayette, CO), has recently published three papers that confirm that RNAi is prone to cause off-target effects. One of the papers also provides encouragement that these problems can be solved, as chemical modifications to the siRNAs that mediate silencing significantly reduce unintended regulation (2–4).

Early on, siRNAs seemed highly specific, as a single mutation in the target site could be demonstrated to completely abolish silencing (5). Perhaps an early warning sign was a report showing that mutations in the target region, which corresponded to the central region of several siRNAs, did not always abrogate knockdown (6). Other reports revealed some tolerance to wobble base pairing (7, 8). Part of the reason a louder alarm did not go off may have been that the first microarray experiments

demonstrated excellent specificity for siRNAs (9, 10). Scientists were therefore caught off guard when Jackson *et al.* (1) reported severe off-target effects for genes with only a stretch of 11 nucleotides of sequence similarity between the target and the guide strand of the siRNA. These effects were clearly not related to the nonspecific interferon effect, as was observed for both siRNAs (11) and short hairpin RNAs (12), but dependent on the sequence of the siRNA.

As is often the case when a promising new technology receives its first major setback, scientists reacted with dismay. This was partly due to uncertainty about the interpretation of results that were obtained in siRNA experiments, but enthusiasm for therapeutic applications was also dampened. Lack of knowledge about the off-target mechanism precluded methods that solve the problem. The recent papers in *RNA* contribute to our understanding of the mechanisms involved.

Jackson *et al.* (1) tested 24 siRNAs against two genes in their initial study. Their recent paper supports the previous results with a dataset comprising six genes (4). Importantly, many of the genes that were down-regulated according to the microarrays had sequence complementarity to the siRNAs in the 3' UTRs of the transcripts. Moreover, the most enriched hexamers in the 3' UTRs were complementary to nucleotides 2–7, the seed region shown to be important for microRNA (miRNA) targeting (13, 14). Jackson *et al.* (1) note that

*Corresponding author.
E-mail: jrossi@bricoh.edu.

Published online June 16, 2006
10.1021/cb6002256 CCC: \$33.50
© 2006 by American Chemical Society

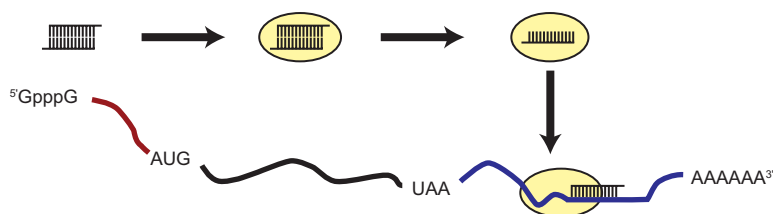


Figure 1. siRNA duplex is bound by RISC, one of the two strands is eliminated, and the remaining strand serves as guide for pairing to a complementary target. The 3' UTR of protein-coding genes (blue) can be targeted for down-regulation by the RNA-induced silencing complex (RISC; yellow) with only the seed sequence (bases 2–8) sharing complementarity with the target. When the siRNAs is completely paired with the target mRNA, all regions (5' UTRs, red; coding; and 3' UTRs) can be targeted, and the effector protein in RISC, Ago 2, can cleave the target. In the example shown, the likely outcome would be translational inhibition as opposed to site-specific cleavage.

hexamers in positions 1–6 and 3–8 were also highly enriched; this means that complementarity to the first 8 nucleotides of the 5'-end of siRNAs is most important for off-targeting. These results confirm previous reports in the literature (15–18). The authors compared messenger RNA (mRNA) levels at 24 h with protein levels at 48 h post-transfection and showed that protein levels drop as expected; hence, the results are not artifacts of the microarrays.

Knockdowns at the protein level were similar to what would be expected from the reduction in mRNA levels. Recent evidence shows that multiple mechanisms are at play when miRNAs silence genes (reviewed in ref 19); a combination of endonucleolytic cleavage, mRNA decay, and translational suppression is therefore entirely possible. Because miRNA–target interactions take place in RNA processing bodies, called P-bodies (20), where the targeted mRNAs

can be degraded by cellular mRNA decay mechanisms, the reductions in mRNA levels are not unexpected.

A valid question is whether moderate knockdown by off-targeting is enough to induce phenotypes. Fedorov *et al.* (2) used a cell viability assay, which showed that ~30% of the tested siRNAs could induce toxicity by regulating genes other than the intended target in a sequence-specific manner. In an effort to show that the toxic phenotype was a direct result of RNAi, the authors knocked down the human Argonaute 2 (Ago2) protein, which is a component of the RNA-induced silencing complex (RISC) and the effector of siRNA-directed cleavage, and observed reduced cell-death ratios. Interestingly, knockdown of human Ago1 (another component of RISC), which does not have cleavage capability, gave similar results. Meister *et al.* (21) showed that Ago2 is the cleaving component of

RNAi; however, the cleavage mechanism is not necessarily the dominating one here, as many of the silenced transcripts have only limited complementarity to the siRNAs, a hallmark of miRNAs that mediate translational inhibition. Bagga *et al.* (22) have shown that miRNAs can target mRNAs for degradation, but by a mechanism independent of Ago2-mediated cleavage, presumably by sequestering the mRNAs in P-bodies. It is important to note that lowering the concentration of siRNAs <1 nM removed the toxic phenotype; this result emphasizes the importance of using highly potent siRNAs at the lowest possible concentration to avoid unwanted toxicities.

siRNAs are clearly not as specific as many had hoped a few years ago. In retrospect, it may have been naïve to expect that a biological pathway would require full-sequence complementarity, as even the most critical mechanisms are usually tolerant to some aberrations. Chemistry has provided a solution to circumvent off-targeting. A collaborative study between Dharmacon Research (Lafayette, CO) and Rosetta Inpharmatics (Seattle, WA) demonstrated that methyl-groups added to the 2' position of the ribosyl ring of the second base of the siRNA significantly reduced off-targeting (3). Importantly, this common RNA oligonucleotide modification does not affect the degree of silencing of the intended target, as could also be expected from previous reports (23). Others have shown that perfect complementarity between the target and nucleotides 2–7/8 of the miRNA is sufficient to effect down-regulation in itself, although additional binding of sequences in the miRNA 3'-end can increase the probability of silencing even when there is not full complementarity to the miRNA 5'-end (24). Interestingly, the 2'-O-methyl modification was much more effective at reducing the levels of off-targeting when it was added to position 2 than when it was added to position 1 of the siRNA.

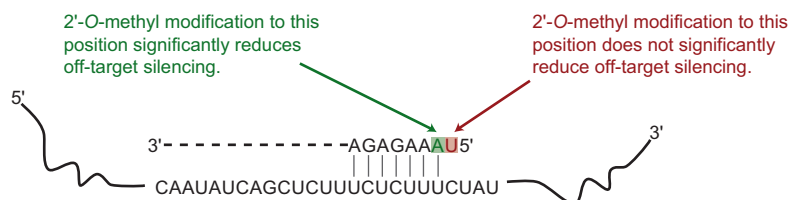


Figure 2. This schematic shows how 7 nucleotides of complementarity (bases 2–8) at the 5'-end of an siRNA may be enough to target a sequence in the 3' UTR of an mRNA. A 2'-O-methyl added to position 2 of the siRNA significantly reduces off-target effects.

These three papers in *RNA* are the most recent addition to the body of evidence showing that the use of siRNAs is not foolproof. siRNAs may induce sequence-specific off-target effects as described, but if they contain certain sequence motifs and structural features, nonspecific activation of the innate immune system, including induction of interferon pathway proteins, may be the result (25, 26). Judicious chemical modifications of the RNA backbone provide promising remedies for both of these problems and offer a research opportunity for chemists, as even better modifications are likely to exist. It should be noted that Dharmascan Research and Rosetta Inpharmatics have opted for the usual selection of various immortalized cancer cell lines. Fedorov *et al.* (2) remark that there were differences in sensitivity to siRNA off-targeting between the cervical, prostate, and breast cancer cell lines that they used; normal diploid cells with all native responses intact may show even more pronounced off-targeting and innate immune responses.

Since the relatively short seed region is definitely the most important determinant for off-target regulation, siRNAs with different seeds are unlikely to produce a similar expression signature. Consequently, researchers who observe the same phenotypic response from two or more siRNAs targeting the same message can still be relatively confident that the phenotype is not a consequence of off-target silencing. These studies clearly enforce a much debated issue about what types of controls should be used in siRNA experiments to prove that an observed phenotype is specific to the knockdown of the intended target.

REFERENCES

1. Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi, *Nat. Biotechnol.* **21**, 635–637.

2. Fedorov, Y., Anderson, E. M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leake, D., Marshall, W. S., and Khvorova, A. (2006) Off-target effects by siRNA can induce toxic phenotype, *RNA*, published online May 8, <http://dx.doi.org/10.1261/ma.28106>.
3. Jackson, A. L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J. M., Lim, L., Karpilow, J., Nichols, K., Marshall, W., Khvorova, A., and Linsley, P. S. (2006) Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing, *RNA*, published online May 8, <http://dx.doi.org/10.1261/ma.30706>.
4. Jackson, A. L., Burchard, J., Schelter, J., Chau, B. N., Cleary, M., Lim, L., and Linsley, P. S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity, *RNA*, published online May 8, <http://dx.doi.org/10.1261/ma.25706>.
5. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature* **411**, 494–498.
6. Boutla, A., Delidakis, C., Livadaras, I., Tzagris, M., and Tabler, M. (2001) Short 5'-phosphorylated double-stranded RNAs induce RNA interference in *Drosophila*, *Curr. Biol.* **11**, 1776–1780.
7. Harborth, J., Elbashir, S. M., Vandenberg, K., Manianga, H., Scaringe, S. A., Weber, K., and Tuschl, T. (2003) Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing, *Antisense Nucleic Acid Drug Dev.* **13**, 83–105.
8. Saxena, S., Jonsson, Z. O., and Dutta, A. (2003) Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells, *J. Biol. Chem.* **278**, 44312–44319.
9. Chi, J. T., Chang, H. Y., Wang, N. N., Chang, D. S., Dunphy, N., and Brown, P. O. (2003) Genomewide view of gene silencing by small interfering RNAs, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6343–6346.
10. Semizarov, D., Frost, L., Sarthy, A., Kroeger, P., Halbert, D. N., and Fesik, S. W. (2003) Specificity of short interfering RNA determined through gene expression signatures, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6347–6352.
11. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H., and Williams, B. R. (2003) Activation of the interferon system by short-interfering RNAs, *Nat. Cell Biol.* **5**, 834–839.
12. Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A. L., and Iggo, R. (2003) Induction of an interferon response by RNAi vectors in mammalian cells, *Nat. Genet.* **34**, 263–264.
13. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell* **120**, 15–20.
14. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003) Prediction of mammalian microRNA targets, *Cell* **115**, 787–798.
15. Birmingham, A., Anderson, E. M., Reynolds, A., Ilesly-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W. S., and Khvorova, A. (2006) 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets, *Nat. Methods* **3**, 199–204.
16. Lim, L. P., Lau, N. C., Garrett-Engle, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs, *Nature* **433**, 769–773.
17. Lin, X., Ruan, X., Anderson, M. G., McDowell, J. A., Kroeger, P. E., Fesik, S. W., and Shen, Y. (2005) siRNA-mediated off-target gene silencing triggered by a 7 nt complementation, *Nucleic Acids Res.* **33**, 4527–4535.
18. Xie, X., Lu, J., Kulbokas, E. J., Golub, T. R., Mootha, V., Lindblad-Toh, K., Lander, E. S., and Kellis, M. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals, *Nature* **434**, 338–345.
19. Roush, S., and Slack, F. (2006) Micromanagement: microRNAs stabilize mRNAs, *ACS Chem. Biol.* **1**, 132–134.
20. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., and Parker, R. (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies, *Nat. Cell Biol.* **7**, 719–723.
21. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs, *Mol. Cell* **15**, 185–197.
22. Bagga, S., Bracht, J., Hunter, S., Massier, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation, *Cell* **122**, 553–563.
23. Amarzguoui, M., Holen, T., Babaie, E., and Prydz, H. (2003) Tolerance for mutations and chemical modifications in a siRNA, *Nucleic Acids Res.* **31**, 589–595.
24. Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005) Principles of microRNA-target recognition, *PLoS Biol.* **3**, e85.
25. Homung, V., Guenther-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., Endres, S., and Hartmann, G. (2005) Sequence-specific potent induction of IFN- α by short interfering RNA in plasmacytoid dendritic cells through TLR7, *Nat. Med.* **11**, 263–270.
26. Judge, A. D., Sood, V., Shaw, J. R., Fang, D., McClintock, K., and MacLachlan, I. (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA, *Nat. Biotechnol.* **23**, 457–462.

Drug Discovery: Here Comes the Worm

Laurent Ségalat*

CNRS-UMR 5534, University of Lyon, Lyon, France

In a landmark paper, Roy and colleagues (1) have used the nematode *Caenorhabditis elegans* to identify new compounds of therapeutic interest by screening a large library of small-molecule chemical candidates. They called the compounds nemadipines. Nemadipines show a structural homology with the well-known antihypertension drugs called dihydropyridines. By taking advantage of the powerful *C. elegans* genetic tools, they have established that nemadipines block the activity of *C. elegans* L-type calcium channels. More interestingly, nemadipines also antagonize vertebrate L-type calcium channels. This study is a proof of concept for the use of this invertebrate system in molecule screens and subsequent target identification.

The tiny, free-living *C. elegans* has been a model organism in biology for 40 years and can now be found in ~1000 laboratories worldwide. *C. elegans* is harmless, ~1 mm long, composed of ~1000 cells, and has a life cycle of only ~3 d. It can be grown in liquid media as well as on agar plates. Although quite a primitive organism, this animal is endowed with many of the basic organs, muscles, and systems common to most members of the animal kingdom (neurons, muscles, intestine, epidermis, and a detoxification and excretory system). During the first few decades of its laboratory career, *C. elegans* was considered primarily as a model organism for tackling developmental biology issues, to the delight of developmental biologists who appreciated its many unique qualities. In recent years, *C. elegans* has been used for two additional

purposes. First, it has become widely used as a model for human diseases, because it provides a cheap and fast alternative to traditional mouse and rat models. Second, it is used in drug screens to identify new compounds of potential medical interest.

Although 5 years ago the idea that this tiny animal could provide a useful means of drug identification was almost heretical, it has recently gained momentum. Drug screens on *C. elegans* can be performed in two ways. First, drugs can be searched that will correct a phenotype that has been induced in *C. elegans* as a result of a mutation or by transgenesis (2, 3). In such cases, *C. elegans* is an alternative to *in vitro* and cell-based screening methods preceding preclinical trials. Because such methods do not exist for numerous genetic diseases, *C. elegans* is of particular interest. In a slightly different perspective, drug screens can also be performed on healthy *C. elegans* nematodes to assay the effects of molecules under development and/or to identify targets for such molecules. The study by Roy and colleagues falls into the latter category (1).

A long-standing bottleneck in drug development has been the identification of bioactive molecules and their targets. *C. elegans* can be added to the existing arsenal, but like any system of this kind, it has pros and cons (4). Supporters appreciate that *C. elegans* is compatible with high-throughput requirements because it can be grown in multiwell plates. Proponents also welcome the fact that, although *C. elegans* is quite a simple animal, it has a level of complexity and regulatory processes higher than that of

ABSTRACT Identification of bioactive molecules and their targets impedes the process of drug development. In a recent paper, a genetically tractable organism, the *Caenorhabditis elegans* worm, is shown to be a viable screening system in which the drug target and the pathway it activates can be readily identified.

*To whom correspondence should be addressed.
E-mail: segalat@cgmcc.univ-lyon1.fr.1662006

Published online June 16, 2006
10.1021/cb600221m CCC: \$33.50
© 2006 by American Chemical Society

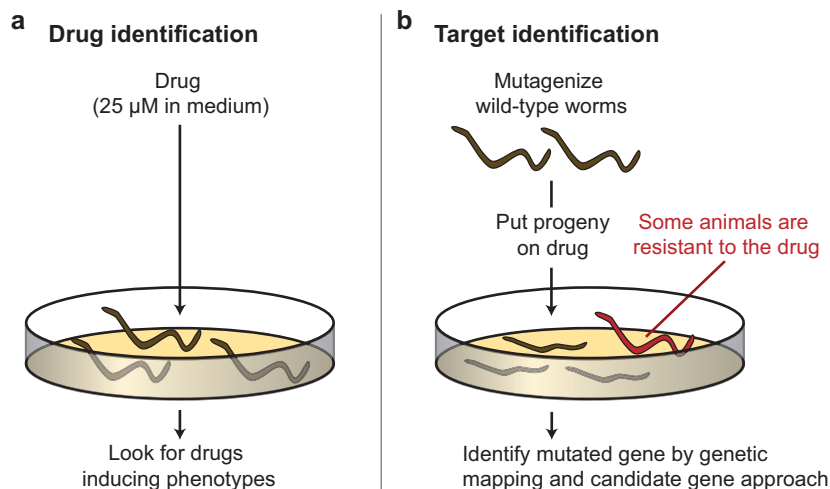


Figure 1. Drug and target identification using the nematode *C. elegans* is performed in two steps. **a)** Bioactive drugs are identified because they induce visible phenotypes (lethality, slow growth, or morphological defects) on wild-type *C. elegans*. **b)** Target identification is obtained by isolating mutants resistant to the drug and identifying the genes mutated.

cells in culture. Therefore, it may allow the exploration of pharmacological domains that could not be investigated with *in vitro* assays. This is particularly true for drugs acting on neurons and muscles: bacteria and yeast, widely used for drug identification, do not have any neurons and muscles and cultured neurons and myotubes have a physiology far different from that of their natural counterparts.

C. elegans also has some cons. First, it is surrounded by a thick cuticle that acts as a barrier to many molecules. In a compound screen, a significant fraction of the molecules tested will show no effects, because they will just not make it into the animals. Second, another fraction of possible hits may be missed: although they make it into the animal, these compounds will not recognize the nematode orthologous target. Protein divergence between nematodes and vertebrates is high. Third, measuring the drug concentration inside the animals is quite challenging: tests are mostly qualitative. Last, but not least, as far as high-throughput screens are concerned, there are not that many phenotypes to look at in worms.

Despite these limitations, among the 14,000 small membrane-permeable molecules screened by Kwok *et al.* (1), 308 (2%) induced a clearly visible phenotype on *C. elegans* (lethality, slow growth, paralysis, and abnormal morphology). Although comparing the hit rates obtained with very different systems is difficult, 2% seems reasonable for a wide-spectrum screen of structurally diverse molecules. Not surprisingly, structurally similar molecules were found to have similar effects on the worms.

Next comes the question of target identification. Remember that *C. elegans* has been the workhorse of hard-core geneticists who have developed elaborate genetic tools. Once a drug produces a specific phenotype in *C. elegans*, such as slow growth and morphological defects in the case of nemadipine, genetics can be used to identify the drug target as well as the pathway it activates. This strategy is similar to the one used to identify the bacterial targets of antibiotics. After a random mutagenesis on thousands of animals, in which virtually all the genes are mutated at least once, one looks for mutant animals

that have become insensitive (resistant) to the drug (Figure 1). In most cases, the mutations will affect either the drug receptor or downstream elements mediating its action. In the case of Nemadipine-A, after a preliminary mapping of 5 independent mutations on the *C. elegans* genetic map, Kwok *et al.* (1) found that 5/5 mutants were located near the gene *egl-19*, which encoded an L-type calcium channel and was one of the candidate targets. This finding was confirmed when the mutants were shown to carry alterations in the *egl-19* DNA sequence.

In conclusion, the worm system is a new addition in the toolbox of therapeutic-drug identification. This system has its limitations and will obviously miss potentially interesting drugs. However, we should regard the glass as half-full rather than half-empty. Despite its drawbacks, *C. elegans* will catch a few drugs that would not have been identified otherwise. This is a good enough reason to justify its use.

REFERENCES

1. Kwok, T. C., Ricker, N., Fraser, R., Chan, A. W., Bums, A., Stanley, M. E. F., McCourt, P., Cutler, S. R., and Roy, P. J. (2006) A small-molecule screen in *C. elegans* yields a new calcium channel antagonist, *Nature* 441, 91–95.
2. Parker, J. A., Arango, M., Abderrahmane, S., Lambert, E., Tourette, C., Catoire, H., and Neri, C. (2005) Resveratrol rescues mutant polyglutamine cytotoxicity in nematode and mammalian neurons, *Nat. Genet.* 37, 349–350.
3. Gaud, A., Simon, J. M., Witzel, T., Carre-Pierrat, M., Wermuth, C. G., and Ségalat, L. (2004) Prednisone reduces muscle degeneration in dystrophin-deficient *Caenorhabditis elegans*, *Neuromuscular Disord.* 14, 365–370.
4. Ségalat, L. (2004) Reply to De Luca, *Neuromuscular Disord.* 14, 698.

Tumor Selectivity of Hsp90 Inhibitors: The Explanation Remains Elusive

Gabriela Chiosis^{†*} and Len Neckers[‡]

[†]Department of Medicine and Program in Molecular Pharmacology and Chemistry, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, and [‡]Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland 20892-1107

The molecular chaperone heat-shock protein 90 (Hsp90) plays an important role in maintaining the functional stability and viability of cells under a transforming pressure. It allows cancer cells to tolerate the deregulation of components of the signaling pathway that occur during cellular transformation. Hsp90 protects the cell by interacting with and stabilizing several client substrates, including kinases, hormone receptors, and transcription factors, which are directly involved in driving multistep malignancy, and also with mutated oncogenic proteins that drive the transformed phenotype (see reviews 1–5). Association of Hsp90 with these client proteins maintains their ability to function in the deregulated state and appears to be essential for their transforming, aberrant activity. Enhanced Hsp90 affinity for mutated or functionally deregulated client proteins has been observed, and several examples of this behavior have been documented. Historically, *v-src* was the first oncogene shown to display unusually stable interaction with the chaperone (6, 7). In fact, the first Hsp90 inhibitors, geldanamycin (GM) and radicicol (RD), were identified in a screen for compounds that could reverse the phenotype of cells transformed by *v-src* (8). In contrast, the non-oncogenic *c-src* requires only limited assistance from the Hsp90 machinery for its maturation and cellular function. Examples may be extended to other transformed phenotypes, and in this regard, almost every protein

involved in cell-specific oncogenic processes has been shown to be regulated by Hsp90 (2, 5). In addition to its role as a chaperone of oncoproteins, Hsp90 is involved in protein folding processes that occur in normal cells. Cells are faced with the task of folding thousands of different polypeptides into a wide range of conformations, a process requiring the concerted action of multiple molecular chaperones. From yeast to mammals, Hsp90 functions together with Hsp70 in the folding of a diverse set of proteins, including transcription factors, regulatory kinases, and numerous other proteins that appear to lack common structural or functional features (9). Even under nonstressed conditions, Hsp90 accounts for as much as 1–2% of total cellular protein. Conceiving that Hsp90 is a viable target in cancer therapy is hard because constitutive Hsp90 genetic knockout in eukaryotes is lethal (3, 5). That it is such a target has become apparent only after the discovery of pharmacological agents that selectively inhibit its function (10). These agents have been useful in probing the biological functions of Hsp90 at the molecular level and in validating it as a novel target for anticancer drugs.

The N-terminal region ATP pocket binders are the first identified inhibitors of Hsp90 activity (Figure 1). Some of these are natural products, such as ansamycins (GM, 17-allylamino-17-demethoxygeldanamycin [17-AAG] and 17-dimethylaminoethylamino-17-demethoxy-geldanamycin

ABSTRACT Two recent papers attempt to solve both the tumor selectivity and the *in vivo* tumor accumulation profiles seen with some Hsp90 inhibitors. They spotlight the higher affinity of ansamycins' hydroquinone over the quinone form for Hsp90 and further discuss its possible contribution to ansamycins' tumor selectivity.

*To whom correspondence should be addressed.
E-mail: chiosisg@mskcc.org.

Published online June 16, 2006
10.1021/cb600224w CCC: \$33.50
© 2006 by American Chemical Society

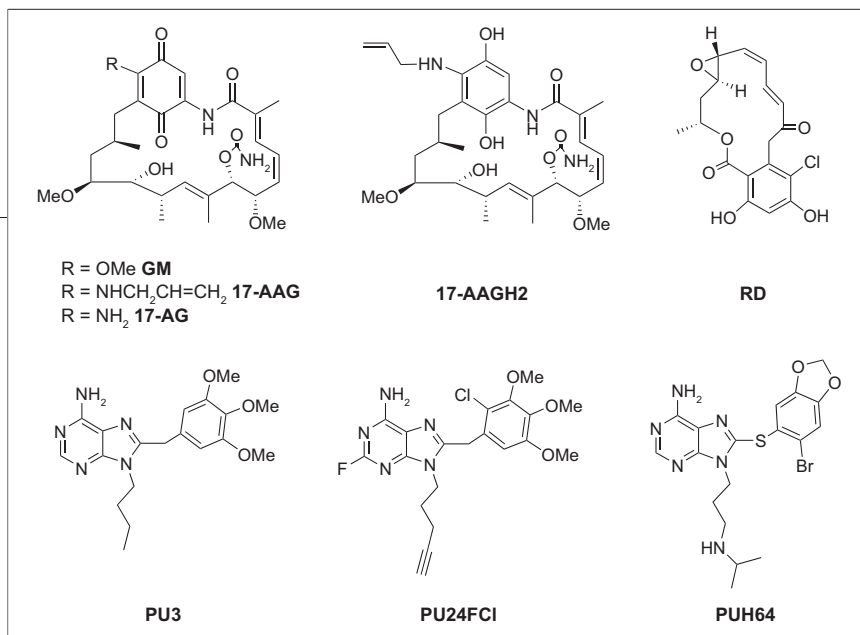


Figure 1. Representative Hsp90 inhibitors. Ansamycins (GM, 17-AAG, 17-AG), RD, and the PU-scaffold derivatives (PU3, PU24FCl, and PUH64).

[17-DMAG] (11–13) and RDs (RD, RD oxime-derivatives, (14); cycloproparadicol, (15); and pochonin D, (16)). Others are synthetic small molecules discovered either by design or by high-throughput screening. These can be subclassified into purine-scaffold derivatives (PU-class), discovered by our laboratory (17, 18), and their further re-scaffolding products, such as triazolopyrimidines, pyrazolopyrimidines, and pyrrolopyrimidines (19), and other scaffolds, such as pyrazole derivatives (20). Compounds that interfere with chaperone cycling by binding to important domains in the C-terminus have also been identified (21, 22). Recently, a peptidomimetic modeled on the binding interface between the molecular chaperone Hsp90 and the antiapoptotic and mitotic regulator survivin was reported (23). This peptidomimetic, termed shepherdin, mimics the survivin sequence I74–L87, the minimal peptide span that retains Hsp90 inhibitory activity. Shepherdin was shown to make extensive contacts with the ATP pocket of Hsp90, resulting in destabilization of chaperone client proteins.

Probing Hsp90 function with these agents in cellular and animal models of cancer has led to some surprising yet rewarding findings. First, several of these inhibitor classes have shown selective binding to Hsp90 in tumor cells (23–28). Second, cancer cells have proven to be significantly more sensitive to Hsp90 inhibition than are nontransformed cells (23–30).

Third, Hsp90 inhibitors at nontoxic doses have demonstrated anticancer activity in multiple animal models (23, 27, 28, 31, 32). Moreover, drug accumulation in tumors, coupled with rapid clearance from blood and normal tissue, has been observed for some Hsp90 inhibitor classes in these animal models (26, 27, 33–35).

Although no comprehensive explanation has yet been presented for the remarkable selectivity of Hsp90 inhibitors toward cancer cells, many hypotheses have emerged. Both *in vitro* and *in vivo* approaches have been used to implicate co-chaperones, tumor-specific drug modification, oncogene addiction, tumor-specific post-translational modification (PTM), and so forth, as the basis of this selectivity.

In Vitro Explanations Using Biochemical Approaches. Several groups have analyzed the Hsp90 inhibitor tumor selectivity at a biochemical level. First, Kamal *et al.* (24) have measured a 100-fold difference for 17-AAG (Figure 1) in affinity between transformed and normal-cell Hsp90. Further, studies with the PU-scaffold Hsp90 inhibitors (Figure 1) have reported similar selectivity with several members of this class of molecules (25–28). Recently, Altieri *et al.* (23) have demonstrated that immobilized shepherdin pulled down Hsp90 only from transformed cells, not from normal ones.

Several attempts to shed light on the higher affinity of these agents for tumor Hsp90 have focused on ansamycins.

Initially, the affinity of these drugs for the chaperone in solution was determined by several biochemical methods to be $\sim 1 \mu\text{M}$. This is contrary to their cellular potency seen at low-nanomolar concentrations. Multiple explanations of this apparent discrepancy have been proposed. In a recent paper, Maroney *et al.* (36) suggested that the dihydroquinone form of 17-AAG, 17-AAGH2 (Figure 1), which may form in cells, has a better affinity for Hsp90 than does the quinone 17-AAG. First, an assay that monitors protein unfolding as a function of temperature by using the environmentally sensitive dye bis-1-anilino-8-naphthalene sulfonate was employed to test Hsp90 thermal stability in the presence and absence of GM, 17-AAG, RD, and nucleotides, with or without added reducing agents. For both GM and 17-AAG, a 40-fold increase in protein stabilization was seen in the presence of reducing agents such as DTT and tris-(2-carboxyethyl)phosphine hydrochloride (TCEP). Further, a filter binding assay that measures drug binding, under equilibrium conditions (48 h), of [³H]-17-AAG to a truncated Hsp90 α containing the N-terminal region determined K_d of 1.1 μM and 2.4 nM in the absence or presence of TCEP, respectively. The same assay was used to determine an extended k_{off} for 17-AAGH2 (Figure 1) as compared to 17-AAG ($\gg 4.5$ h vs several minutes). In a different approach, using a fluorescence polarization assay, Llauger *et al.* (37) reported that two fluorescently labeled GMs, GM-BODIPY and GM-FITC, bound tightly to Hsp90 α with K_d of 33.8 ± 1.2 and 23.3 ± 0.9 nM, respectively. These same authors further demonstrated that binding was affected by DTT, a higher affinity with K_d of 6.6 ± 1.3 nM being favored by more DTT in the assay buffer and longer incubation times (38, 39). These authors also reported that equilibration to the high-affinity state was slow at low-DTT concentrations but accelerated as the DTT content was increased. Gooljarsingh *et al.* (40) repro-

duced these findings in their recent *PNAS* publication. They reported that GM-BODIPY induced a time-dependent conformational change in both Hsp90 α and β isoforms leading to a high-affinity state. A similar profile was observed for Hsp90 assayed in cell lysates obtained from ovarian epithelial cancer (transformed) and human umbilical vein endothelial (HUVEC, normal) cells. These data are in agreement with previous reports by Llauger *et al.* (25, 26) on Hsp90 from breast-cancer cells. The dissociation rate of this tightly bound GM was measured to be very slow, with a $t_{1/2}$ of 4.5 h. All high-affinity binding constants for ansamycins were observed in the presence of 2 mM DTT in the assay buffer.

These findings indicated that reducing agents affected affinity, so we further analyzed buffer conditions for several other assays used previously to determine the affinity of ansamycins for Hsp90. Roe *et al.* (41) used isothermal titration calorimetry to calculate a dissociation constant of 1.22 μM for GM binding to intact yeast Hsp90 and 0.78 μM for GM binding to the Hsp90 N-terminal domain. These values were obtained in 20 mM Tris-HCl, pH 7.5, and 1 mM EDTA. Chiosis *et al.* (17, 42) used immobilized GM to obtain an apparent relative affinity of 17-AAG for Hsp90 α of 1 μM . This assay used buffer conditions comparable to Roe *et al.* (50 mM Tris-HCl, pH 7.4, and 1 mM EDTA with added 1% NP40 to reduce background interference from the solid support). Carreras *et al.* (43) reported that in 10 mM Tris-HCl and 5 mM MgCl_2 , pH 7.0, the binding of [^3H]-17-AAG to Hsp90 reached equilibrium in a few minutes with a K_d of $0.4 \pm 0.1 \mu\text{M}$. Using immobilized biotinylated GM, Le Brazidec *et al.* (44) reported a relative affinity of 17-AAG for Hsp90 in PBS of 800 nM.

Collectively, these data support the hypothesis that reducing agents can influence the affinity of ansamycins for Hsp90, suggesting that the hydroquinone is a better Hsp90 binder than the quinone itself. How-

ever, they do not support a role for this phenomenon in explaining the higher tumor affinity of Hsp90 inhibitors, because neither PU-scaffold derivatives nor shepherdin are DTT-sensitive. In addition, these findings do not preclude an additional effect of DTT on the structural conformation of Hsp90. Further, the concept of an “encounter complex” between GM and Hsp90 with a relatively weak K_i of 450 nM, which then equilibrates to a tight complex with K_i of 10 nM (*e.g.*, “time-dependent” binding of ansamycins), presented by Gooljarsingh *et al.* (40) can be explained by either an induced fit model or the active ansamycin isomerization model described by Lee *et al.* (45). In any case, such a binding schema is ansamycin-specific.

Another biochemical explanation for tumor selectivity of Hsp90 inhibitors comes from work by Kamal *et al.* (24), who reported that Hsp90 from tumor cells is present entirely in multichaperone complexes with high ATPase activity and also high affinity for ligands, whereas in normal tissues, Hsp90 exists in a latent uncomplexed state. Maroney *et al.* (36) confirmed this finding by demonstrating that the amount of Hsp90 complexed to co-chaperones is higher in tumor cells than in resting ones. Gooljarsingh *et al.* (40), however, contradicted the finding and argued that Hsp90 co-chaperones such as Hsp70, Hsp40, Hop, and p23 do not alter the GM-Hsp90 binding profile, even though they are sufficient to refold a denatured Hsp90 client protein. Interpretation of such *in vitro* data has its limitations, however, because these co-chaperones are not bound to the same conformation of Hsp90. Further, there are many more co-chaperones whose interaction with Hsp90 could significantly impact (at least theoretically) access to or the shape of the nucleotide pocket (or surrounding regions).

Putting the Cell to Work. Another finding for pharmacological Hsp90 inhibition is that cancer cells are significantly more sensitive

to these drugs than are nontransformed cells. Such selectivity in inhibiting the growth of cancer cells has been reported for ansamycins, RD derivatives, and PU-scaffold derivatives (23–30). Such sensitivity was first believed to be class-specific and due to an intracellular reduction of 17-AAG. Kelland *et al.* (46) observed that ansamycins were substrates for purified human NAD(P)H:quinone oxidoreductase 1 (NQO1; DT-diaphorase, EC 1.6.99.2). This flavoenzyme can use either NADH or NADPH as reducing cofactors to catalyze the direct two-electron reduction of quinones to hydroquinones. Elevated cellular activity of this enzyme sensitized cells to 17-AAG, but surprisingly, not to GM or 17-amino, 17-demethoxygeldanamycin (17-AG). Although 17-AAG was a reasonable substrate for human DT-diaphorase, it was not an appreciably better substrate than GM or 17-AG. Ross *et al.* (47) confirmed the 17-AAG findings and further demonstrated that the metabolism of 17-AAG by recombinant human NQO1 led to the appearance of 17-AAGH2. The formation of 17-AAGH2 was NQO1-dependent and could be inhibited by the addition of a mechanism-based (suicide) inhibitor of NQO1. Maroney *et al.* (36) further demonstrated that 17-AAGH2 does not spontaneously form in aqueous media. They do, however, find that cells possess enough reductive potential to conduct this transformation, and a 68% transformation of 17-AAG to 17-AAGH2 was observed in MCF7 breast-cancer cells. However, quiescent HUVEC cells reduced 17-AAG to the same extent. Nonetheless, the inhibitory activity of 17-AAG is 100 nM in MCF7 but $>10 \mu\text{M}$ in quiescent HUVECs; thus, it is unlikely that inhibitor selectivity for tumor Hsp90 is explained by tumor-specific conversion of 17-AAG to 17-AAGH2.

Chiosis *et al.* (48) reported another peculiar behavior of ansamycins in cells. In tissue-culture experiments, when the molar amount of ansamycins was kept constant but the volume of the medium was in-

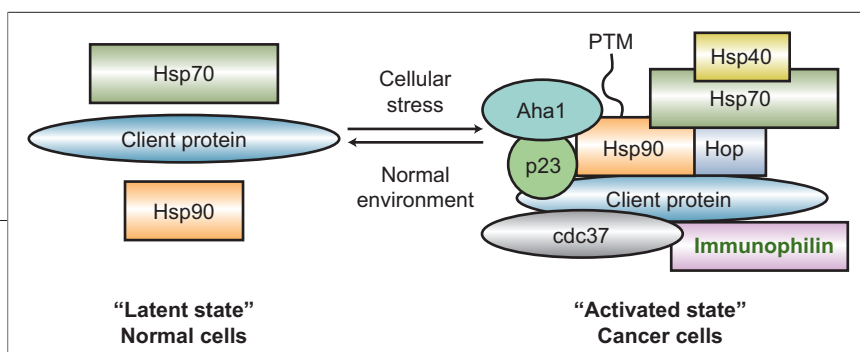


Figure 2. Hsp90 may exist in an equilibrium between an “activated” state prevalent in cancer cells and a “latent” state predominant in normal cells. The activation state of the chaperone may be regulated by co-chaperones and perhaps PTMs. This is a schematic representation of Hsp90 states and does not represent actual individual complexes.

creased, no change was observed in MCF7 cells’ growth-inhibitory potency. In contrast, when such an experiment was repeated with the PU-scaffold derivatives PU3 and PU24FCl (Figure 1), activity diminished with drug concentration. Upon addition to aqueous tissue culture media, ansamycins rapidly accumulated in cells and were mostly depleted from the media, producing higher intracellular concentrations than expected. This behavior was also observed in cultured normal epithelial cells. The nonspecific accumulation of GM and 17-AAG into cells in tissue culture conflicted with their specific tumor accumulation observed *in vivo*. It is thus important to differentiate between nonselective drug accumulation into cells in tissue culture, which may be due to physicochemical peculiarities of a compound in those settings, and selective accumulation into cancer (*vs* normal) cells due to different biological characteristics of these cells. These data warn us that any peculiarities seen in tissue culture may not explain the *in vivo* tumor accumulation of Hsp90 inhibitors.

Whole-Animal Case Study: No Plastic to Blame. Hsp90 inhibitors have anticancer activity in multiple-animal xenograft models at nontoxic doses. Such effects have been observed with ansamycins such as 17-AAG, IPI-504, and 17-DMAG (32–35) and also with RD derivatives (31), PU-scaffold inhibitors (27, 28), and shepherdin (23). In addition, drug accumulation in tumors coupled to rapid clearance from normal tissue has been observed for multiple Hsp90 inhibitor classes in these models. The first agent reported to be retained in tumors was our PU-scaffold derivative PU24FCl (Figure 1) (27). While this agent was rapidly cleared from blood, pharmacologically relevant

concentrations were recorded in MCF7 xenografts at >24 h post-administration. Second-generation, water-soluble PU-class agents such as PUH64 (Figure 1) were similarly shown to be retained in MDA-MB-468 xenografts, suggesting that tumor retention is not influenced by drug lipophilicity (26). Similar results were later reported by Eisman *et al.* (33) for 17-DMAG in MDA-MB-231 breast-cancer xenografts and by Workman *et al.* (34) for 17-AAG in human-ovarian-cancer xenografts. Recently, IPI-504, the reduced form of 17-AAG, was shown to be rapidly cleared from plasma while selectively retained in tumor tissue. These results were reported in a multiple-myeloma xenograft model (35). Further, Serenex has claimed a tumor retention profile for their structurally novel SNX-2112 Hsp90 inhibitor in an HT29 human colon adenocarcinoma model (www.serenex.com). Thus, the tumor-specific accumulation of such diverse chemical classes of Hsp90 inhibitors clearly cannot be entirely explained by the biochemical and cellular hypotheses presented above (which focus only on the peculiarities of ansamycins).

The Answer May Lie in Hsp90 Itself. In tumors, Hsp90 comprises as much as 4–6% of total cellular protein, and this translates into an intracellular concentration in the ~500 μM range. It is thus noteworthy that administration of 50–200 mg kg^{-1} of diverse Hsp90 inhibitors led to only 0.5–5.0% of drug being retained in tumors. Eisman *et al.* (33) reported an ~5 $\mu\text{g mL}^{-1}$ level of 17-DMAG retained in MDA-MB-231 tumors at ~12 h post-administration when drug was injected intravenously at 75 mg kg^{-1} . Banerji *et al.* (34) observed in A2780 and CH1 human-ovarian-cancer xenografts treated with a single dose of 17-AAG

(80 mg kg^{-1} intraperitoneally) ~5 μM of 17-AAG and 17-AG (an active metabolite of 17-AAG, Figure 1) at 24 h. Similar data were reported for the PU-scaffold Hsp90 inhibitors. Administration of 200 mg kg^{-1} PU24FCl to mice bearing MCF7 xenografts led to 5–10 μM tumor drug levels at 24 h (27). PUH64 also accumulated to low-micromolar concentration in tumors at 24 h when administered intraperitoneally at 50–100 mg kg^{-1} to mice bearing MDA-MB-468 xenografts (26). The Serenex derivative SNX-2112 is reported to reach a concentration of 3 μM in HT29 colon-cancer xenografts at 24 h when administered at 100 mg kg^{-1} orally (www.serenex.com). At a concentration of 3–10 μM , these drugs occupy only a small fraction (0.5–2.0%) of total cellular Hsp90 binding sites. However, in all of the examples described above, the doses of Hsp90 inhibitors used resulted in tumor growth inhibition. One may infer from such observations that these drugs target only a relatively low-abundance but high-affinity conformation of Hsp90, likely found in a multichaperone complex with transformation-specific oncoproteins. This may represent the small fraction of client proteins regulating the transformed phenotype. The latent Hsp90 complexes regulating normal misfolding processes and comprising at any time >95% of total cellular Hsp90 may not be effectively inhibited by these drugs at the relatively nontoxic doses used. Such an interpretation leads to the hypothesis that, under normal conditions, Hsp90 interacts with client proteins in a dynamic, low-affinity manner regulated by low-affinity binding and release of ATP and ADP “latent state”. Upon mutation or deregulation characteristic of the cancer phenotype, many of these client proteins may display (and require) unusually stable association with Hsp90-containing chaperone complexes “activated state” (Figure 2). This state also exhibits a high affinity for ATP and ADP or other ligands of this regulatory pocket (*i.e.*, N-terminal Hsp90 inhibitors). The shift in equilibrium

from the latent to the activated state may be dictated by the degree of transformation or amount of “stress” on the system (abundance of mutated and deregulated proteins, a hypoxic and/or low-nutrient environment, etc.).

The mechanism of this enhanced affinity has not yet been entirely elucidated but is likely due to tumor-specific modification of Hsp90 itself and not to any unique tumor-specific metabolism of its inhibitors. Possible factors responsible for Hsp90 modification include its specific interaction with co-chaperones (24, 36) and/or alterations in the post-translational state of Hsp90, co-chaperones, or both (49).

While they confirm previous observations about ansamycins, the two new papers by Maroney *et al.* (36) and Gooljarsingh *et al.* (40) do not extend our understanding of the specific biology of tumor Hsp90, particularly its remarkable sensitivity to pharmacological inhibition. Many questions remain unanswered about the uniqueness of Hsp90 in tumors.

REFERENCES

- Workman, P. (2004) Combinatorial attack on multi-step oncogenesis by inhibiting the Hsp90 molecular chaperone, *Cancer Lett.* 206, 149–157.
- Zhang, H., and Burrows, F. (2004) Targeting multiple signal transduction pathways through inhibition of Hsp90, *J. Mol. Med.* 82, 488–499.
- Whitesell, L., and Lindquist, S. L. (2005) HSP90 and the chaperoning of cancer, *Nat. Rev. Cancer* 5, 761–772.
- Neckers, L., and Neckers, K. (2005) Heat-shock protein 90 inhibitors as novel cancer chemotherapeutics—An update, *Expert Opin. Emerging Drugs* 10, 137–149.
- Chiosis, G. (2006) Targeting chaperones in transformed systems—A focus on Hsp90 and cancer, *Expert Opin. Ther. Targets* 10, 37–50.
- Xu, Y., and Lindquist, S. (1993) Inhibition of heat shock protein hsp90 governs the activity of pp60v-src kinase, *Proc. Natl. Acad. Sci. U.S.A.* 90, 7074–7078.
- Whitesell, L., Mimnaugh, E. G., De Costa, B., Myers, C. E., and Neckers, L. M. (1994) Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: Essential role for stress proteins in oncogenic transformation, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8324–8328.
- Uehara, Y. (2003) Natural product origins of Hsp90 inhibitors, *Curr. Cancer Drug Targets* 3, 325–330.
- Wegele, H., Muller, L., and Buchner, J. (2004) Hsp70 and Hsp90—A relay team for protein folding, *Rev. Physiol. Biochem. Pharmacol.* 151, 1–44.
- Neckers, L. (2006) Using natural product inhibitors to validate Hsp90 as a molecular target in cancer, *Curr. Top. Med. Chem.*, in press.
- Neckers, L., Schulte, T. W., and Mimnaugh, E. (1999) Geldanamycin as a potential anti-cancer agent: Its molecular target and biochemical activity, *Invest. New Drugs* 17, 361–373.
- Schulte, T. W., and Neckers, L. M. (1998) The benzoquinone ansamycin 17-allylamino-17-demethoxygeldanamycin binds to Hsp90 and shares important biologic activities with geldanamycin, *Cancer Chemother. Pharmacol.* 42, 273–279.
- Smith, V., Sausville, E. A., Camalier, R. F., Fiebig, H. H., and Burger, A. M. (2005) Comparison of 17-dimethylaminoethylamino-17-demethoxy-geldanamycin (17-DMAG17-DMAG) and 17-allylamino-17-demethoxygeldanamycin (17-AAG) in vitro: Effects on Hsp90 and client proteins in melanoma models, *Cancer Chemother. Pharmacol.* 56, 126–137.
- Soga, S., Shiotsu, Y., Akinaga, S., and Sharma, S. V. (2003) Development of radicicol analogues, *Curr. Cancer Drug Targets* 3, 359–369.
- Yamamoto, K., Garbaccio, R. M., Stachel, S. J., Solit, D. B., Chiosis, G., Rosen, N., and Danishefsky, S. J. (2003) Target oriented total synthesis as a resource in the discovery of potentially valuable agents in oncology: Cyclopropradicicol, *Angew. Chem.* 42, 1280–1284.
- Moulin, E., Zoete, V., Barluenga, S., Karplus, M., and Wintsinger, N. (2005) Design, synthesis, and biological evaluation of HSP90 inhibitors based on conformational analysis of radicicol and its analogues, *J. Am. Chem. Soc.* 127, 6999–7004.
- Chiosis, G., Timaul, M. N., Lucas, B., Munster, P. N., Zheng, F. F., Sepp-Lorenzino, L., and Rosen, N. (2001) A small molecule designed to bind to the adenine nucleotide pocket of Hsp90 causes Her2 degradation and the growth arrest and differentiation of breast cancer cells, *Chem. Biol.* 8, 289–299.
- Chiosis, G., and Rosen, N. (2002) Small molecule composition for binding to Hsp90, *PCT Int. Appl. WO-20020236075*.
- Kasibhatla, S. R., Boehm, M. F., Hong, K., Biamonte, M. A., Shi, J., Le Brazidec, J.-Y., Zhang, L., and Hurst, D. (2005) Novel heterocyclic compounds as Hsp90 inhibitors, *PCT Int. Appl. WO-2005028434*.
- Cheung, K. M., Matthews, T. P., James, K., Rowlands, M. G., Proxall, K. J., Sharp, S. Y., Maloney, A., Roe, S. M., Prodromou, C., Pearl, L. H., Aheme, G. W., McDonald, E., and Workman, P. (2005) The identification, synthesis, protein crystal structure and in vitro biochemical evaluation of a new 3,4-dianilpyrazole class of Hsp90 inhibitors, *Bioorg. Med. Chem. Lett.* 15, 3338–3343.
- Marcu, M. G., Schulte, T. W., and Neckers, L. (2000) Novobiocin and related coumarins and depletion of heat shock protein 90-dependent signaling proteins, *J. Natl. Cancer Inst.* 92, 242–248.
- Yu, X. M., Shen, G., Neckers, L., Blake, H., Holzbeierlein, J., Cronk, B., and Blagg, B. S. (2005) Hsp90 inhibitors identified from a library of novobiocin analogues, *J. Am. Chem. Soc.* 127, 12778–12779.
- Plescia, J., Salz, W., Xia, F., Pennati, M., Zaffaroni, N., Daidone, M. G., Meli, M., Dohi, T., Fortugno, P., Nefedova, Y., Gabrilovich, D. I., Colombo, G., and Altieri, D. C. (2005) Rational design of shepherdin, a novel anticancer agent, *Cancer Cell* 5, 457–468.
- Kamal, A., Thao, L., Sensintaffar, J., Zhang, L., Boehm, M. F., Fritz, L. C., and Burrows, F. J. (2003) A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors, *Nature* 425, 407–410.
- Llauger, L., He, H., Kim, J., Aguirre, J., Rosen, N., Peters, U., Davies, P., and Chiosis, G. (2005) 8-Arylsulfanyl and 8-arylsulfoxyl adenine derivatives as inhibitors of the heat shock protein 90, *J. Med. Chem.* 48, 2892–2905.
- He, H., Zatorska, D., Kim, J., Aguirre, J., Llauger, L., She, Y., Wu, N., Immormino, R. M., Gewirth, D. T., and Chiosis, G. (2006) Identification of potent water-soluble purine-scaffold inhibitors of the heat shock protein 90, *J. Med. Chem.* 49, 381–390.
- Vilenchik, M., Solit, D., Basso, M., Huezio, H., Lucas, B., Huazhong, H., Rosen, N., Spampinato, C., Modrich, P., and Chiosis, G. (2004) Targeting wide-range oncogenic transformation via PU24FC1, a specific inhibitor of tumor Hsp90, *Chem. Biol.* 11, 787–797.
- Biamonte, M. A., Shi, J., Hong, K., Hurst, D. C., Zhang, L., Fan, J., Busch, D. J., Karjian, P. L., Maldonado, A. A., Sensintaffar, J. L., Yang, Y.-C., Kamal, A., Lough, R. E., Lundgren, K., Burrows, F. J., Timony, G. A., Boehm, M. F., and Kasibhatla, S. R. (2006) Orally active purine-based inhibitors of the heat shock protein 90, *J. Med. Chem.* 49, 817–828.
- Soga, S., Shiotsu, Y., Akinaga, S., and Sharma, S. V. (2003) Development of radicicol analogues, *Curr. Cancer Drug Targets* 3, 359–369.
- Whitesell, L., Shiffrin, S. D., Schwab, G., and Neckers, L. M. (1992) Benzoquinonoid ansamycins possess selective tumoricidal activity unrelated to src kinase inhibition, *Cancer Res.* 52, 1721–1728.
- Soga, S., Shiotsu, Y., Akasaka, K., Narumi, H., Agatsuma, T., Ikuina, Y., Murakata, C., Tamaoki, T., Schulte, T. W., Neckers, L. M., and Akinaga, S. (1999) KF25706, a novel oxime derivative of radicicol exhibits in vivo antitumor activity via selective depletion of Hsp90 binding signaling molecules, *Cancer Res.* 59, 2931–2938.
- Solit, D. B., Zheng, F. F., Drobnyak, M., Munster, P. N., Higgins, B., Verbel, D., Heller, G., Tong, W., Cordon-Cardo, C., Agus, D. B., Scher, H. I., and Rosen, N. (2002) 17-Allylamino-17-demethoxygeldanamycin induces the degradation of androgen receptor and HER-2/neu and inhibits the growth of prostate cancer xenografts, *Clin. Cancer Res.* 5, 986–993.
- Eiseman, J. L., Lan, J., Lagattuta, T. F., Hamburger, D. R., Joseph, E., Covey, J. M., and Egorin, M. J. (2005) Pharmacokinetics and pharmacodynamics of 17-demethoxy 17-[[[2-dimethylamino]ethyl]amino]geldanamycin (17-DMAG, NSC 707545) in C.B-17 SCID mice bearing MDA-MB-231 human breast cancer xenografts, *Cancer Chemother. Pharmacol.* 55, 21–32.

34. Banerji, U., Walton, M., Raynaud, F., Grimshaw, R., Kelland, L., Valenti, M., Judson, I., and Workman, P. (2005) Pharmacokinetic–pharmacodynamic relationships for the heat shock protein 90 molecular chaperone inhibitor 17-allylamino, 17-demethoxygeldanamycin in human ovarian cancer xenograft models, *Clin. Cancer Res.* **11**, 7023–7032.
35. Sydor, J. R., Pien, C. S., Zhang, Y., Ali, J., Dembski, M. S., Ge, J., Grenier, L., Hudak, J., Normant, E., Pak, R., Patterson, J., Pink, M., Sang, J., Woodward, C., Mitsiades, C. S., Anderson, K. C., Grayzel, D. S., Wright, J., Tong, J. K., Adams, J., Palombella, V. J., and Barret, J. A. (2005) Anti-tumor activity of a novel, water soluble Hsp90 inhibitor IPI-504 in multiple myeloma, *Proc. Am. Assoc. Cancer Res.* **46**, Abstract No. 6160.
36. Maroney, A. C., Marugan, J. J., Mezzasalma, T. M., Bamakov, A. N., Garrabrant, T. A., Weaner, L. E., Jones, W. J., Bamakova, L. A., Koblisch, H. K., Todd, M. J., Masucci, J. A., Deckman, I. C., Galemno, Jr., R. A., and Johnson, D. L. (2006) Dihydroquinone ansamycins: Toward resolving the conflict between low in vitro affinity and high cellular potency of geldanamycin derivatives, *Biochemistry* **45**, 5678–5685.
37. Llauger, L., Felts, S., Huezio, H., Rosen, N., and Chiosis, G. (2003) Synthesis of novel fluorescent probes for the molecular chaperone Hsp90, *Bioorg. Med. Chem. Lett.* **13**, 3975–3978.
38. Llauger, L., He, H., Kim, J., Rosen, N., and Chiosis, G. (2003) Development of a fluorescence polarization assay for the molecular chaperone Hsp90, Proceedings NCI-EORTC-AACR Molecular Targets and Cancer Therapeutics Meeting, Boston, MA, Nov 17–21, Abstract No. B154.
39. Kim, J., Felts, S., He, H., Llauger, L., Huezio, H., Rosen, N., and Chiosis, G. (2004) Development of a fluorescence polarization assay for the molecular chaperone Hsp90, *J. Biomol. Screening* **9**, 375–381.
40. Gooljarsingh, L. T., Fernandes, C., Yan, K., Zhang, H., Grooms, M., Johanson, K., Sinnamon, R. H., Kirkpatrick, R. B., Kerrigan, J., Lewis, T., Arnone, M., King, A. J., Lai, Z., Copeland, R. A., and Tummino, P. J. (2006) A biochemical rationale for the anticancer effects of Hsp90 inhibitors: Slow, tight binding inhibition by geldanamycin and its analogues, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7625–7630.
41. Roe, S. M., Prodromou, C., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1999) Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin, *J. Med. Chem.* **42**, 260–266.
42. Chiosis, G., Rosen, N., and Sepp-Lorenzino, L. (2001) LY294002-geldanamycin heterodimers as selective inhibitors of the PI3K and PI3K-related family, *Bioorg. Med. Chem. Lett.* **11**, 909–913.
43. Carreras, C. W., Schirmer, A., Zhong, Z., and Santi, D. V. (2003) Filter binding assay for the geldanamycin-heat shock protein 90 interaction, *Anal. Biochem.* **317**, 40–46.
44. Le Brazidec, J. Y., Kamal, A., Busch, D., Thao, L., Zhang, L., Timony, G., Grecko, R., Trent, K., Lough, R., Salazar, T., Khan, S., Burrows, F., and Boehm, M. F. (2004) Synthesis and biological evaluation of a new class of geldanamycin derivatives as potent inhibitors of Hsp90, *J. Med. Chem.* **47**, 3865–3873.
45. Lee, Y. S., Marcu, M. G., and Neckers, L. (2004) Quantum chemical calculations and mutational analysis suggest heat shock protein 90 catalyzes trans-cis isomerization of geldanamycin, *Chem. Biol.* **11**, 991–998.
46. Kelland, L. R., Sharp, S. Y., Rogers, P. M., Myers, T. G., and Workman, P. (1999) DT-Diaphorase expression and tumor cell sensitivity to 17-allylamino, 17-demethoxygeldanamycin, an inhibitor of heat shock protein 90, *J. Natl. Cancer Inst.* **91**, 1940–1949.
47. Guo, W., Reigan, P., Siegel, D., Zirrolli, J., Gustafson, D., and Ross, D. (2005) Formation of 17-allylamino-demethoxygeldanamycin (17-AAG) hydroquinone by NAD(P)H:quinone oxidoreductase 1: Role of 17-AAG hydroquinone in heat shock protein 90 inhibition, *Cancer Res.* **65**, 10006–10015.
48. Chiosis, G., Huezio, H., Rosen, N., Mimnaugh, E., Whitesell, L., and Neckers, L. (2003) 17-AAG—Low target binding affinity and potent cell activity: Finding an explanation, *Mol. Cancer Ther.* **2**, 123–129.
49. Chiosis, G., Vilenchik, M., Kim, J., and Solit, D. (2004) Hsp90: The vulnerable chaperone, *Drug Discovery Today* **9**, 881–888.

Small Molecules: Big Players in the Evolution of Protein Synthesis

Sandro F. Ataíde[†] and Michael Ibba^{†,*}

[†]Department of Microbiology and ^{*}Ohio State Biochemistry Program, The Ohio State University, Columbus, Ohio 43210

The amino acids (aa) required for translation of messenger RNA are delivered to the ribosome esterified to the 3'-ends of tRNAs (Figure 1). The aminoacylation of tRNAs is catalyzed by aminoacyl-tRNA synthetases (aaRSs), which must discriminate their unique cognate pair of aa and tRNA from among the vast number of similar molecules that exist in the cell (1). Cell survival is totally dependent on the correct functioning of aaRSs, and several different strategies have evolved to ensure the accurate recruitment of aa during protein synthesis (2–4). In addition to accurate aa recognition, other approaches adopted by aaRSs to maintain fidelity include editing (also known as proofreading), gene duplication, and the use of alternative biosynthetic pathways (5). In-depth studies of the aaRS family have demonstrated how the chemistry of a particular aa influenced the evolution of these enzymes in the different kingdoms of life. Here, we will review how the requirement for strict aa discrimination during protein synthesis to ensure accurate translation of the genetic code played a role in forging enzymes with dedicated active sites and discuss the additional functions required to prevent degeneracy during decoding (summarized in Figure 2). In the first section, background information about the importance of aaRSs for cell viability is presented, indicating the diversity of these enzymes. A description of the mechanism of aa discrimination used by the different aaRSs follows, presenting different modes for targeting these enzymes. Different strategies used by aaRSs for discrimination are presented, such as secondary sites to prevent infiltration of the genetic code, divergent pathways of aa-tRNA biosynthesis in different organisms, and different aspects of duplication of aaRSs and their implications for the development of better aaRS inhibitors.

Aminoacyl-tRNA Synthetases and Translation. The formation of aa-tRNA is a two-step reaction: after binding

ABSTRACT The aminoacyl-tRNA synthetases (aaRSs) are responsible for selecting specific amino acids for protein synthesis, and this essential role in translation has garnered them much attention as targets for novel antimicrobials. Understanding how the aaRSs evolved efficient substrate selection offers a potential route to develop useful inhibitors of microbial protein synthesis. Here, we discuss discrimination of small molecules by aaRSs, and how the evolutionary divergence of these mechanisms offers a means to target inhibitors against these essential microbial enzymes.

*To whom correspondence should be addressed.
E-mail: ibba.1@osu.edu.

Received for review May 14, 2006
and accepted May 22, 2006.

Published online June 16, 2006
10.1021/cb0600200k CCC: \$33.50

© 2006 by American Chemical Society

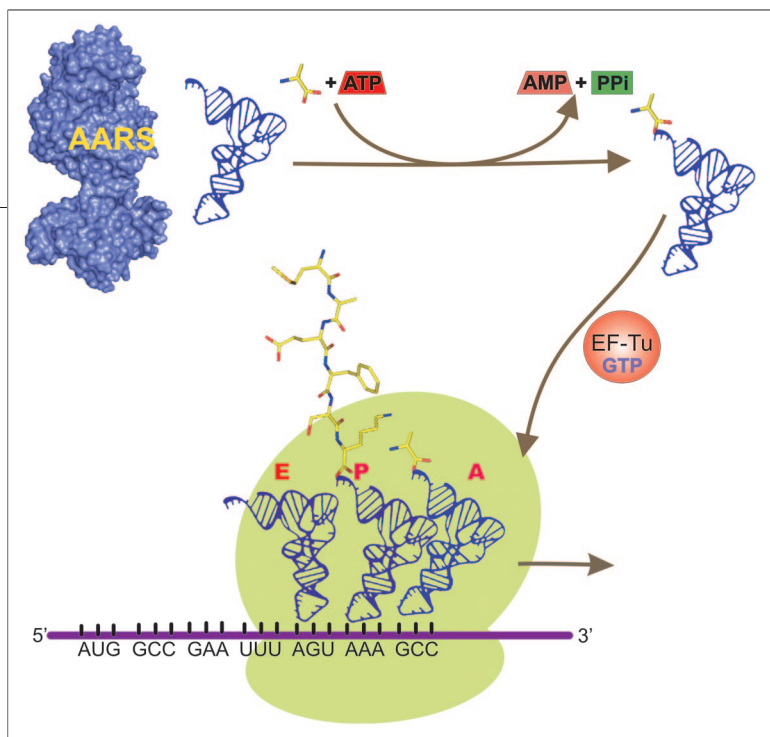


Figure 1. Scheme for the co-translational insertion of an aa in response to a particular codon. tRNA is first aminoacylated with cognate aa. Elongation factor EF-Tu binds aa-tRNA^{aa} forming the aa-tRNA^{aa}-EF-Tu-GTP ternary complex, which delivers the aa-tRNA^{aa} to the ribosomal A site when it is occupied by the corresponding codon on the mRNA.

to the active site, the α -carboxylate of the aa attacks the α -phosphate of ATP leading to the formation of an enzyme-bound mixed anhydride (aminoacyl-adenylate [aa-AMP]) and an inorganic pyrophosphate leaving group; in the second step, the 2'- or 3'-hydroxyl of the terminal ribose of the corresponding tRNA performs a nucleophilic attack on the aminoacyl-adenylate leading to formation of aa-tRNA and an AMP leaving group (1). The overall two-step reaction is certainly common to all synthetases, but whether a common mechanism exists for all aaRS is currently unknown (5). The 20 canonical aaRSs, as found in *Escherichia coli* and eukaryotes, are found in two highly conserved structural groups with 10 members each, classes I and II (6, 7). The class assignments of aaRSs with particular aa specificities have been almost completely conserved through evolution; the only known exception is the representation of lysyl-tRNA synthetase (LysRS) in both groups (see below) (8, 9). Specific structural and mechanistic elements define the members of a class (6, 10). AaRSs from class I possess a Rossman dinucleotide binding domain flanked by two signature motifs, HIGH and KMSKS, while class II contains an active site formed by an extended antiparallel β -sheet structure characterized by three degenerate sequence motifs.

The aaRSs are believed to have evolved from two common ancestors, one from each class, which diverged according to the necessity to discriminate

particular cognate tRNAs and aa (10–12). The essential role of the aaRSs in translation places a strong selective pressure on the evolution of these enzymes to prevent mistakes during cognate aa-tRNA formation. The recognition of tRNA requires the identification of a unique set of elements, nucleotides, or modified nucleotides at particular positions (13). These so-called identity elements of a tRNA are often placed in the acceptor and anticodon stems, the anticodon loop, and the variable arm of the tRNA. Because of their size and complexity, tRNAs offer sufficiently diverse recognition elements to allow their specific selection by the corresponding aaRS. Distinguishing between structurally related aa and other small molecules is considerably more problematic, and occasional errors in substrate selection are unavoidable (14). Consequently, during evolution, certain aaRSs have acquired appended domains that serve to proof-read noncognate aa (15). Other strategies to enhance the specificity of small molecule discrimination have also appeared during aaRS evolution such as gene duplication, *trans*-editing factors, and pre-translational modification. Both cognate recognition and noncognate aa discrimination have had major roles in shaping aaRSs at the levels of individual residues, modules, and

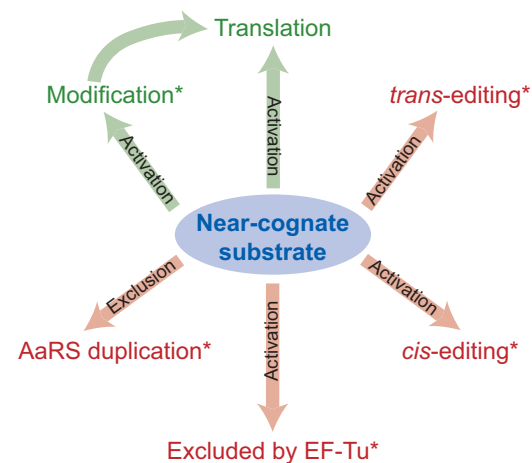


Figure 2. The fate of near-cognate substrates in protein synthesis. Pathways in green lead to translation, while those in red indicate mechanisms by which near-cognate substrates can be excluded. The * denotes pathways that may be exploitable as drug targets due to corresponding differences between bacteria and eukaryotes. Exclusion indicates substrates unable to bind the active site productively, and activation indicates substrates able to enter the aminoacylation pathway.

subunit recruitment. These changes are still evident as a major source of heterogeneity in aaRS structures, presenting the potential for specific antimicrobial targeting (16).

Amino Acid Discrimination in aaRS Active Sites. The existence of heterogeneity in active site discrimination among aaRSs in the different kingdoms of life mainly consists of variation in aa composition of the active site and its surroundings. An example of active site divergence is seen in a seryl-tRNA synthetase (SerRS) from the archaeon *Methanosarcina barkeri* which has a zinc-dependent aa discrimination, while the bacterial type, present in all kingdoms of life, does not require a zinc for aa discrimination (17). A detailed analysis of the different strategies used by aaRSs to discriminate their cognate aa provides the first step toward assessing the use of aaRS as potential drug targets.

AaRSs must specifically recognize their cognate aa from among the vast number of small molecules in the cell with similar physical and chemical properties (14). The presence of both D- and L-enantiomers for each aa, precursors from aa biosynthesis, products of aa degradation, and the natural aa together impose a strong selective pressure for a very specific active site, since all have the potential to disrupt translation (18). Initial selection of certain aa and analogues is dependent on significant differences in size, such as between Gly and Trp, or charge when comparing for example Arg and Glu. The discrimination of aa with smaller differences, for example Asp and Asn, is achieved through a network of highly specific interactions during substrate binding (19, 20). Aspartyl-tRNA synthetase (AspRS) takes advantage of the negative charge of Asp and uses mainly electrostatic interactions with two Arg, one Lys, and one His residues to specifically interact with the two carboxylate groups of Asp. The His residue is important in preventing the binding of Asn as shown by structural, biochemical, and theoretical studies (19). In the closely related asparaginyl-tRNA synthetase (AsnRS), the Lys conserved in the active site of AspRS is replaced by a Gly or Leu and the critical His is absent, allowing preferential binding of Asn rather than Asp (20). Interestingly, asparagine synthase B shares the same binding residues as yeast AspRS; however, Asp binds to the active site in a reverse orientation in order to activate the β -carboxylate with AMP instead of the α -carboxylate as in AspRS (21). The AspRS/AsnRS discrimination mode illustrates how the expansion of the genetic code to accommodate both

Asp and Asn was facilitated by divergent evolution from an ancestral enzyme to generate two aaRSs with high substrate specificities. The discrimination between Glu and Gln relies on the same principle. The aa specificity was demonstrated by replacing the residues required for Gln binding for the residues required for Glu binding in human glutamyl-tRNA synthetase (GlnRS). The new enzyme was able to activate and attach Glu to tRNA^{Gln}, demonstrating how active site specificity could be modified among related aaRSs (22).

In most cases, the fidelity of aa selection is achieved by a network of H-bond and hydrophobic interactions between the aa and the cognate aaRS. Tyrosyl-tRNA synthetase (TyrRS) is the best characterized example of how aaRSs achieve fidelity in aa discrimination (reviewed in ref 23). TyrRS uses an extensive H-bonding network to discriminate Tyr from Phe, and the Y34F replacement disrupts the H-bond network within the active site reducing substrate discrimination (24). However, the substitution W126L enhances the discrimination of Tyr over Phe even with a disruption of an H-bond network with Asp176 (25). The fidelity of aa recognition is dependent on the plasticity of the active site of each aaRS to accommodate the side chains without steric clashes between the substrate and the active site residues of the synthetase (26). Generally, each aaRS has evolved to achieve a useful level of specificity without necessarily maximizing discrimination between the cognate aa and cognate analogues (25–28). This was strikingly illustrated in a recent study that detected substrate analogues for 17 aaRSs that could be aminoacylated to cognate tRNA by the wild-type aaRSs (18). Most of the 92 aa analogues found to be substrates were synthetic compounds and likely never exerted a selective pressure on the aaRS to discriminate against them.

In addition to aa binding by specific interactions in the active site, aaRSs also increase specificity in recognition by induced fit (29). AaRSs which use this mechanism (GlnRS, TyrRS, arginyl- [ArgRS], [GluRS], histidyl- [HisRS], [LysRS1], and threonyl-tRNA synthetase [ThrRS]) rely on communication between distal regions of the protein to sense the binding of the correct substrate in order to recruit the appropriate catalytic residues into a productive position in the active site. ArgRS, GluRS, GlnRS, and LysRS1 share much of the same induced fit mechanism in which the binding of the cognate tRNA is required in order to form a productive active site confor-

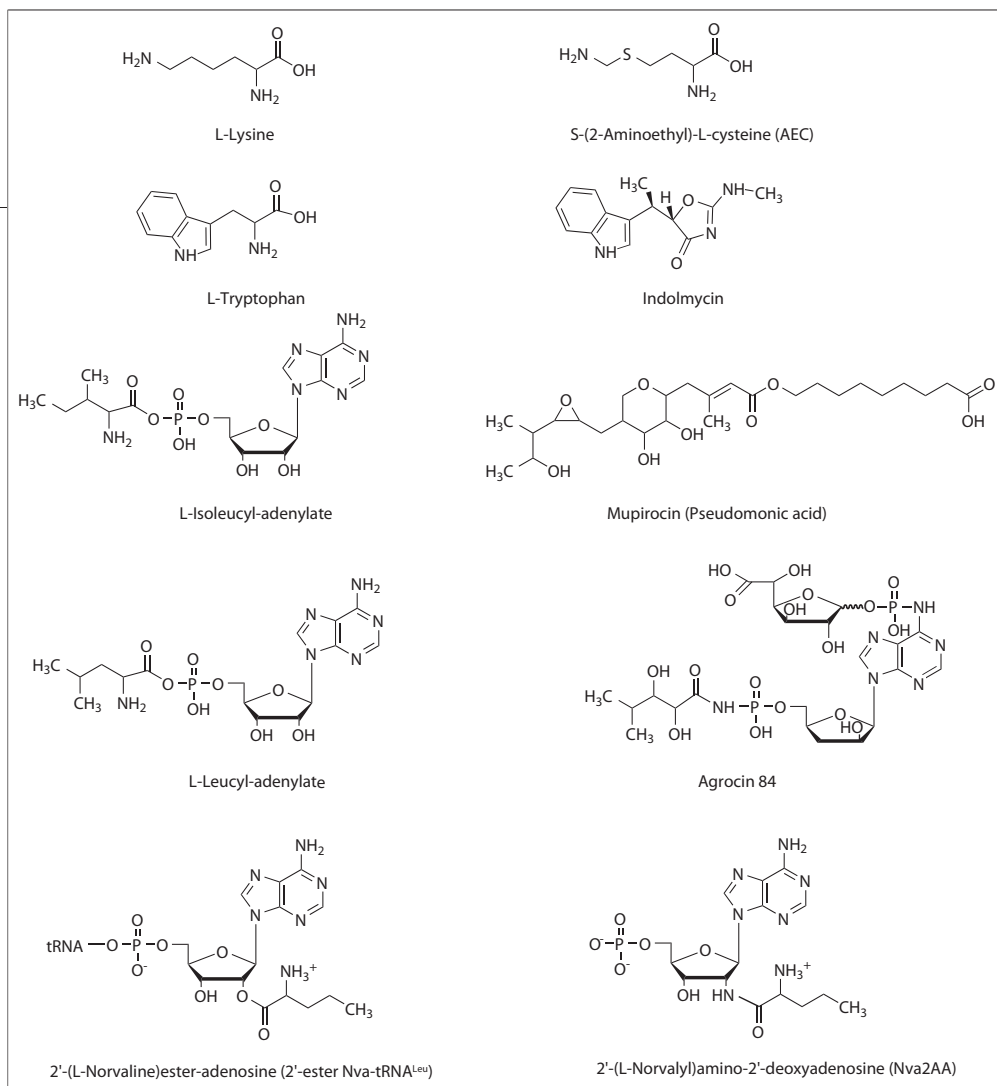


Figure 3. Chemical structures of cognate substrates and inhibitors of aaRSs.

mation and trigger aa activation (19, 30–35). In these examples ATP and cognate aa binding are enhanced upon recognition of the correct tRNA by the anticodon binding region. Furthermore, in GlnRS, binding of cognate tRNA is enhanced if the cognate aa (Gln) is bound to the active site, while the near-cognate aa (Glu) reduces tRNA binding 60-fold (31). In ThrRS and TyrRS, induced fit upon binding of the correct substrate leads directly to formation of a productive active site (36, 37). The indication that induced fit serves as an extra discrimination factor in binding of both cognate aa and cognate tRNA demonstrates how aaRSs have evolved different mechanisms for selection of the correct substrate.

A number of known inhibitors of aaRSs act by interfering with recognition of the cognate aa at the active site. The search for aaRS inhibitors has identified many natural products such as indolmycin (inhibitor of tryptophanyl-tRNA synthetase [TrpRS]) (38), borrelidin (inhibitor of ThrRS) (39), and mupirocin (inhibitor of IleRS) (40, 41). Borrelidin is available for research only

and used strictly as a eukaryotic ThrRS inhibitor in clinical studies against malaria and antiangiogenesis (42, 43). The mode of borrelidin inhibition is by binding to a hydrophobic noncatalytic domain, cluster A, which impairs catalytic conformational changes in ThrRS, resulting in reduced binding of ATP and Thr (44). The antiangiogenesis action of borrelidin is due to inhibition of ThrRS and activation of caspases 3 and 8 to induce apoptosis (42).

Indolmycin is a potent bacterial TrpRS inhibitor that acts as a competitive inhibitor due to a differential binding to the enzyme compared to Trp (Figure 3) (38, 45). Since indolmycin is a biosynthetic derivative of Trp, it has other mechanisms of action inside the cell which cumulatively affect viability (46, 47). Unfortunately, the inhibitory action is not widespread among pathogens, probably due to its hydrophobicity that impairs uptake by certain organisms. Although indolmycin is commercially available for research only, it is not FDA-approved. Indolmycin acts as a bacteriostatic agent against *Staphylococcus aureus*, which can acquire elevated resistance against indolmycin via a point mutation that causes the H43N replacement in TrpRS (48). A bacteriocidal effect of indolmycin was observed against *Helicobacter pylori*, which was unable to develop resistance (49). Also, *in vitro* and *in vivo* studies in *Streptomyces griseus*, which produces indolmycin, revealed a second copy of TrpRS that confers the resistance to indolmycin (45).

Of the numerous aaRS inhibitors, only mupirocin (pseudomonic acid), which inhibits IleRS, is commercially available and FDA-approved (50). Structural and biochemical studies indicate the same overall mode of binding by mupirocin as the Ile-AMP intermediate (Figure 4), the difference being that the inhibitor contains a nonanionic acid moiety which fits into an

unoccupied hydrophobic pocket in IleRS. The binding of mupirocin in the hydrophobic pocket of IleRS then blocks the subsequent binding of isoleucine and ATP, as confirmed by substitution of the nonanionic acid moiety by a shorter alkyl group that resulted in loss of inhibition (40, 41, 50, 51). Because of the competitive inhibition of IleRS and the bacteriostatic effect, several mupirocin-resistant *S. aureus* strains were isolated with point mutations in IleRS (16, 41). Most of the resistance against mupirocin derives from the V588F replacement, which was shown to interact with the nonanionic acid moiety. The mode of action of mupirocin illustrates how exploiting noncatalytic motifs and hydrophobic regions around the active sites can effectively inhibit function and may be exploitable for other aaRSs.

The most potent competitive inhibitors targeted against any of the aaRSs are variants of the aa-AMP, which is the key intermediate of the aminoacylation reaction (52). All the aminoacyl-adenylates present a K_i in the low nanomolar range. Several aa-AMP variants have been synthesized in a nonhydrolyzable form with a sulfamoyl or an aryl-replacement of the α -phosphate group of the adenine (53–55). The nonhydrolyzable analogues are potent inhibitors, but the selectivity of these compounds is not restricted to the bacterial aaRSs, and they can potentially also inhibit the host enzyme. Adenylate analogues are widely used to determine X-ray structures of aaRSs and to understand the interactions of the active site with the substrates in mechanistic studies (56). Several compounds that mimic the aa-AMP intermediate have been characterized in screenings of natural products and in larger synthetic library screenings. Of the identified aa-AMP analogues, Agrocin 84 is known to be a biocontrol of plant tumors caused by *Agrobacterium tumefaciens* and mimics leucyl-adenylate (Figure 3) (57). Agrocin 84 possesses a D-glucofuranosyloxyposphoryl group linked to the adenine moiety, which is important for the uptake of the compound by the pathogen but must be cleaved to release the toxic moiety which can act as a competitive inhibitor of LeuRS. The toxic moiety has a stable phosphoroamidate bond instead of the labile phosphoroanhydride of the genuine aa-AMP, which makes it a better inhibitor (57). A similar strategy is used by the AspRS inhibitor microcin C, a pentapeptide processed by the cell to generate an Asp-AMP analogue with an *N*-acyl phosphoroamidate linkage (58).

The variation displayed by bacterial and eukaryotic aaRSs such as aa insertions and variations around the active site are the key features exploited to date when using these enzymes as potential drug targets. Protein sequence alignments and modeling on 3D X-ray structures are not, however, sufficient to predict such sites (17). Details of these variations will only come from a combination of biochemical and structural studies of different aaRS candidates for drug targeting. Also, understanding the different mechanisms of compound uptake by pathogens can lead to a better design of compounds such as Agrocin 84 and microcin C, which have the potential to be used as the basis for the design of highly selective targeted aaRS inhibitors. The design of these compounds will require a better understanding of the physiology of interesting pathogens and their hosts allied with genomics, proteomics, and the screening of compound libraries. Unfortunately, much still remains to be characterized to provide a comprehensive starting point for the rationale design of drugs with the same, targeted, characteristics of Agrocin 84 and microcin C.

Amino Acid Discrimination in Editing Sites. As in the case of Asp/AsnRS and Glu/GlnRS, another subgroup of aaRSs also clearly shares common origins, but the degree of similarity between their aa substrates imposes additional constraints on accurate recognition. Discrimination between the aliphatic aa Val, Ile, and Leu (59–62) poses an obvious problem, as the small differences in their potential binding energies preclude the high level of specificity observed for Asp and Asn. The corresponding aaRSs have evolved a proofreading mechanism, named editing, which consists of a secondary site which is able to recognize and hydrolyze aa-AMP and misacylated tRNAs (Table 1). The requirement for an editing site is conserved among the different kingdoms of life and strictly required to maintain cell viability (63). In a few cases, when the editing domain is absent or inactive in the aaRS, a *trans*-acting factor with editing activity may come into play. Editing is not, however, ubiquitous; for example, both phenylalanyl-tRNA synthetase (PheRS) and LeuRS from mitochondria have lost their respective editing

KEYWORDS

Transfer RNA (tRNA): Small adaptor RNA responsible for delivering amino acids to the ribosome during translation of messenger RNA.

Aminoacyl-tRNA Synthetases (aaRS): Family of enzymes responsible for attaching amino acids to the 3'-ends of tRNAs.

Amidotransferases: Enzymes responsible for chemical conversion of Glu into Gln or Asp into Asn in a tRNA-dependent manner.

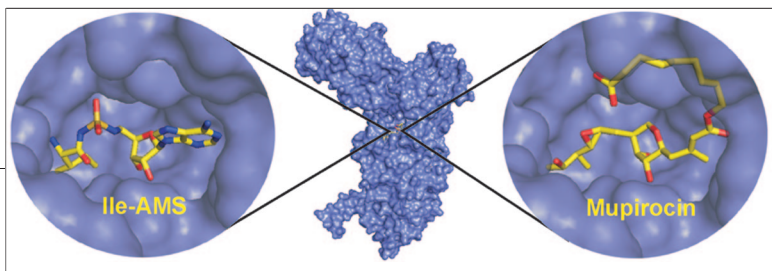


Figure 4. Visualization of the IleRS structure from *Thermus thermophilus* with inhibitor (Mupirocin) or nonhydrolyzable adenylated analogue (Ile-AMS, 5'-O-[N-(L-isoleucyl)sulphamoyl]adenosine) bound to the active site (50). Insets indicate the location of the active site on the IleRS structure. The surface representation is transparent blue in order to provide a realistic view of the active site. Color code for stick structures: carbon, yellow; oxygen, red; nitrogen, blue; and sulfur, orange.

activities (64, 65). Several editing activities of aaRSs are presented below, together with examples of cases in which the editing function was lost and a *trans*-acting factor rescues the activity.

IleRS activates noncognate Val only ~200-fold less efficiently than cognate Ile, creating a potentially high level of misincorporation of noncognate aa by the ribosome during protein synthesis (66). IleRS corrects this potentially catastrophic misactivation of Val by using an editing mechanism that allows hydrolysis at a secondary active site, named the editing site (67–70). This leads to a substantially lower error rate of <1:3000, compatible with the overall level of fidelity observed in translation. LeuRS (62, 71), valyl- (ValRS) (60), methionyl- (MetRS) (72, 73), prolyl- (ProRS) (74), alanyl-tRNA synthetases (AlaRS) (75), ThrRS (76), and PheRS (77, 78) all possess comparable editing activities that minimize the degeneracy of the genetic code by clearing misactivated aa and misacylated tRNAs.

The domains responsible for the hydrolysis of misacylated tRNA differ between class I and class II and even within the same class of aaRS (15). Usually, the proof-

reading domain excludes binding of the cognate aa-tRNA and binds only the misacylated tRNA. The mechanism by which the connective peptide 1 (CP1) domains of IleRS, ValRS, and LeuRS proofread the misacylated tRNA is by accepting the small noncognate aa in the editing site, while excluding the large cognate aa. Subsequently, the amino group of the noncognate aa interacts with a conserved aspartic acid residue present in the editing site, which is responsible for the correct positioning of the substrate for hydrolysis of the ester bond (71, 79, 80). CP1 is a globular domain positioned ~40 Å away from the catalytic site and can be easily accessed by the movement of the CCA end of the tRNA to reposition the misacylated 3'-end into the proofreading site (62). The editing sites of ValRS, LeuRS, and IleRS are present in all kingdoms of life, although LeuRS from human mitochondria has lost the editing function (64).

Class II ThrRS possesses another proofreading motif which is believed to hydrolyze the misacylated tRNA using an H₂O molecule that is deprotonated by a conserved histidine residue present in the proofreading domain (81). The proofreading domains in ThrRS (76) and AlaRS (75) are distinct modules present at the N-terminal and inserted in the C-terminal portion of the enzymes, respectively, while in PheRS, the active and editing sites are in different subunits (65, 77, 78). ProRS can be found in two versions, one “prokaryotic-like”, which contains the editing domain, while the “eukaryotic-like” version present in human does not edit. This feature is not conserved among all eukaryotic-like ProRS, since the archaeal *Methanococcus jannaschii* enzyme has editing activity (82).

In other instances, additional proteins are recruited in *trans* to edit misacylated tRNAs as in the case of YbaK and D-Tyr-tRNA deacylase. YbaK specifically hydrolyzes Cys-tRNA^{Pro} upon interaction with ProRS, but in a free form is unable to compete with EF-Tu for the misacylated tRNA or even discriminate the correct substrate for deacylation (83–86). Two other forms of deacylases are also known: in bacteria and eukaryotes, D-Tyr-tRNA deacylase hydrolyzes D-Tyr-tRNA^{Tyr} and also D-Asp-tRNA^{Asp} and D-Trp-tRNA^{Trp} (87–89); in some archaea, a paralog of D-Tyr-tRNA deacylase instead hydrolyzes misacylated Ser-tRNA^{Thr}, compensating for the lack of an N-terminal proofreading domain in certain ThrRSs (90, 91).

Inhibitory metabolites, or precursors of the cognate aa, which can bypass active and editing site discrimination by aaRSs are good candidates for drug design. For

TABLE 1. AaRS known to possess editing activity and their noncognate substrates

AaRS	Edited noncognate substrates
Class I	
IleRS	Ala, α -aminobutyrate, Cys, homocysteine, homoserine, Thr, Val
LeuRS	Homocysteine, γ -, δ -hydroxyisoleucine, γ -, δ -hydroxyleucine, Ile, Met, norleucine, norvaline
MetRS	Homocysteine
ValRS	Ala, α -aminobutyrate, Cys, Ser, Thr
Class II	
AlaRS	Gly, Ser
LysRS2	Homocysteine, ornithine
PheRS	Ile, Tyr
ProRS	Ala, Cys
ThrRS	Ser

example, α -aminobutyrate is a competitive inhibitor of ValRS (63, 92), and norvaline is a competitive inhibitor of LeuRS (93) which can be incorporated during protein synthesis. A series of new synthetic compounds which can be aminoacylated onto the cognate tRNA by aaRSs was reported recently (18), and the different chemical functionalities present in those compounds can be useful to investigate candidates for inhibition of both, active and editing sites. A potential class of inhibitors that has not been investigated are compounds that mimic the noncognate aminoacyl-tRNA form such as 2'-(L-norvalyl)amino-2'-deoxyadenosine (Nva2AA) which inhibits LeuRS (71) (Figure 5), thereby targeting the proofreading domain of a specific aaRS (16). While evolutionary divergence of editing sites confirms these as promising targets, difficulties in synthesizing the equivalent of misacylated tRNA in a stable form and analysis of the specific inhibitory mode and its cellular consequences still need to be investigated. Nevertheless, in many cases, cellular viability necessitates careful proofreading activity, suggesting this class of compounds (63, 94), which would target the proofreading domains of aaRSs, as promising new candidates with a strong potential for drug target development. In addition to these elaborations of aaRSs, other post-aminoacylation pathways have also evolved both to allow code expansion and to provide new aa biosynthetic pathways.

Heterogeneity in Alternative aa-tRNA Synthesis.

Some bacteria and archaea lack one or more aaRSs, and a tRNA-dependent pathway is used in those organisms to synthesize the cognate aa. Because of the heterogeneity of these pathways and their obligate activity to sustain life, understanding their mechanism of action and developing compounds to target these enzymes is a promising source of drug targets.

Many prokaryotes lack AsnRS and/or GlnRS, and in those organisms, the AspRS and/or GluRS are able to mischarge tRNA^{Asn} or tRNA^{Gln} with Asp or Glu, respectively. An amidotransferase then converts the Asp-tRNA^{Asn} or Glu-tRNA^{Gln} into Asn-tRNA^{Asn} or Gln-tRNA^{Gln}, respectively (Table 2) (95–99). While these enzymes are predominantly microbial, the Glu-tRNA^{Gln} amidotransferase is found in chloroplasts and was thought to be present in mitochondria (100). However, it was recently shown that mitochondria instead use GlnRS, and the amidotransferases have become strong candidates to be exploited as bacterial drug targets

(101). The synthesis of initiator fMet-tRNA^{fMet} is also dependent on a similar mechanism in bacteria and organelles (102, 103). Initially, tRNA^{fMet} is aminoacylated with Met by MetRS and further converted by Met-tRNA formylase into fMet-tRNA^{fMet}, which is then used in translation initiation (Table 2) (104).

A related method used to ensure the correct charging of the cognate tRNA is the exclusive biosynthesis of a certain aa on the 3'-end of the tRNA. The best known example is the synthesis of selenocysteine-tRNA^{Sec}, which starts with seryl-tRNA synthetase (SerRS) misacylating tRNA^{Sec} with serine. Selenocysteine synthase (SelA) then catalyzes the conversion of serine to selenocysteine on tRNA^{Sec} (105–107). Another recently described example of the synthesis of an aa directly on tRNA involves Cys biosynthesis in certain archaea. In these organisms, a new aaRS named *ortho*-phosphoseryl-tRNA synthetase (SepRS) first catalyzes the attachment of *ortho*-phosphoserine to tRNA^{Cys} which is further converted into Cys-tRNA^{Cys} by the Sep-tRNA^{Cys}-tRNA synthase (SepCysS) enzyme (108).

The overall Sec pathway is universally conserved, including the synthesis of Sec-tRNA^{Sec}, the decoding of a UGA codon as a Sec, the presence of a Selenocysteine Insertion Sequence (SECIS) element, and the presence of a dedicated EF-Tu homologue (SelB) for delivery of Sec-tRNA^{Sec} (107). However, among the kingdoms of life, the differences in the Sec pathway can potentially be exploited as drug targets; for example, no homologue of bacterial SelA has yet been found in eukaryotes, and the pathway is believed to be substantially different (109). Also, the SECIS elements of bacterial mRNA are often placed downstream of the AUG site in a coding region, while in eukaryotes and archaea, they are always in the 3' untranslated region (110). The differences in SelB lie in the fact that the bacterial version recognizes and interacts with SECIS through its C-terminal region, while the eukaryal and archaeal forms lack the C-terminal domain and instead interact with SECIS binding protein 2 (SBP2) (111, 112). All these differences in the Sec synthesis and insertion pathways have the potential to be exploited as antimicrobial drug targets.

KEYWORDS

Deacylation: Hydrolysis of ester bond between amino acid and 2'- or 3'-hydroxyl of the 3'-end of the tRNA.

Misacylation: Mistake committed by aaRS in attaching a near- or noncognate amino acid to the cognate tRNA.

Editing site: Secondary site in aaRS responsible for deacylation of misacylated tRNA.

Genetic code degeneracy: Error in decoding the genetic code upon incorporation of the wrong amino acid in response to a particular mRNA codon.

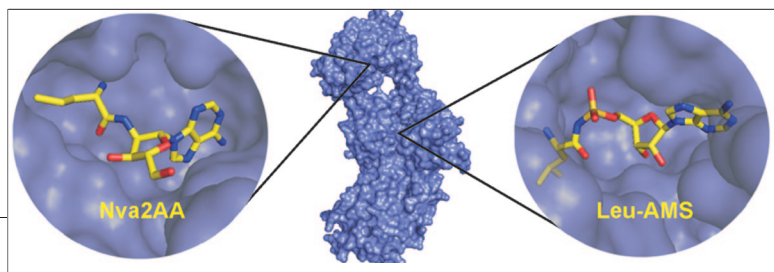


Figure 5. Visualization of the LeuRS structure from *Thermus thermophilus* with editing site inhibitor (Nva2AA) and nonhydrolyzable adenylated analogue (Leu-AMS, 5'-O-[N-(L-leucyl)sulphamoyl]adenosine) bound to the editing and active site, respectively (71). Insets indicate the location of the editing site at the left and active site at the right side of the LeuRS structure. The surface representation is transparent blue in order to provide a realistic view of the editing and active sites. Color code for stick structures: carbon, yellow; oxygen, red; nitrogen, blue; and sulfur, orange.

The heterogeneity of tRNA-dependent aa synthesis provides many promising targets; however, structural and functional studies must still be performed to understand the potential for inhibition and the cellular consequences of inhibiting these pathways. Compounds targeting the different enzymes involved in amidotransferase, formylase, and selenocysteine synthesis are promising candidates for drug development. In addition to these post-aminoacylation pathways, orthologous and nonorthologous duplication of aaRSs provides another new opening both for drug design and studies of potential antibiotic resistance.

Discrimination of Small Molecules through aaRS Duplication. In addition to the various mechanisms described above, some organisms possess a second ortholog of the same aaRS, which often confers resistance to certain conditions or inhibitors by discriminating against small molecules. The existence of such duplicates occurs in organisms exposed to certain inhibitory compounds that, in order to survive, evolved a resistant version of the same aaRS. Although this implies that gene duplication events can be a pitfall for drug targeting, in fact they can also be used as models, since the resistance mechanism can be predicted, studied, and characterized to provide a means to develop better inhibitors that circumvent known problems.

TrpRS can be found in duplicate in streptomycetes that produce indolmycin, where only one constitutively expressed copy is sensitive to indolmycin, while the second is expressed to rescue the cell when the inhibitor is synthesized (45). TrpRS can also be found in *Deinococcus radiodurans* as two less similar variants, TrpRS1 and TrpRS2. In this case, the housekeeping TrpRS1 is resistant to inhibition and misacylation of tRNA^{Trp} with Trp analogues, while TrpRS2

is able to charge tRNA^{Trp} with 4-nitro-tryptophan and 5-hydroxy-tryptophan (113, 114).

The most intriguing case of functional duplication involves LysRS, which can be present either as two versions of class II-type LysRS (LysRS2) with a house-keeping version *lysS* and a heat-shock-inducible version *lysU*, or as the structurally unrelated class I-type LysRS (LysRS1) and LysRS2. In organisms that harbor two copies of LysRS2, the second copy (*lysU*) is expressed under specific conditions such as heat shock and in media containing leucine. The inducible LysRS2 is resistant to inhibition by lysine analogues such as cadaverine, indicating that this duplication of activity is important to avoid misincorporation of lysine analogues under stress conditions (115–117). LysRS1 exists mainly in archaea and some pathogenic bacteria, and this enzyme is more selective and less prone to inhibition when compared to LysRS2 (118). The different versions of LysRS take advantage of different networks of interactions in their active sites to discriminate the correct substrate; where LysRS2 relies on a large network of H-bonding interactions, LysRS1 employs a minimal H-bond interaction in a specific part of the active site. Discrimination against near-cognate analogues which are natural metabolites such as S-(2-aminoethyl)-L-cysteine (AEC), homocysteine, and ornithine can be a serious problem for LysRS2 which does not possess post-transfer editing activity (115, 118). LysRS2 can activate homocysteine and ornithine with ATP, but before they can be transferred to tRNA^{Lys}, they are further cyclized into homocysteinethiolactone and ornithine lactone, respectively, and released (115).

AEC is potentially more problematic since, once activated by LysRS2, it can be transferred to tRNA^{Lys} and used in protein synthesis by the ribosome, inhibiting cell

TABLE 2: Noncanonical aminoacyl-tRNA synthesis in translation

Aminoacyl-tRNA	AaRS	Substrate for	Final product for translation
Asp-tRNA ^{Asn}	AspRS	gatCAB (amidotransferase)	Asn-tRNA ^{Asn}
Sep-tRNA ^{Cys}	SepRS	SepCysS	Cys-tRNA ^{Cys}
Met-tRNA ^{fMet}	MetRS	Met-tRNA formylase	fMet-tRNA ^{fMet}
Glu-tRNA ^{Gln}	GluRS	gatCAB/gatDE (amidotransferase)	Gln-tRNA ^{Gln}
Pyl-tRNA ^{Pyl}	PylRS	Translation	Pyl-tRNA ^{Pyl}
Ser-tRNA ^{Sec}	SerRS	SelA	Sec-tRNA ^{Sec}

growth (118). The class I version of LysRS is always less active when compared to the class II version, which seems to be more robust with respect to cognate aminoacylation (27, 28, 118). Interestingly, LysRS1 is not significantly inhibited by AEC either *in vitro* or *in vivo*, indicating different mechanisms of recognition and discrimination of lysine and analogues compared to LysRS2 (Figure 3) (118, 119). The fact that AEC is a naturally occurring metabolite (discussed in ref 119), and represents one of the possible compounds which might have exerted a selective pressure in the evolution of two distinct LysRSs with different discrimination of lysine analogues, makes LysRS1 a good candidate to be exploited as an antimicrobial drug target. The requirements of an organism to retain either LysRS1 or LysRS2 are likely based on the necessity for a more selective enzyme *versus* a more active one, possibly depending on growth physiology and the presence of inhibitors in the environment (27, 28).

TyrRS, ThrRS, IleRS, MetRS, and TrpRS can also be found in certain genomes as apparent duplicates, and the expression of the second copy is often dependent upon the presence of inhibitors for the housekeeping version or changes in cellular physiology (45, 120–124). The extensive duplication seen among aaRSs also has important practical implications. Clinical treatment with mupirocin has resulted in the emergence of two resistance mechanisms: high-level resistance, which derives from acquisition of a plasmid-borne IleRS ortholog, or low-level resistance resulting from point mutations at two sites in the gene encoding IleRS (41, 125–127). Studies have demonstrated that the low level of resistance can be reverted by compensatory mutations acquired when the organism is not under antimicrobial pressure (41, 125).

Screening of new potent synthetic drugs targeting bacterial MetRS has identified resistant organisms harboring two copies of MetRS. The second copy, which is resistant to the synthetic compounds, possesses an insertion of 27 aa around the active site and is present in several organisms (121). Bioinformatics studies have identified the origin of the second copy of MetRS as deriving from soil-dwelling bacteria which have never been exposed to these synthetic inhibitors (120). These examples summarize how aaRSs can readily adapt and evolve to gain specific functions and avoid degeneracy during decoding. The robust scaffold on which aaRSs are built allows such adaptations to different conditions to which an organism might be exposed, such as the

presence of potent inhibitors. While a drawback in some instances, gene duplication can be exploited to understand the extent of plasticity of aaRSs and characterize the weaknesses of both copies, thereby using this feature of aaRSs to develop new drugs potentially less susceptible to resistance.

Conclusions. In recent years, the discovery of new functions and activities among the aaRSs has broadened our knowledge of these essential housekeeping enzymes. Proteomics studies, along with the search for other substrates within the cell, have shown that aaRSs are more versatile than previously believed. In parallel, efforts to expand the genetic code to allow the insertion of unnatural aa have focused on first modifying aaRS specificity (14, 128). On the basis of the often severe loss of activity in many of these modified aaRSs, and studies with inhibitors on the mechanism of substrate recognition, it is becoming clearer how specificity has evolved in these enzymes. Certain aaRSs have apparently evolved to acquire the best balance between activity and specificity, depending on how easily a mistake could be made in misacylating cognate tRNA under particular physiological conditions.

Contrary to the common belief that aaRS evolution was primarily driven by tRNA recognition, there has been significant selective pressure to discriminate the cognate aa from other small molecules in the cell. A particularly striking illustration of this comes from the recent discovery of an aaRS specific for the rare amino acid pyrrolysine (129, 130). Understanding the mechanisms by which aaRSs discriminate small molecules provides the opportunity to design drugs that are small molecule inhibitors of specific aaRSs. Characterization of the mechanisms of substrate recognition at the active and editing sites provides a rationale for the development of potent inhibitors for a specific aaRS, which will require structural and biochemical data to improve the design of drugs targeting each of these sites. In parallel, genomics, bioinformatics, and proteomics offer a better potential to identify *trans*-editing proteins and tRNA-dependent aa biosynthesis and circumvent the most common mode of aaRS resistance, namely, duplication, further enhancing the promise of success for new inhibitors.

Acknowledgment: We thank C. Hausmann, J. Ling, and T. E. Rogers for critical reading of the manuscript, and the anonymous reviewers for their helpful and constructive comments. This work was supported by Grant GM 65183 from the National Institutes of Health. S.F.A. is supported by an American Heart Association Predoctoral Fellowship.

REFERENCES

- Amez, J. G., and Moras, D. (1997) Structural and functional considerations of the aminoacylation reaction, *Trends Biochem. Sci.* **22**, 211–216.
- O'Donoghue, P., and Luthey-Schulten, Z. (2003) On the evolution of structure in aminoacyl-tRNA synthetases, *Microbiol. Mol. Biol. Rev.* **67**, 550–573.
- Ibba, M., and Söll, D. (1999) Quality control mechanisms during translation, *Science* **286**, 1893–1897.
- Dale, T., and Uhlenbeck, O. C. (2005) Amino acid specificity in translation, *Trends Biochem. Sci.* **30**, 659–665.
- Ibba, M., and Söll, D. (2004) Aminoacyl-tRNAs: setting the limits of the genetic code, *Genes Dev.* **18**, 731–738.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs, *Nature* **347**, 203–206.
- Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N., and Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å, *Nature* **347**, 249–255.
- Ibba, M., Morgan, S., Cumow, A. W., Pridmore, D. R., Voithknecht, U. C., Gardner, W., Lin, W., Woese, C. R., and Söll, D. (1997) A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases, *Science* **278**, 1119–1122.
- Terada, T., Nureki, O., Ishitani, R., Ambrogelly, A., Ibba, M., Söll, D., and Yokoyama, S. (2002) Functional convergence of two lysyl-tRNA synthetases with unrelated topologies, *Nat. Struct. Biol.* **9**, 257–262.
- Cusack, S. (1993) Sequence, structure and evolutionary relationships between class 2 aminoacyl-tRNA synthetases: an update, *Biochimie* **75**, 1077–1081.
- Nagel, G. M., and Doolittle, R. F. (1995) Phylogenetic analysis of the aminoacyl-tRNA synthetases, *J. Mol. Evol.* **40**, 487–498.
- Ribas de Pouplana, L., and Schimmel, P. (2001) Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem, *Cell* **104**, 191–193.
- Giegé, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res.* **26**, 5017–5035.
- Hendrickson, T. L., de Crécy Lagard, V., and Schimmel, P. (2004) Incorporation of nonnatural amino acids into proteins, *Annu. Rev. Biochem.* **73**, 147–176.
- Geslain, R., and Ribas de Pouplana, L. (2004) Regulation of RNA function by aminoacylation and editing?, *Trends Genet.* **20**, 604–610.
- Hurdle, J. G., O'Neill, A. J., and Chopra, I. (2005) Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents, *Antimicrob. Agents Chemother.* **49**, 4821–4833.
- Bilokapic, S., Maier, T., Ahel, D., Gruic-Sovulj, I., Söll, D., Weygand-Durasevic, I., and Ban, N. (2006) Structure of the unusual seryl-tRNA synthetase reveals a distinct zinc-dependent mode of substrate recognition, *EMBO J.* **25**, 2498–2509.
- Hartman, M. C., Josephson, K., and Szostak, J. W. (2006) Enzymatic aminoacylation of tRNA with unnatural amino acids, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4356–4361.
- Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J. C., and Moras, D. (1994) The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction, *EMBO J.* **13**, 327–337.
- Berthet-Colominas, C., Seignovert, L., Härtlein, M., Grotli, M., Cusack, S., and Leberman, R. (1998) The crystal structure of asparaginyl-tRNA synthetase from *Thermus thermophilus* and its complexes with ATP and asparaginyl-adenylate: the mechanism of discrimination between asparagine and aspartic acid, *EMBO J.* **17**, 2947–2960.
- Nakatsu, T., Kato, H., and Oda, J. (1998) Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase, *Nat. Struct. Biol.* **5**, 15–19.
- Agou, F., Quevillon, S., Kerjan, P., and Mirande, M. (1998) Switching the amino acid specificity of an aminoacyl-tRNA synthetase, *Biochemistry* **37**, 11309–11314.
- First, E. A. (1997) in *Comprehensive Biological Catalysis: A Mechanistic Reference* (Sinnott, M., Garner, C. D., First, E. A., and Davies, G., Eds.) pp 573–607, Academic Press Limited, San Diego, CA.
- Wells, T. N., and Fersht, A. R. (1986) Use of binding energy in catalysis analyzed by mutagenesis of the tyrosyl-tRNA synthetase, *Biochemistry* **25**, 1881–1886.
- de Prat Gay, G., Duckworth, H. W., and Fersht, A. R. (1993) Modification of the amino acid specificity of tyrosyl-tRNA synthetase by protein engineering, *FEBS Lett.* **318**, 167–171.
- Tumer, J. M., Graziano, J., Spraggon, G., and Schultz, P. G. (2006) Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6483–6488.
- Ataide, S. F., and Ibba, M. (2004) Discrimination of cognate and noncognate substrates at the active site of class II lysyl-tRNA synthetase, *Biochemistry* **43**, 11836–11841.
- Wang, S., Prætorius-Ibba, M., Ataide, S. F., Roy, H., and Ibba, M. (2006) Discrimination of cognate and noncognate substrates at the active site of class I lysyl-tRNA synthetase, *Biochemistry* **45**, 3646–3652.
- Alexander, R. W., and Schimmel, P. (2001) Domain-domain communication in aminoacyl-tRNA synthetases, *Prog. Nucleic Acid Res. Mol. Biol.* **69**, 317–349.
- Sherlin, L. D., and Perona, J. J. (2003) tRNA-dependent active site assembly in a class I aminoacyl-tRNA synthetase, *Structure* **11**, 591–603.
- Uter, N. T., Gruic-Sovulj, I., and Perona, J. J. (2005) Amino acid-dependent transfer RNA affinity in a class I aminoacyl-tRNA synthetase, *J. Biol. Chem.* **280**, 23966–23977.
- Cavarelli, J., Delagoutte, B., Eriani, G., Gangloff, J., and Moras, D. (1998) L-Arginine recognition by yeast arginyl-tRNA synthetase, *EMBO J.* **17**, 5438–5448.
- Delagoutte, B., Moras, D., and Cavarelli, J. (2000) tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding, *EMBO J.* **19**, 5599–5610.
- Sekine, S., Nureki, O., Dubois, D. Y., Bernier, S., Chenevert, R., Lapointe, J., Vassilyev, D. G., and Yokoyama, S. (2003) ATP binding by glutamyl-tRNA synthetase is switched to the productive mode by tRNA binding, *EMBO J.* **22**, 676–688.
- Ibba, M., Losey, H. C., Kawarabayasi, Y., Kikuchi, H., Bunjun, S., and Söll, D. (1999) Substrate recognition by class I lysyl-tRNA synthetases: a molecular basis for gene displacement, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 418–423.
- Bovee, M. L., Pierce, M. A., and Francklyn, C. S. (2003) Induced fit and kinetic mechanism of adenylation catalyzed by *Escherichia coli* threonyl-tRNA synthetase, *Biochemistry* **42**, 15102–15113.
- Kobayashi, T., Takimura, T., Sekine, R., Kelly, V. P., Kamata, K., Sakamoto, K., Nishimura, S., and Yokoyama, S. (2005) Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase, *J. Mol. Biol.* **346**, 105–117.
- Werner, R. G., Thorpe, L. F., Reuter, W., and Nierhaus, K. H. (1976) Indolmycin inhibits prokaryotic tryptophanyl-tRNA ligase, *Eur. J. Biochem.* **68**, 1–3.
- Nass, G., and Thomale, J. (1974) Alteration of structure of level of threonyl-tRNA-synthetase in Borrelidin resistant mutants of *E. coli*. *FEBS Lett.* **39**, 182–186.
- Yanagisawa, T., Lee, J. T., Wu, H. C., and Kawakami, M. (1994) Relationship of protein structure of isoleucyl-tRNA synthetase with pseudomonic acid resistance of *Escherichia coli*. A proposed mode of action of pseudomonic acid as an inhibitor of isoleucyl-tRNA synthetase, *J. Biol. Chem.* **269**, 24304–24309.

41. Hurdle, J. G., O'Neill, A. J., Ingham, E., Fishwick, C., and Chopra, I. (2004) Analysis of mupirocin resistance and fitness in *Staphylococcus aureus* by molecular genetic and structural modeling techniques, *Antimicrob. Agents Chemother.* **48**, 4366–4376.
42. Kawamura, T., Liu, D., Towle, M. J., Kageyama, R., Tsukahara, N., Wakabayashi, T., and Littlefield, B. A. (2003) Anti-angiogenesis effects of borrelidin are mediated through distinct pathways: threonyl-tRNA synthetase and caspases are independently involved in suppression of proliferation and induction of apoptosis in endothelial cells, *J. Antibiot. (Tokyo)* **56**, 709–715.
43. Otoguro, K., Ui, H., Ishiyama, A., Kobayashi, M., Togashi, H., Takahashi, Y., Masuma, R., Tanaka, H., Tomoda, H., Yamada, H., and Omura, S. (2003) In vitro and in vivo antimalarial activities of a non-glycosidic 18-membered macrolide antibiotic, borrelidin, against drug-resistant strains of *Plasmodia*, *J. Antibiot.* **56**, 727–729.
44. Ruan, B., Bovee, M. L., Sacher, M., Stathopoulos, C., Poralla, K., Francklyn, C. S., and Söll, D. (2005) A unique hydrophobic cluster near the active site contributes to differences in borrelidin inhibition among threonyl-tRNA synthetases, *J. Biol. Chem.* **280**, 571–577.
45. Kitabatake, M., Ali, K., Demain, A., Sakamoto, K., Yokoyama, S., and Söll, D. (2002) Indolmycin resistance of *Streptomyces coelicolor* A3(2) by induced expression of one of its two tryptophanyl-tRNA synthetases, *J. Biol. Chem.* **277**, 23882–23887.
46. Werner, R. G., and Reuter, W. (1979) Interaction of indolmycin in the metabolism of tryptophan in rat liver, *Arzneim-Forsch.* **29**, 59–63.
47. Pohlmann, J., and Brotz-Oesterhelt, H. (2004) New aminoacyl-tRNA synthetase inhibitors as antibacterial agents, *Curr. Drug Targets: Infect. Disord.* **4**, 261–272.
48. Hurdle, J. G., O'Neill, A. J., and Chopra, I. (2004) Anti-staphylococcal activity of indolmycin, a potential topical agent for control of staphylococcal infections, *J. Antimicrob. Chemother.* **54**, 549–552.
49. Kanamaru, T., Nakano, Y., Toyoda, Y., Miyagawa, K. I., Tada, M., Kaisho, T., and Nakao, M. (2001) In vitro and in vivo antibacterial activities of TAK-083, an agent for treatment of *Helicobacter pylori* infection, *Antimicrob. Agents Chemother.* **45**, 2455–2459.
50. Nakama, T., Nureki, O., and Yokoyama, S. (2001) Structural basis for the recognition of isoleucyl-adenylate and an antibiotic, mupirocin, by isoleucyl-tRNA synthetase, *J. Biol. Chem.* **276**, 47387–47393.
51. Silvian, L. F., Wang, J., and Steitz, T. A. (1999) Insights into editing from an Ile-tRNA synthetase structure with tRNA^{Ile} and mupirocin, *Science* **285**, 1074–1077.
52. Bernier, S., Dubois, D. Y., Therrien, M., Lapointe, J., and Chênevert, R. (2000) Synthesis of glutaminyl adenylate analogues that are inhibitors of glutaminyl-tRNA synthetase, *Bioorg. Med. Chem. Lett.* **10**, 2441–2444.
53. Brown, M. J., Mensah, L. M., Doyle, M. L., Broom, N. J., Osbourne, N., Forrest, A. K., Richardson, C. M., O'Hanlon, P. J., and Pope, A. J. (2000) Rational design of femtomolar inhibitors of isoleucyl tRNA synthetase from a binding model for pseudomonadic acid-A, *Biochemistry* **39**, 6003–6011.
54. Forrest, A. K., Jarvest, R. L., Mensah, L. M., O'Hanlon, P. J., Pope, A. J., and Sheppard, R. J. (2000) Aminoalkyl adenylate and aminoacyl sulfamate intermediate analogues differing greatly in affinity for their cognate *Staphylococcus aureus* aminoacyl tRNA synthetases, *Bioorg. Med. Chem. Lett.* **10**, 1871–1874.
55. Tao, J., and Schimmel, P. (2000) Inhibitors of aminoacyl-tRNA synthetases as novel anti-infectives, *Expert Opin. Invest. Drugs* **9**, 1767–1775.
56. Chênevert, R., Bernier, S., and Lapointe, J. (2003) In *Translation Mechanisms* (Lapointe, J., and Brakier-Gingras, L., Eds.), pp 416–427, Springer, New York.
57. Reader, J. S., Ordoukhanian, P. T., Kim, J. G., de Crécy Lagard, V., Hwang, I., Farrand, S., and Schimmel, P. (2005) Major biocontrol of plant tumors targets tRNA synthetase, *Science* **309**, 1533.
58. Metlitskaya, A., Kazakov, T., Kommer, A., Pavlova, O., Prætorius-Ibba, M., Ibba, M., Krasheninnikov, I., Kolb, V., Khmel, I., and Severinov, K. (2006) Aspartyl-tRNA synthetase is the target of peptidenucleotide antibiotic microcin C, *J. Biol. Chem.*, published online Mar 30, <http://dx.doi.org/10.1074/jbc.M513174200>.
59. Nureki, O., Vassilyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P., and Yokoyama, S. (1998) Enzyme structure with two catalytic sites for double-sieve selection of substrate, *Science* **280**, 578–582.
60. Fukai, S., Nureki, O., Sekine, S., Shimada, A., Tao, J., Vassilyev, D. G., and Yokoyama, S. (2000) Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA^{Val} and valyl-tRNA synthetase, *Cell* **103**, 793–803.
61. Fukunaga, R., and Yokoyama, S. (2005) Structural basis for non-cognate amino acid discrimination by the valyl-tRNA synthetase editing domain, *J. Biol. Chem.* **280**, 29937–29945.
62. Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S., and Cusack, S. (2005) The crystal structure of leucyl-tRNA synthetase complexed with tRNA^{Leu} in the post-transfer-editing conformation, *Nat. Struct. Mol. Biol.* **12**, 923–930.
63. Nangle, L. A., de Crécy Lagard, V., Döring, V., and Schimmel, P. (2002) Genetic code ambiguity. Cell viability related to the severity of editing defects in mutant tRNA synthetases, *J. Biol. Chem.* **277**, 45729–45733.
64. Lue, S. W., and Kelley, S. O. (2005) An aminoacyl-tRNA synthetase with a defunct editing site, *Biochemistry* **44**, 3010–3016.
65. Roy, H., Ling, J., Alfonso, J., and Ibba, M. (2005) Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase, *J. Biol. Chem.* **280**, 38186–38192.
66. Schmidt, E., and Schimmel, P. (1995) Residues in a class I tRNA synthetase which determine selectivity of amino acid recognition in the context of tRNA, *Biochemistry* **34**, 11204–11210.
67. Nomanbhoy, T. K., Hendrickson, T. L., and Schimmel, P. (1999) Transfer RNA-dependent translocation of misactivated amino acids to prevent errors in protein synthesis, *Mol. Cell* **4**, 519–528.
68. Hendrickson, T. L., Nomanbhoy, T. K., and Schimmel, P. (2000) Errors from selective disruption of the editing center in a tRNA synthetase, *Biochemistry* **39**, 8180–8186.
69. Hendrickson, T. L., Nomanbhoy, T. K., de Crécy Lagard, V., Fukai, S., Nureki, O., Yokoyama, S., and Schimmel, P. (2002) Mutational separation of two pathways for editing by a class I tRNA synthetase, *Mol. Cell* **9**, 353–362.
70. Pezo, V., Metzgar, D., Hendrickson, T. L., Waas, W. F., Hazebrouck, S., Döring, V., Marlière, P., Schimmel, P., and de Crécy-Lagard, V. (2004) Artificially ambiguous genetic code confers growth yield advantage, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8593–8597.
71. Lincecum, T. L., Jr., Tukalo, M., Yaremchuk, A., Mursinna, R. S., Williams, A. M., Sproat, B. S., Van Den, E. W., Link, A., Van Calenberg, S., Grotli, M., Martinis, S. A., and Cusack, S. (2003) Structural and mechanistic basis of pre- and posttransfer editing by leucyl-tRNA synthetase, *Mol. Cell* **11**, 951–963.
72. Gao, W., Goldman, E., and Jakubowski, H. (1994) Role of carboxy-terminal region in proofreading function of methionyl-tRNA synthetase in *Escherichia coli*, *Biochemistry* **33**, 11528–11535.
73. Jakubowski, H. (1996) The synthetic/editing active site of an aminoacyl-tRNA synthetase: evidence for binding of thiols in the editing subsite, *Biochemistry* **35**, 8252–8259.
74. Beuning, P. J., and Musier-Forsyth, K. (2000) Hydrolytic editing by a class II aminoacyl-tRNA synthetase, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8916–8920.
75. Beebe, K., Ribas de Pouplana, L., and Schimmel, P. (2003) Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability, *EMBO J.* **22**, 668–675.

76. Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C. S., Ehresmann, C., and Moras, D. (2000) Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem, *Cell* **103**, 877–884.
77. Roy, H., Ling, J., Imov, M., and Ibba, M. (2004) Post-transfer editing in vitro and in vivo by the beta subunit of phenylalanyl-tRNA synthetase, *EMBO J.* **23**, 4639–4648.
78. Kotik-Kogan, O., Moor, N., Tworowski, D., and Saforo, M. (2005) Structural basis for discrimination of L-phenylalanine from L-tyrosine by phenylalanyl-tRNA synthetase, *Structure*. **13**, 1799–1807.
79. Mursinna, R. S., Lee, K. W., Briggs, J. M., and Martinis, S. A. (2004) Molecular dissection of a critical specificity determinant within the amino acid editing domain of leucyl-tRNA synthetase, *Biochemistry* **43**, 155–165.
80. Williams, A. M., and Martinis, S. A. (2006) Mutational unmasking of a tRNA-dependent pathway for preventing genetic code ambiguity, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3586–3591.
81. Dock-Bregeon, A. C., Rees, B., Torres-Larios, A., Bey, G., Caillet, J., and Moras, D. (2004) Achieving error-free translation; the mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution, *Mol. Cell* **16**, 375–386.
82. Beuning, P. J., and Musier-Forsyth, K. (2001) Species-specific differences in amino acid editing by class II prolyl-tRNA synthetase, *J. Biol. Chem.* **276**, 30779–30785.
83. Zhang, H., Huang, K., Li, Z., Banerjee, L., Fisher, K. E., Grishin, N. V., Eisenstein, E., and Herzberg, O. (2000) Crystal structure of YbaK protein from *Haemophilus influenzae* (HI1434) at 1.8 Å resolution: functional implications, *Proteins* **40**, 86–97.
84. An, S., and Musier-Forsyth, K. (2004) Trans-editing of Cys-tRNA^{Pro} by *Haemophilus influenzae* YbaK protein, *J. Biol. Chem.* **279**, 42359–42362.
85. Wong, F. C., Beuning, P. J., Silvers, C., and Musier-Forsyth, K. (2003) An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing, *J. Biol. Chem.* **278**, 52857–52864.
86. Ruan, B., and Söll, D. (2005) The bacterial YbaK protein is a Cys-tRNA^{Pro} and Cys-tRNA^{Cys} deacylase, *J. Biol. Chem.* **280**, 25887–25891.
87. Lim, K., Tempczyk, A., Bonander, N., Toedt, J., Howard, A., Eisenstein, E., and Herzberg, O. (2003) A catalytic mechanism for D-Tyr-tRNA^{Tyr} deacylase based on the crystal structure of *Hemophilus influenzae* HI0670, *J. Biol. Chem.* **278**, 13496–13502.
88. Soutourina, J., Plateau, P., Delort, F., Peirotes, A., and Blanquet, S. (1999) Functional characterization of the D-Tyr-tRNA^{Tyr} deacylase from *Escherichia coli*, *J. Biol. Chem.* **274**, 19109–19114.
89. Ferri-Fioni, M. L., Schmitt, E., Soutourina, J., Plateau, P., Mechulam, Y., and Blanquet, S. (2001) Structure of crystalline D-Tyr-tRNA^{Tyr} deacylase. A representative of a new class of tRNA-dependent hydrolases, *J. Biol. Chem.* **276**, 47285–47290.
90. Korencic, D., Ahel, I., Schelet, J., Sacher, M., Ruan, B., Stathopoulos, C., Blum, P., Ibba, M., and Söll, D. (2004) A freestanding proofreading domain is required for protein synthesis quality control in Archaea, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10260–10265.
91. Dwivedi, S., Kruparani, S. P., and Sankaranarayanan, R. (2005) A D-amino acid editing module coupled to the translational apparatus in archaea, *Nat. Struct. Mol. Biol.* **12**, 556–557.
92. Nomanbhoy, T. K., and Schimmel, P. R. (2000) Misactivated amino acids translocate at similar rates across surface of a tRNA synthetase, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5119–5122.
93. Apostol, I., Levine, J., Lippincott, J., Leach, J., Hess, E., Glascock, C. B., Weickert, M. J., and Blackmore, R. (1997) Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in *Escherichia coli*, *J. Biol. Chem.* **272**, 28980–28988.
94. Döring, V., Mootz, H. D., Nangle, L. A., Hendrickson, T. L., de Crécy-Lagard, V., Schimmel, P., and Marlière, P. (2001) Enlarging the amino acid set of *Escherichia coli* by infiltration of the valine coding pathway, *Science* **292**, 501–504.
95. Curnow, A. W., Ibba, M., and Söll, D. (1996) tRNA-dependent asparagine formation, *Nature* **382**, 589–590.
96. Tumbula, D. L., Becker, H. D., Chang, W. Z., and Söll, D. (2000) Domain-specific recruitment of amide amino acids for protein synthesis, *Nature* **407**, 106–110.
97. Raczniak, G., Becker, H. D., Min, B., and Söll, D. (2001) A single amidotransferase forms asparaginyl-tRNA and glutaminyl-tRNA in *Chlamydia trachomatis*, *J. Biol. Chem.* **276**, 45862–45867.
98. Tumbula-Hansen, D., Feng, L., Toogood, H., Stetter, K. O., and Söll, D. (2002) Evolutionary divergence of the archaeal aspartyl-tRNA synthetases into discriminating and nondiscriminating forms, *J. Biol. Chem.* **277**, 37184–37190.
99. Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (2005) Structural basis for tRNA-dependent amidotransferase function, *Structure*. **13**, 1421–1433.
100. Ibba, M., Curnow, A. W., and Söll, D. (1997) Aminoacyl-tRNA synthesis: divergent routes to a common goal, *Trends Biochem. Sci.* **22**, 39–42.
101. Rinehart, J., Krett, B., Rubio, M. A., Alfonso, J. D., and Söll, D. (2005) *Saccharomyces cerevisiae* imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion, *Genes Dev.* **19**, 583–592.
102. Varshney, U., and RajBhandary, U. L. (1992) Role of methionine and formylation of initiator tRNA in initiation of protein synthesis in *Escherichia coli*, *J. Bacteriol.* **174**, 7819–7826.
103. Vial, L., Gomez, P., Panvert, M., Schmitt, E., Blanquet, S., and Mechulam, Y. (2003) Mitochondrial methionyl-tRNA^{Met} formyltransferase from *Saccharomyces cerevisiae*: gene disruption and tRNA substrate specificity, *Biochemistry* **42**, 932–939.
104. Guillon, J. M., Heiss, S., Soutourina, J., Mechulam, Y., Laalami, S., Grunberg-Manago, M., and Blanquet, S. (1996) Interplay of methionine tRNAs with translation elongation factor Tu and translation initiation factor 2 in *Escherichia coli*, *J. Biol. Chem.* **271**, 22321–22325.
105. Low, S. C., Hamey, J. W., and Berry, M. J. (1995) Cloning and functional characterization of human selenophosphate synthetase, an essential component of selenoprotein synthesis, *J. Biol. Chem.* **270**, 21659–21664.
106. Low, S. C., and Berry, M. J. (1996) Knowing when not to stop: selenocysteine incorporation in eukaryotes, *Trends Biochem. Sci.* **21**, 203–208.
107. Böck, A., Thanbichler, M., Rother, M., and Resh, A. (2005) in *The Aminoacyl-tRNA Synthetases* (Ibba, M., Francklyn, C. S., and Cusack, S., Eds.), pp 320–327, Landes Bioscience, Georgetown, TX.
108. Sauerwald, A., Zhu, W., Major, T. A., Roy, H., Palioura, S., Jahn, D., Whitman, W. B., Yates, J. R., III, Ibba, M., and Söll, D. (2005) RNA-dependent cysteine biosynthesis in archaea, *Science* **307**, 1969–1972.
109. Driscoll, D. M., and Copeland, P. R. (2003) Mechanism and regulation of selenoprotein synthesis, *Annu. Rev. Nutr.* **23**, 17–40.
110. Krol, A. (2002) Evolutionarily different RNA motifs and RNA-protein complexes to achieve selenoprotein synthesis, *Biochimie* **84**, 765–774.
111. Low, S. C., Grundner-Culemann, E., Hamey, J. W., and Berry, M. J. (2000) SECIS-SBP2 interactions dictate selenocysteine incorporation efficiency and selenoprotein hierarchy, *EMBO J.* **19**, 6882–6890.
112. Leibundgut, M., Frick, C., Thanbichler, M., Böck, A., and Ban, N. (2005) Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors, *EMBO J.* **24**, 11–22.

113. Buddha, M. R., Keery, K. M., and Crane, B. R. (2004) An unusual tryptophanyl tRNA synthetase interacts with nitric oxide synthase in *Deinococcus radiodurans*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15881–15886.
114. Buddha, M. R., and Crane, B. R. (2005) Structures of tryptophanyl-tRNA synthetase II from *Deinococcus radiodurans* bound to ATP and tryptophan. Insight into subunit cooperativity and domain motions linked to catalysis, *J. Biol. Chem.* **280**, 31965–31973.
115. Jakubowski, H. (1999) Misacylation of tRNA^{Lys} with noncognate amino acids by lysyl-tRNA synthetase, *Biochemistry* **38**, 8088–8093.
116. Brevet, A., Chen, J., Leveque, F., Blanquet, S., and Plateau, P. (1995) Comparison of the enzymatic properties of the two *Escherichia coli* lysyl-tRNA synthetase species, *J. Biol. Chem.* **270**, 14439–14444.
117. Desogus, G., Todone, F., Brick, P., and Onesti, S. (2000) Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction, *Biochemistry* **39**, 8418–8425.
118. Levengood, J., Ataide, S. F., Roy, H., and Ibba, M. (2004) Divergence in noncognate amino acid recognition between class I and class II lysyl-tRNA synthetases, *J. Biol. Chem.* **279**, 17707–17714.
119. Jester, B. C., Levengood, J. D., Roy, H., Ibba, M., and Devine, K. M. (2003) Nonorthologous replacement of lysyl-tRNA synthetase prevents addition of lysine analogues to the genetic code, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 14351–14356.
120. Brown, J. R., Gentry, D., Becker, J. A., Ingraham, K., Holmes, D. J., and Stanhope, M. J. (2003) Horizontal transfer of drug-resistant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens, *EMBO Rep.* **4**, 692–698.
121. Gentry, D. R., Ingraham, K. A., Stanhope, M. J., Rittenhouse, S., Jarvest, R. L., O'Hanlon, P. J., Brown, J. R., and Holmes, D. J. (2003) Variable sensitivity to bacterial methionyl-tRNA synthetase inhibitors reveals subpopulations of *Streptococcus pneumoniae* with two distinct methionyl-tRNA synthetase genes, *Antimicrob. Agents Chemother.* **47**, 1784–1789.
122. Henkin, T. M., Glass, B. L., and Grundy, F. J. (1992) Analysis of the *Bacillus subtilis* *tyrS* gene: conservation of a regulatory sequence in multiple tRNA synthetase genes, *J. Bacteriol.* **174**, 1299–1306.
123. Gendron, N., Putzer, H., and Grunberg-Manago, M. (1994) Expression of both *Bacillus subtilis* threonyl-tRNA synthetase genes is autogenously regulated, *J. Bacteriol.* **176**, 486–494.
124. Putzer, H., Gendron, N., and Grunberg-Manago, M. (1992) Coordinate expression of the two threonyl-tRNA synthetase genes in *Bacillus subtilis*: control by transcriptional antitermination involving a conserved regulatory sequence, *EMBO J.* **11**, 3117–3127.
125. Fujimura, S., Tokue, Y., and Watanabe, A. (2003) Isoleucyl-tRNA synthetase mutations in *Staphylococcus aureus* clinical isolates and in vitro selection of low-level mupirocin-resistant strains, *Antimicrob. Agents Chemother.* **47**, 3373–3374.
126. Hodgson, J. E., Cummock, S. P., Dyke, K. G., Morris, R., Sylvester, D. R., and Gross, M. S. (1994) Molecular characterization of the gene encoding high-level mupirocin resistance in *Staphylococcus aureus* J2870, *Antimicrob. Agents Chemother.* **38**, 1205–1208.
127. Thomas, D. G., Wilson, J. M., Day, M. J., and Russell, A. D. (1999) Mupirocin resistance in staphylococci: development and transfer of isoleucyl-tRNA synthetase-mediated resistance in vitro, *J. Appl. Microbiol.* **86**, 715–722.
128. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struct.*, published online Jan 13, <http://dx.doi.org/10.1146/annurev.biophys.35.101105.121507>.
129. Polycarpo, C., Ambrogelly, A., Berube, A., Winbush, S. M., McCloskey, J. A., Crain, P. F., Wood, J. L., and Söll, D. (2004) An aminoacyl-tRNA synthetase that specifically activates pyrrolysine, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12450–12454.
130. Blight, S. K., Larue, R. C., Mahapatra, A., Longstaff, D. G., Chang, E., Zhao, G., Kang, P. T., Green-Church, K. B., Chan, M. K., and Krzycki, J. A. (2004) Direct charging of tRNA^{CUA} with pyrrolysine in vitro and in vivo, *Nature* **431**, 333–335.

Direct and Rapid Cytosolic Delivery Using Cell-Penetrating Peptides Mediated by Pyrenebutyrate

Toshihide Takeuchi[†], Michie Kosuge[†], Akiko Tadokoro[†], Yukio Sugiura[†], Mayumi Nishi[‡], Mitsuhiro Kawata[‡], Naomi Sakai[§], Stefan Matile[§], and Shiroh Futaki^{†,1,*}

[†]Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan, [‡]Department of Anatomy and Neurobiology, Kyoto Prefectural University of Medicine, Kyoto 602-8566, Japan, [§]Department of Organic Chemistry, University of Geneva, Geneva, Switzerland, ¹and SORST, JST, Kawaguchi, Saitama 332-0012, Japan

Arginine-rich peptides such as oligoarginines (1, 2) and Tat (3) have been shown in numerous studies to translocate across biomembranes. These peptides have been utilized as delivery vectors that bring exogenous proteins and various molecules into cells to modulate cellular functions. There is still dispute as to exactly how these cationic peptides enter cells, but we and others have shown that the guanidinium cations on arginine residues play crucial roles in intracellular delivery (4–8). We have also recently shown that negatively charged counteranions with high hydrophobicity, such as phosphatidylglycerol and pyrenebutyrate, can exert a great influence on the translocation behavior of arginine peptides in artificial membranes (9–11). The positively charged arginine peptides electrostatically interact with the counteranions to increase their net hydrophobicity, thus, facilitating direct translocation through the lipid bilayers. The particular effectiveness of polyaromatic tails as found in pyrenebutyrate is thought to originate from ion pair stabilization of the aromatic surface as well as accelerated translocation due to their interfacial preference (10, 11).

Focusing on the concept of counteranion-mediated delivery, we first investigated whether counteranions can have effects on the translocation of arginine-rich peptides through biological membranes of live cells. Pyrenebutyrate was selected as one of the

representative counteranions, and octa-arginine (R8) as a representative peptide vector (Figure 1, panel a). Preliminary quantification assay using fluorescence-activated cell sorter (FACS) has shown the significant increase in total cellular uptake of R8 into living cells in the presence of pyrenebutyrate (10). Therefore, microscopic observation was performed to examine the effect of pyrenebutyrate on the methods of internalization of R8 peptide as well as the intracellular distribution.

HeLa cells were preincubated with pyrenebutyrate, followed by incubation with Alexa 488-labeled R8 peptide. The R8 peptide diffusely localized throughout the cytosol and more strongly in the nucleus (Figure 1, panels b and c). Surprisingly, this pattern of fluorescence was observed within a few minutes' treatment of the cells with the R8 peptide after the initial 2-min preincubation with pyrenebutyrate (Figure 1, panel b). This diffuse labeling of internalized peptide was seen in almost all the cells, and very few endosome-like punctate signals were observed in this time period. Conversely, no significant R8 peptide fluorescence was observed in the cells incubated in the absence of pyrenebutyrate (Figure 1, panel d). This suggests a critical role for this compound in peptide translocation. Internalization of the R8 peptide was also observed by the treatment of the cells with the R8 peptide prior to addition of pyrenebutyrate or by the simultaneous addi-

ABSTRACT Intracellular delivery of bioactive molecules using arginine-rich peptides, including oligoarginine and HIV-1 Tat peptides, is a recently developed technology. Here, we report a dramatic change in the methods of internalization for these peptides brought about by the presence of pyrenebutyrate, a counteranion bearing an aromatic hydrophobic moiety. In the absence of pyrenebutyrate, endocytosis plays a major role in cellular uptake. However, the addition of pyrenebutyrate results in direct membrane translocation of the peptides yielding diffuse cytosolic peptide distribution within a few minutes. Using this method, rapid and efficient cytosolic delivery of the enhanced green fluorescent protein (EGFP) was achieved in cells including rat hippocampal primary cultured neurons. Enhancement of bioactivity on the administration of an apoptosis-inducing peptide is also demonstrated. Thus, coupling arginine-rich peptides with this hydrophobic anion dramatically improved their ability to translocate cellular membranes, suggesting the great impact of this approach on exploring and controlling cell function.

*To whom correspondence should be addressed.
E-mail: futaki@scl.kyoto-u.ac.jp.

Received for review March 18, 2006
and accepted May 18, 2006

Published online June 16, 2006
10.1021/cb600127m CCC: \$33.50

© 2006 by American Chemical Society

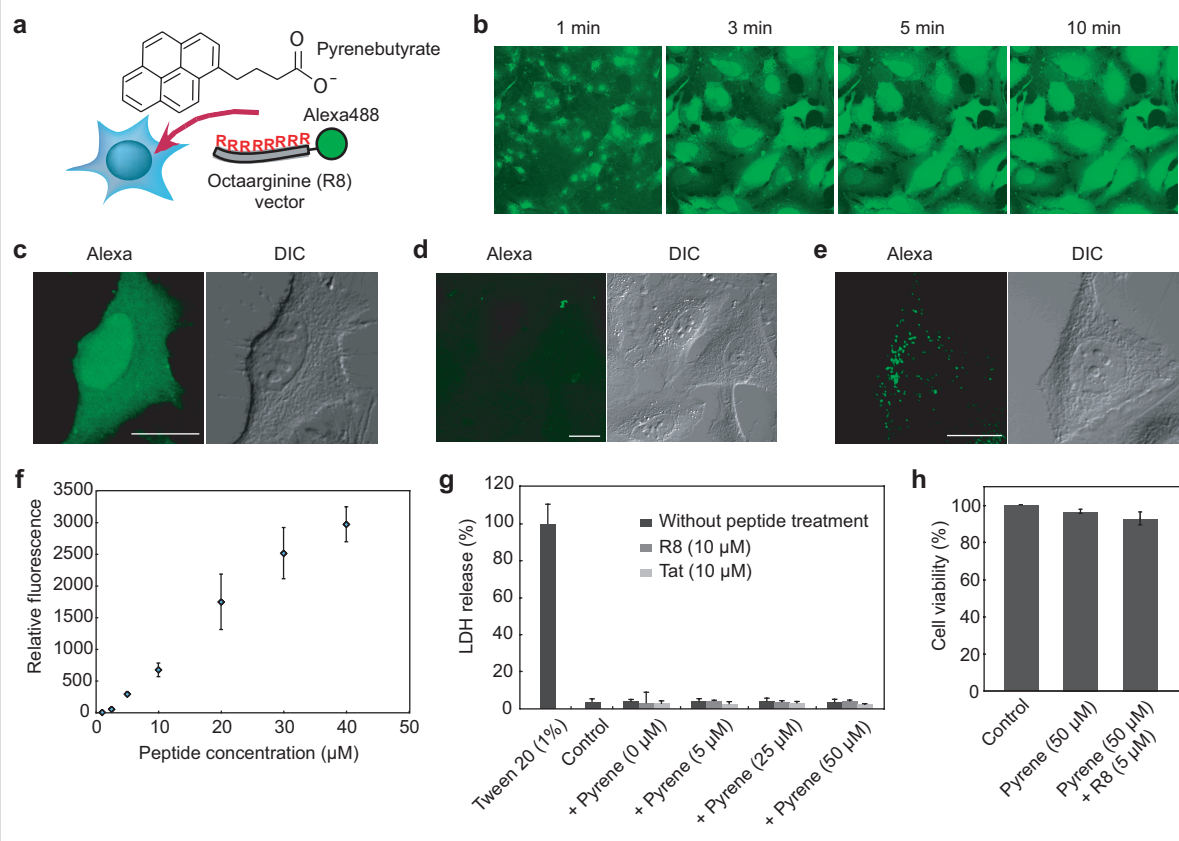


Figure 1. Direct and facile translocation of the R8 peptide into the cytosol in the presence of pyrenebutyrate without using endocytosis. **a**) A schematic representation of pyrenebutyrate (1-pyrenebutyric acid)-mediated delivery of the octaarginine (R8) vector into cells. **b**) A time-course observation of peptide internalization. HeLa cells were preincubated with pyrenebutyrate in PBS for 2 min prior to the addition of R8-Alexa in PBS (see Methods) at 37 °C and analyzed by confocal microscopy without washing the cells. **c**) Prominent cytosolic diffusion of the R8 peptide in the presence of pyrenebutyrate in PBS. HeLa cells were incubated for 4 min with R8-Alexa in the presence of pyrenebutyrate after pretreatment with pyrenebutyrate as in panel b. **d**) In the absence of pyrenebutyrate, no significant signals of R8-Alexa were observed under the same conditions. **e**) When HeLa cells were incubated with R8-Alexa (10 μ M) in a serum-containing medium for 1 h at 37 °C, punctate signals of the peptide suggestive of endocytic uptake were predominantly observed. Scale bar, 20 μ m (panels c–e). **f**) Peptide concentration-dependent increase in cell-associated fluorescence in pyrenebutyrate-treated cells. **g**) No significant effect of the pyrenebutyrate treatment on cell membrane integrity (LDH release assay). **h**) No significant pyrenebutyrate-induced cytotoxicity. Error bars represent the mean \pm standard deviation (SD) of three samples (panels f–h).

tions of the R8 peptide and pyrenebutyrate. However, the internalization efficiency of the peptide in the latter two cases seems to be slightly lower than in the former case, presumably due to the more prominent aggregation of pyrenebutyrate with R8 peptide (Supplementary Figure 1). Therefore, later experiments were conducted by incubating cells with pyrenebutyrate prior to addition of peptide. The quantity of internalized peptide in the presence of 50 μ M pyrenebutyrate increased linearly with increasing peptide concentration, and no significant threshold in the peptide concentration was observed for the cellular uptake under the given conditions (Figure 1, panel f). Damage to the plasma membranes by this pyrenebutyrate/R8 treatment would

yield nuclear staining with propidium iodide. However, no nuclear staining was observed, suggesting that this cytosolic peptide labeling was not due to membrane disruption (data not shown). The integrity of the membranes and the absence of cytotoxicity by this treatment were further confirmed by the lactate dehydrogenase (LDH) release assay and MTT [=3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide] assay (Figure 1, panels g and h, respectively). This efficient cytosolic delivery of the R8 peptide in the presence of pyrenebutyrate was similarly observed in other cell lines such as COS-7, PC12, CHO-K1, and RAW264.7 (data not shown) and was also observed with other membrane-permeable peptides derived

from HIV-1 Tat (12), HIV-1 Rev (4), and Antennapedia (Penetratin) (13) (Supplementary Figure 2), indicating the wide applicability of this concept.

The ultimate goal of this technology should be the delivery of large molecular weight proteins into the cytosol. To assess the effectiveness of this approach to cytosolic protein delivery, an enhanced green fluorescent protein (EGFP) bearing an R8 segment on its N-terminus (R8-EGFP, molecular mass \sim 30 kDa) was prepared. Surprisingly, a similarly diffuse R8-EGFP fluorescence was observed in more than 70% of the HeLa cells (Figure 2, panel a, left and middle panels). As in the case of the R8 peptide (Figure 1, panel c), R8-EGFP was distributed diffusely in the cytosol

but also strongly in the nucleus (Figure 2, panel a, middle panel), which is very similar to that of the corresponding protein intracellularly expressed by transfection (Figure 2, panel b). The efficiency of obtaining EGFP fluorescent cells was much higher compared to transfection. In contrast, when the cells were similarly incubated with EGFP lacking the R8 segment in the presence of pyrenebutyrate, no significant signals of EGFP were observed (Figure 2, panel a, right panel).

Transfection of nondividing cells such as neuronal primary cultures is difficult. We therefore determined whether the effects of pyrenebutyrate were also observed in rat hippocampal primary cultured neurons. Confocal microscopic observation of the cells revealed that these cells stained strongly for R8-EGFP (Figure 2, panel c). Interestingly, it was often observed that neurites were as effectively stained as the cell bodies. At this concentration, internalization of the R8-EGFP was observed for about 40% of the hippocampal cells. When the concentration of R8-EGFP was raised to 20 μM , this increased to >80% (Figure 2, panel d). No significant toxicity was observed under these treatments.

It is important to determine whether these pyrenebutyrate effects can be expanded to show a biological effect due to increased translocation of a delivery vector; the results of quantification of cell-associated peptides and the eventual biological effects are not universally parallel (14). We therefore applied this counteranion-mediated delivery to investigate whether this compound would enhance the effects of an apoptosis-inducing peptide (pro-apoptotic domain peptide, PAD) (15) that mediates its effect in the cytosol. PAD is an amphiphilic basic peptide including 14 residues of *D*-amino acids, *D*-(KLAKLAK)₂. It has been reported that cytosolic delivery of the PAD peptide leads to mitochondrial membrane disruption and eventual apoptosis (14–16). On the other

hand, the peptide is nontoxic without being internalized into cells.

A dramatic effect of the addition of pyrenebutyrate was confirmed by the observation of mitochondrial depolarization using a mitochondrial membrane-potential-sensitive dye, JC-1. This marker exhibits a potential-dependent accumulation in mitochondria (J-aggregates), and they are observed as red fluorescence. When the mitochondrial membrane is depolarized, the fluorescence changes to green (17). When HeLa cells were treated with R8-PAD in the absence of pyrenebutyrate, red fluorescent aggregation was predominantly observed (Figure 3, panel a, left and center panels). In contrast, when the cells were pretreated with pyrenebutyrate, green fluorescence was observed in substantially all of the cells (Figure 3, panel a, right panel) only after this treatment with R8-PAD. This

suggests that effective mitochondrial depolarization was accomplished by the accelerated internalization of R8-PAD in the presence of pyrenebutyrate. The biological effect was also assessed on the basis of cell death induction mediated by the intracellularly delivered apoptosis-inducing peptide (R8-PAD) using MTT assays. Cell viability was approximately 50% (Figure 3, panel b). However, preincubation of cells in pyrenebutyrate (10 min) before addition of the R8-PAD peptide reduced cell viability, and the enhancement of cell death was dependent on the concentration of pyrenebutyrate. The highest activity (<20% cell viability) was observed in cells treated in the presence of 50 μM pyrenebutyrate. These results suggest that not only cellular uptake but also the biological effect can be enhanced by this counteranion-mediated delivery.

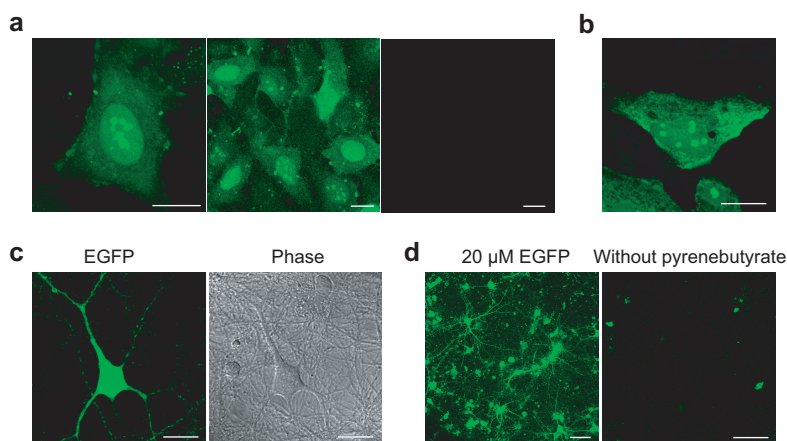


Figure 2. Efficient cytosolic delivery of EGFP protein attained. **a)** Internalization of the enhanced green fluorescent protein (EGFP) bearing the R8 peptide (R8-EGFP) into HeLa cells. Cells were treated with R8-EGFP (10 μM) in PBS in the presence of pyrenebutyrate as described in Figure 1, panel c (left and middle panels). EGFP without the R8 segment (10 μM) showed very little internalization even in the presence of 50 μM pyrenebutyrate (right panel). **b)** A similar cellular localization pattern as in panel a was observed in cells transfected with the plasmid coding R8-EGFP (pR8-EGFP). **c)** Internalization of R8-EGFP into neurites including cell bodies of hippocampal primary cultured neuronal cells. The cells were treated as in panel a. **d)** Diffuse labeling was observed in more than 80% of the cells within 4 min after treatment with pyrenebutyrate (50 μM) and R8-EGFP (20 μM) in PBS (left panel). No significant internalization of R8-EGFP was observed for the cells similarly treated in the absence of pyrenebutyrate (right panel). Scale bars, 20 μm (panel a–c) and 50 μm (panel d).

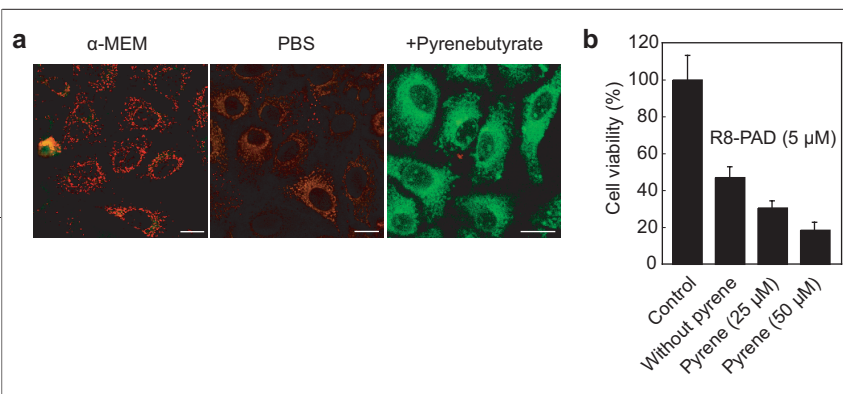


Figure 3. Reinforcement of cytosolic activity by the proapoptotic domain (PAD) peptide intracellularly delivered in the presence of

pyrenebutyrate. **a)** Accelerated cancellation of the mitochondrial membrane potential by the R8-PAD peptide in the presence of pyrenebutyrate. As controls, HeLa cells were treated with R8-PAD (5 μM , 4 min) in serum containing α -MEM (left panel) or PBS (middle panel) in the absence of pyrenebutyrate. Alternatively, the cells were treated with pyrenebutyrate (2 min) prior to incubation with R8-PAD (4 min) (final concentration of pyrenebutyrate and R8-PAD, 50 and 5 μM , respectively) (right panel). The effect of the PAD peptide on the respective cells was then analyzed by confocal microscopic observation after washing the cells twice with PBS and incubating with a mitochondrial membrane potential-sensitive marker, JC-1 (5 μM) at 37 $^{\circ}\text{C}$ for 30 min. Scale bar, 20 μm . **b)** Enhanced cytosolic activity of the PAD peptide by pyrenebutyrate. Error bars represent the mean \pm SD of six samples.

As for mechanisms that enable this efficient internalization, a direct membrane translocation without a requirement for endocytosis is postulated (Figure 1, panels b and c). As support of this, a very similar cellular distribution, as well as increase in cellular uptake, was observed in the presence of pyrenebutyrate even when the cells were treated with R8-Alexa at 4 $^{\circ}\text{C}$, where endocytosis does not work (Supplementary Figure 3). This yields a marked difference from observations of cells treated with 10 μM R8 peptide at 37 $^{\circ}\text{C}$ in the absence of pyrenebutyrate (Figure 1, panel e); the punctate signals observed in the perinuclear area suggested that the majority of the peptide in this case was taken up and was trapped in endocytic vesicles (14, 18). In addition, the importance of the membrane potential is proposed as a driving force in this counteranion-mediated translocation (Supplementary Figure 4), as was observed for the internalization of arginine-rich peptides into liposomes (19) and suspension cells (20) in the absence of pyrenebutyrate.

Interestingly, the fluorescence microscopic images of the pyrenebutyrate indicated that the majority of the pyrenebutyrate stayed in the membrane compartments in the cells, especially those in the perinuclear region (Figure 4). The intracellular distribution of pyrenebutyrate was observed after only 2 min incubation with the cells; this molecule *freely* diffuses through the plasma mem-

brane and cytosol to reach the membrane compartments. No significant differences in cellular distribution were observed by the treatment of the cells with the R8 peptide. Therefore, pyrenebutyrate acts “like a translocation catalyst” to accelerate the translocation of the peptides. In addition, once the R8 peptide was delivered into cells, pyrenebutyrate could be immediately removed from cells simply by washing in PBS, and this would prevent possible damage to the cells that may be induced by a prolonged incubation with pyrenebutyrate (Figure 4).

As shown above, the pyrenebutyrate-mediated translocation of arginine-rich peptides achieves direct and efficient delivery of bioactive proteins bearing arginine vectors into the cytosol. Importantly, this increase in translocation was shown to manifest in a biological response. As there was no toxicity associated with this proce-

dure, these results should open new opportunities to the utilization of this approach to the studies in cell biology. This approach can also be applicable to the introduction of chemically modified proteins, *e.g.*, with fluorescent moieties and cross-linking agents into cells, as well as natural bioactive proteins. High-throughput screening systems may be established by using this system. Great possibility in the interface of chemistry and biology will come up to explore and control cell function.

This approach has a limitation in that it is not applicable in the presence of a medium or serum; the competition with various ionic species in the medium would hamper the interaction of the arginine vectors with pyrenebutyrate. Screening efforts to find serum-insensitive counteranion activators are ongoing. However, the high efficiency of this approach would outweigh this limitation. A few minutes' treatment of the cells with pyrenebutyrate/peptides in PBS is sufficient for enhancing translocation, and the incubation medium could then be replaced with a more complete medium for longer incubations. For therapeutic applications, an ‘*ex-vivo*’-like approach may be employable, where isolated cells are treated with target proteins and then incorporated into the body again.

Overall, considering these outstanding features, the pyrenebutyrate-mediated delivery using arginine-rich peptides will provide a new concept in membrane translocation.

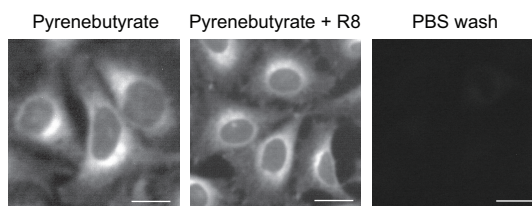


Figure 4. PBS wash of the cells yields immediate removal of pyrenebutyrate from cells. HeLa cells were treated with pyrenebutyrate for 2 min (left panel), followed by nonfluorescently labeled R8 peptide (final concentration of pyrenebutyrate and R8, 50 and 10 μM , respectively) for 4 min, and immediately analyzed by fluorescence microscopy without washing (middle panel) or, alternatively, washed 3 \times with PBS prior to microscopic observation (right panel). Scale bar, 20 μm .

METHODS

Peptides and Proteins. All the peptides used in this report were chemically synthesized by Fmoc (9-fluorenylmethyloxycarbonyl) solid-phase peptide synthesis as reported (16). Actual sequences of the synthesized peptides: R8-Alexa, RRRRRRRRGC(Alexa)-amide; Tat-Alexa, GRKKRRQR-RRPPQC(Alexa)-amide; Rev-Alexa, TRQARRNR-RRRWREQRGC(Alexa)-amide; Penetratin-Alexa, RQKIWFQNRMRKWKKG(Alexa)-amide; R8-PAD (16), RRRRRRRRGG-D-(KLAKLAK)₂-amide. The EGFP protein bearing (His)₆ and (Arg)₈ segments and a thrombin cleavage site on its N-terminus (R8-EGFP) and the EGFP bearing (His)₆ and thrombin cleavage site (lacking R8) were prepared as previously reported (16). Details in preparation of the peptides and protein were described (see Supporting Information).

Confocal Microscopy. 2×10^5 cells were plated into 35-mm glass-bottomed dishes (Iwaki) and cultured for 48 h. After removing the medium, the cells were washed twice with PBS. The cells were first incubated with pyrenebutyrate (1-pyrenebutyric acid) (67 μ M) in PBS for 2 min, and then the peptide solution in PBS (40 μ M) was added to yield the final concentration of pyrenebutyrate and peptides (50 and 10 μ M, respectively). A 20 μ M peptide solution in PBS was employed to obtain a final peptide concentration of 5 μ M. After incubation for 4 min, the cells were washed with PBS ($\times 5$). Distribution of the fluorescently labeled peptides was analyzed *without fixing* using a confocal scanning laser microscope (Olympus FV300) equipped with a 40 \times objective lens. For the observation of hippocampal primary cultures, a Zeiss LSM510 equipped with a 63 \times objective lens was employed. For 4 $^{\circ}$ C experiments, cells were preincubated in a refrigerator (4 $^{\circ}$ C) for 1 h. Washing and incubation of the cells were then conducted using cold PBS and the 4 $^{\circ}$ C-refrigerator, respectively, prior to observation of the cells in cold PBS. For the experiments with the mitochondrial membrane potential probe JC-1, the cells were preincubated with pyrenebutyrate (2 min) prior to R8-PAD treatment (4 min) (final concentration of pyrenebutyrate and R8-PAD, 50 and 5 μ M, respectively). After washing twice with PBS and incubation with JC-1 (5 μ M) for 30 min in α -MEM containing 10% (v/v) calf serum, the cells were analyzed with confocal microscopy.

Fluorescence Microscopy. The cells were treated as stated above using nonfluorescently labeled R8 peptide instead of R8-Alexa. Distribution of pyrenebutyrate was analyzed by a fluorescence microscope (Olympus IX-50) equipped with a 20 \times objective lens using Hg lamp as a light source (excitation, 330–385 nm; emission, >420 nm).

Quantitation of Peptide Uptake. HeLa cells were preincubated with pyrenebutyrate in PBS for 5 min prior to the addition of increasing concentrations of R8-Alexa in PBS at 37 $^{\circ}$ C; final concentration of pyrenebutyrate was fixed at 50 μ M. After incubation for 15 min, the internalized peptide was analyzed by FACS.

LDH Release Assay. HeLa cells were incubated with pyrenebutyrate in PBS for 5 min and then with peptide in the presence of pyrenebutyrate for

30 min. LDH release was calculated from the percentage of the released LDH in pyrenebutyrate-treated cells relative to that of cells treated with 1% (v/v) Tween 20 under the same conditions.

MTT Assay. HeLa cells were treated with 50 μ M pyrenebutyrate for 10 min followed by 5 μ M R8 or Tat in the continued presence of pyrenebutyrate for 20 min in PBS. Following 2 \times PBS washes, the cells were further incubated in the absence of pyrenebutyrate and peptides for 24 h in α -MEM containing 10% (v/v) calf serum. Cell viability was then analyzed by MTT assay. HeLa cells were pretreated with pyrenebutyrate for 10 min, followed by incubation with R8-PAD (5 μ M) for 20 min in PBS. The cells were washed with PBS and incubated in α -MEM containing 10% (v/v) calf serum for 24 h. Cell viability was then analyzed by the MTT assay.

Acknowledgment: This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Institute of Sustainability Science, Kyoto University (SF), and the Swiss NSF (S.M.). T.T. is grateful for a JSPS Research Fellowship for Young Scientists.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Futaki, S. (2005) Membrane-permeable arginine-rich peptides and the translocation mechanisms, *Adv. Drug Delivery Rev.* 57, 547–558.
- Rothbard, J. B., Jessop, T. C., and Wender, P. A. (2005) Adaptive translocation: the role of hydrogen bonding and membrane potential in the uptake of guanidinium-rich transporters into cells, *Adv. Drug Delivery Rev.* 57, 495–504.
- Wadia, J. S., and Dowdy, S. F. (2005) Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer, *Adv. Drug Delivery Rev.* 57, 579–596.
- Futaki, S., Suzuki, T., Ohashi, W., Yagami, T., Tanaka, S., Ueda, K., and Sugiura, Y. (2001) Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery, *J. Biol. Chem.* 276, 5836–5840.
- Wender, P. A., Mitchell, D. J., Pattabiraman, K., Pelkey, E. T., Steinman, L., Rothbard, J. B. (2000) The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters, *Proc. Natl. Acad. Sci. U.S.A.* 97, 13003–13008.
- Wright, L. R., Rothbard, J. B., and Wender, P. A. (2003) Guanidinium rich peptide transporters and drug delivery, *Curr. Protein Pept. Sci.* 4, 105–124.
- Umezawa, N., Gelman, M. A., Haigis, M. C., Raines, R. T., and Gellman, S. H. (2002) Translocation of a β -peptide across cell membranes, *J. Am. Chem. Soc.* 124, 368–369.
- Rueping, M., Mahajan, Y., Sauer, M., and Seebach, D. (2002) Cellular uptake studies with beta-peptides, *ChemBioChem* 3, 257–259.
- Sakai, N., Takeuchi, T., Futaki, S., and Matile, S. (2005) Direct observation of anion-mediated translocation of fluorescent oligoarginine carriers into and across bulk liquid and anionic bilayer membranes, *ChemBioChem* 6, 114–122.
- Perret, F., Nishihara, M., Takeuchi, T., Futaki, S., Lazar, A. N., Coleman, A. W., Sakai, N., and Matile, S. (2005) Anionic fullerenes, calixarenes, coronenes, and pyrenes as activators of oligo/polyarginines in model membranes and live cells, *J. Am. Chem. Soc.* 127, 1114–1115.
- Nishihara, M., Perret, F., Takeuchi, T., Futaki, S., Lazar, A. N., Coleman, A. W., Sakai, N., and Matile, S. (2005) Arginine magic with new counterions up the sleeve, *Org. Biomol. Chem.* 3, 1659–1669.
- Vivès, E., Brodin, P., and Lebleu, B. (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus, *J. Biol. Chem.* 272, 16010–16017.
- Derosi, D., Joliot, A. H., Chassaing, G., and Prochiantz, A. (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes, *J. Biol. Chem.* 269, 10444–10450.
- Nakase, I., Niwa, M., Takeuchi, T., Sonomura, K., Kawabata, N., Koike, Y., Takehashi, M., Tanaka, S., Ueda, K., Simpson, J. C., Jones, A. T., Sugiura, Y., and Futaki, S. (2004) Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement, *Mol. Ther.* 10, 1011–1022.
- Ellerby, H. M., Arap, W., Ellerby, L. M., Kain, R., Andrusiak, R., Rio, G. D., Krajewski, S., Lombardo, C. R., Rao, R., Ruoslahti, E., Bredesen, D. E., and Pasqualini, R. (1999) Anti-cancer activity of targeted proapoptotic peptides, *Nat. Med.* 5, 1032–1038.
- Futaki, S., Niwa, M., Nakase, I., Tadokoro, A., Zhang, Y., Nagaoka, M., Wakako, N., and Sugiura, Y. (2004) Arginine carrier peptide bearing Ni(II) chelator to promote cellular uptake of histidine-tagged proteins, *Bioconjugate Chem.* 15, 475–481.
- Minamikawa, T., Williams, D. A., Bowser, D. N., and Nagley, P. (1999) Mitochondrial permeability transition and swelling can occur reversibly without inducing cell death in intact human cells, *Exp. Cell Res.* 246, 26–37.
- Richard, J. P., Melikov, K., Vivès, E., Ramos, C., Verbeure, B., Gait, M. J., Chemomordik, L. V., and Lebleu, B. (2003) Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake, *J. Biol. Chem.* 278, 585–590.
- Terrone, D., Sang, S. L., Roudaia, L., and Silvius, J. R. (2003) Penetratin and related cell-penetrating cationic peptides can translocate across lipid bilayers in the presence of a transbilayer potential, *Biochemistry* 42, 13787–13799.
- Rothbard, J. B., Jessop, T. C., Lewis, R. S., Murray, B. A., and Wender, P. A. (2004) Role of membrane potential and hydrogen bonding in the mechanism of translocation of guanidinium-rich peptides into cells, *J. Am. Chem. Soc.* 126, 9506–9507.

Mechanistic Basis for Therapeutic Targeting of the Mitochondrial F_1F_0 -ATPase

Kathryn M. Johnson^{†,*}, Joanne Cleary[†], Carol A. Fierke[†], Anthony W. Opipari, Jr.[§], and Gary D. Glick^{†,*}

Department of [†]Chemistry, ^{*}Graduate Program in Immunology, and [§]Obstetrics & Gynecology, University of Michigan, Ann Arbor, Michigan 48109-1055

ABSTRACT Altered cellular bioenergetics are implicated in many disease processes, and modulating the F_1F_0 -ATPase, the enzyme responsible for producing the majority of ATP in eukaryotic cells, has been proposed to have therapeutic utility. Bz-423 is a 1,4-benzodiazepine that binds to the oligomycin sensitivity-conferring protein subunit of the mitochondrial F_1F_0 -ATPase and inhibits the enzyme. In response to Bz-423, cells moderately decrease ATP synthesis and significantly increase superoxide, resulting in redox-regulated apoptosis. Administering Bz-423 to autoimmune mice leads to apoptosis of pathogenic cells and potent attenuation of disease progression. To determine if a mechanism of action distinguishes Bz-423 from toxic F_1F_0 -ATPase inhibitors like oligomycin, we studied how both compounds inhibit the enzyme. Oligomycin is a high-affinity mixed inhibitor, displaying time-dependent inhibition, resulting in severe depletion of ATP. In contrast, Bz-423 is an allosteric inhibitor with lower affinity that rapidly dissociates from the enzyme. Our data support a model in which the interplay of these features underlies the favorable properties of Bz-423. They also represent key criteria for the development of therapeutic F_1F_0 -ATPase inhibitors, which should have utility across a range of areas.

Inhibiting the mitochondrial F_1F_0 -ATPase in respiring cells has two direct consequences: ATP concentration decreases and mitochondria transition from active to resting respiration (state 3 to 4) (1). Each of these responses affects a series of different physiologic processes. During a state 3 to 4 transition, the proton motive force within mitochondria becomes sufficiently high that intermediates capable of one electron reduction reactions have extended half-lives (2). These conditions favor the production of O_2^- , which can function as a signaling molecule initiating apoptosis. In cells that rely on oxidative phosphorylation, ATP levels decrease in proportion to the extent that the F_1F_0 -ATPase is inhibited, and if they drop too much, the metabolic consequences are severe and necrosis ensues (3).

Bz-423 is a cytotoxic 1,4-benzodiazepine (2, 4) that binds to the oligomycin sensitivity-conferring protein (OSCP) subunit of the F_1F_0 -ATPase and inhibits the enzyme (5) resulting in redox-regulated apoptosis (2). While how OSCP binding inhibits the enzyme is not clear, Bz-423 likely blocks and/or induces a conformational change required for catalysis. Administering Bz-423 to lupus-prone mice causes lineage-specific apoptosis in splenocytes and significant disease amelioration (2, 4). Unlike current lupus drugs, Bz-423 suppresses autoimmune disease without affecting normal immune function.

Prior studies have identified abnormalities in lupus lymphocytes, which, given the cellular mechanism of Bz-423-induced apoptosis, can account for its selectivity *in vivo* for autoimmune *versus* normal

lymphocytes (6). However, these findings cannot explain why, unlike most F_1F_0 -ATPase inhibitors which are poisonous (*e.g.*, oligomycin) (7), Bz-423 is therapeutic and not toxic to tissues that utilize aerobic ATP production. Therefore, we investigated if the differences between oligomycin and Bz-423 could be distinguished based on their effects on F_1F_0 -ATPase activity (8).

Bz-423 decreases the apparent V_{max} (Figure 1, panel a) and K_m (Figure 1, panel b) of ATP hydrolysis, while the apparent V_{max}/K_m (Figure 1, panel c) remains unchanged; Bz-423 only inhibits an enzyme–substrate (ES) complex (*i.e.*, it is an uncompetitive inhibitor (8–10)). The concentration range of MgATP examined spans that for transitioning from a slow rate of hydrolysis with only one catalytic site filled to a physiologic rate (11). These data suggest that Bz-423 mainly inhibits ATP hydrolysis when two or three catalytic sites are occupied with nucleotide (12–14). A three-dimensional fit of the dependence of kinetic data confirmed that the uncompetitive model is the best fit of the data (Table 1, top). In contrast, oligomycin is a noncompetitive inhibitor of ATP hydrolysis; it inhibits both E and ES (Table 1, top; Figure 1, panels d–f).

For ATP synthesis, Bz-423 decreases the apparent V_{max} (Figure 2, panel a), increases the apparent K_m (Figure 2, panel b), and decreases V_{max}/K_m (Figure 2, panel c), consistent with mixed inhibition. To fit the dependence of V_{max}/K_m on Bz-423 concentration, it was necessary to include a cooperativity factor accounting for binding to E (n_E) with a stoichiometry >1 , which generated a competitive inhibition constant ($K_{i(E)}$) of

*To whom correspondence should be addressed.
E-mail: gglick@umich.edu.

Received for review March 31, 2006 and accepted May 16, 2006

Published online June 9, 2006

10.1021/cb600143j CCC: \$33.50

© 2006 by American Chemical Society

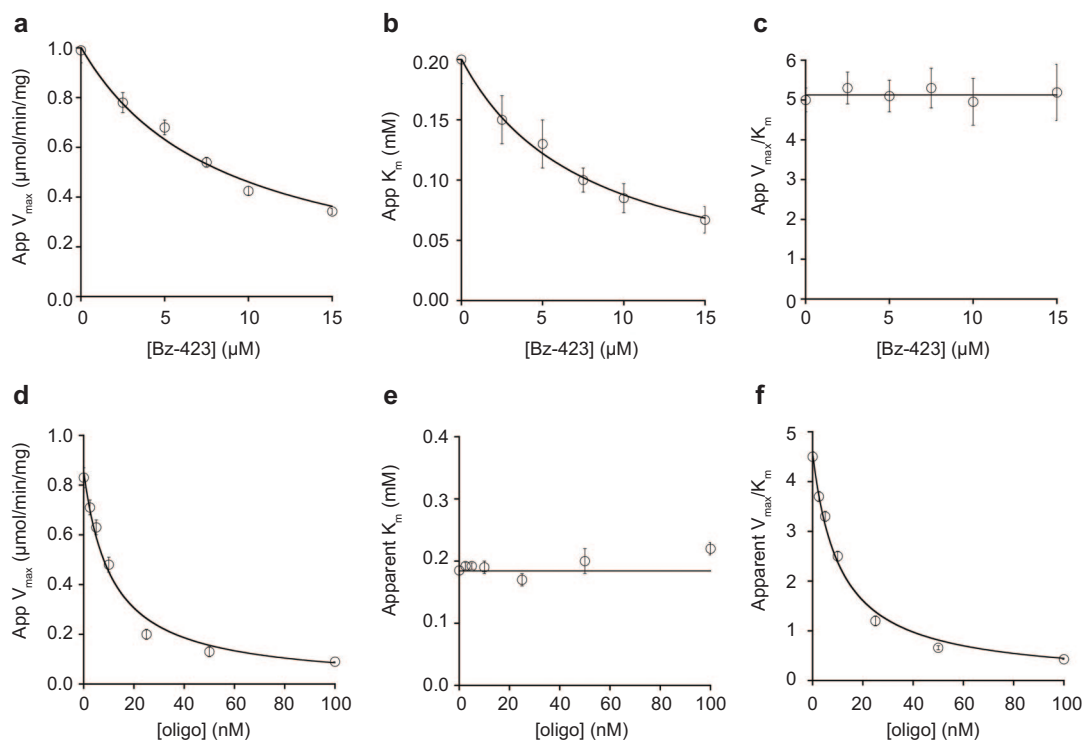


Figure 1. Hydrolysis inhibition kinetics measured using bovine submitochondrial particles (SMPs). Bz-423 decreases both a) the apparent V_{\max} b) and K_m of ATP hydrolysis, with no change in c) V_{\max}/K_m . In contrast, oligomycin decreases d) the apparent V_{\max} f) and V_{\max}/K_m , with e) no change in K_m . The plots of the apparent kinetic parameters in panels a, b, d, and f were fit using eq 1 as described in Methods. The data in this figure are taken from ref 5.

$7.2 \pm 0.3 \mu\text{M}$, with n_E of 2.4 ± 0.3 . The dependence of V_{\max} on Bz-423 concentration can be reasonably described assuming no cooperativity for binding ES (*i.e.*, $n_{ES} = 1$; Figure 2, panel a); including a cooperativity factor does not significantly alter the kinetic parameters and leads to an uncompetitive inhibition constant ($K_{i(ES)}$) of $11 \pm 2 \mu\text{M}$, with n_{ES} of 1.4 ± 0.4 . A three-dimensional fit of the dependence of initial velocity on

concentrations of both MgADP and Bz-423 using the cooperativity factors determined above closely models the raw data (Figure 2, panel d) and affords $K_{i(E)} = 6.9 \pm 0.3 \mu\text{M}$ and an uncompetitive inhibition constant $K_{i(ES)} = 10 \pm 1 \mu\text{M}$ (Table 1, bottom).

Since there is no cooperativity under hydrolytic conditions, during ATP synthesis, it is likely that the proton motive force alters the conformation of the enzyme such that in the

free state, particularly at low MgATP concentration, binding sites not present during hydrolysis are revealed (15). The location of these additional binding sites is not clear, since only one copy of the OSCP is present in each F_1F_0 -ATPase (1). However, the F_1F_0 -ATPase contains several nonspecific recognition sites for hydrophobic molecules, and it is possible that binding to these sites accounts for the increased stoichiometry (16). Since

TABLE 1. Kinetic parameters^a

Inhibitor	Mechanism of inhibition	V_{\max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	K_m (mM)	n_E^d	$K_{i(E)}^e$	n_{ES}^f	$K_{i(ES)}^g$
Bz-423 ^b	uncompetitive	0.72 ± 0.01	0.30 ± 0.01	—	—	1.0 ± 0.1	11.0 ± 0.4
oligomycin ^b	noncompetitive	0.77 ± 0.03	0.12 ± 0.02	1.1 ± 0.3	11.0 ± 3.0	1.2 ± 0.1	13 ± 1
Bz-423 ^c	mixed	0.088 ± 0.005	0.08 ± 0.01	2.4 ± 0.3	6.9 ± 0.3	1.4 ± 0.4	10 ± 1
oligomycin ^c	mixed	0.089 ± 0.003	0.12 ± 0.01	3.6 ± 0.5	10.5 ± 0.9	1.6 ± 0.3	12 ± 1

^a K_i values are μM for Bz-423 and nM for oligomycin. Parameters (\pm SE) were obtained from three-dimensional fits of velocity, [S], and [I], as described in the legends of Figures 1 and 2. ^bHydrolysis. ^cSynthesis. ^d n_E = cooperativity factor for competitive inhibition constant. ^e $K_{i(E)}$ = inhibition constant describing competitive portion of the mixed inhibition. ^f n_{ES} = cooperativity factor for uncompetitive inhibition constant. ^g $K_{i(ES)}$ = inhibition constant describing competitive portion of the mixed inhibition.

While drug discovery typically focuses on competitive antagonists, allosteric inhibitors with properties like Bz-423 can afford a superior therapeutic profile.

1,4-benzodiazepines can also bind to ATP recognition sites within proteins (17), the increased stoichiometry may result from binding of Bz-423 to the catalytic sites in F_1 .

Analysis of the ATP synthesis inhibition kinetics showed that oligomycin is a mixed inhibitor and, like Bz-423, binds cooperatively to the enzyme (Table 1, bottom). The binding site for oligomycin is in the transmembrane region of the c-subunits, which compose part of the proton channel in F_0 (1, 16). Within this region of the enzyme, macrolides such as apoptolidin bind, and the recognition sites for these compounds overlap (18). Hence, the increased stoichi-

ometry observed compared to hydrolysis may result from binding to these overlapping sites.

Inhibition of Na,K-ATPases with oligomycin is time-dependent and results from a slow rate of dissociation from the enzyme (0.05 s^{-1}) (19). Since time-dependent inhibition can have profound effects on the function of the F_1F_0 -ATPase, we examined the off-rate of Bz-423 and oligomycin during ATP synthesis. Submitochondrial particles (SMPs) were incubated with inhibitor above the K_i (14 μM Bz-423 or 35 nM oligomycin), which reduces ATP synthesis activity of the enzyme to 27% (Bz-423) and 14% (oligo-

mycin) of that observed in the absence of inhibitor. Inhibited SMP suspensions were diluted into buffer, and activity was measured over time. Enzymatic activity of the SMPs incubated with Bz-423 had recovered to 90% of control at 10 s, suggesting rapid dissociation (Figure 3). In contrast, the activity of the enzyme incubated with oligomycin was only 54% of the uninhibited activity at 10 s, and increased for several minutes after dilution (Figure 3). The time dependence of the recovery of activity fit to a single exponential, yielding an apparent dissociation rate constant for Bz-423 $>0.3\text{ s}^{-1}$, which is fast. However, dissociation of oligomycin is complex, and a portion of the inhibitor dissociates at least 10-fold slower, consistent with a tight-binding, slow-dissociating inhibitor (20). Presumably, the constraints of a high-affinity inhibitor dissociating from the membrane-embedded proton channel present a significant barrier for dissociation resulting in a slow off-rate.

Time-dependent inhibition explains why even low oligomycin concentrations lead to large decreases in ATP levels (5, 21). Inhibiting the F_1F_0 -ATPase decreases the rate of ATP synthesis, which serves to increase levels of mitochondrial ADP (22). While increased substrate partly offsets competitive inhibition by oligomycin ($K_i \ll K_m$), it also potentiates the uncompetitive arm of the mechanism (8). Hence, the slow off-rate, mechanism, and affinity of oligomycin combine to potentially block all states of the F_1F_0 -ATPase. Although oligomycin also causes mitochondria to undergo a state 3 to 4 transition like Bz-423, these effects are outweighed by the large decrease in cellular ATP levels (3).

The recognition site for Bz-423 is the OSCP, which is solvent-exposed (23). The K_i values for the competitive and uncompetitive components of inhibition of ATP synthesis by Bz-423 are micromolar and above K_m . Estimates of *in vivo* mitochondrial ADP concentration range from high micro-

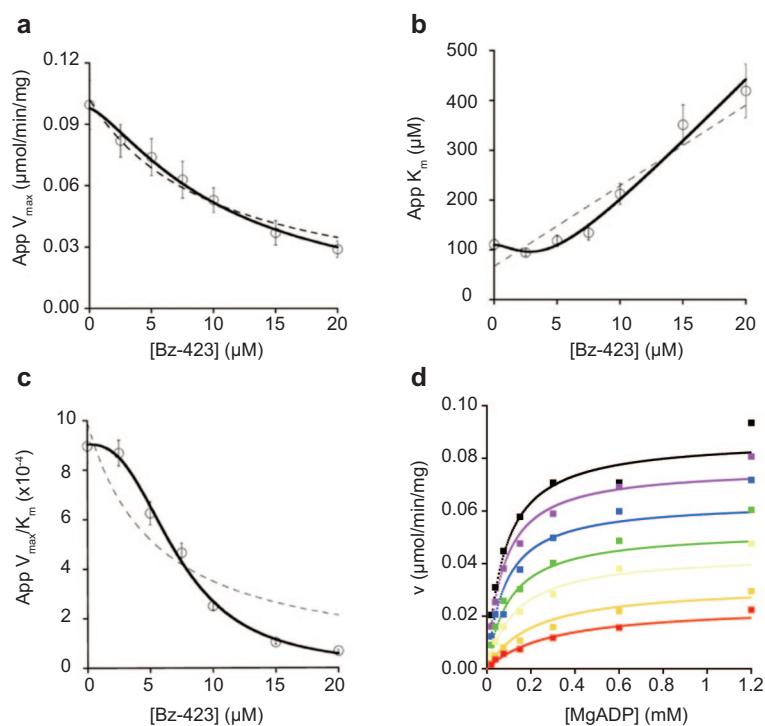


Figure 2. Bz-423 synthesis inhibition kinetics measured using SMPs. Under conditions of varying [MgADP], a) Bz-423 decreases the apparent V_{max} , b) increases the apparent K_m , and c) decreases the apparent V_{max}/K_m of ATP synthesis. The plots of the kinetic parameters in panels a, b, and c were initially fit using eqs 2, 3, and 4, respectively, with n_E and n_{ES} set to 1 (dashed lines). The plots were re-fit using these equations allowing n_E and n_{ES} to vary (solid lines), yielding $n_{ES} = 1.4 \pm 0.4$ and $n_E = 2.4 \pm 0.3$. Similar results were obtained under conditions of varying $[P_i]$. d) The mixed model described by eq 7 in Methods, with n_E and n_{ES} set to 2.4 and 1.4, respectively, was then used to generate a series of theoretical curves of velocity (v) versus [MgADP] for each [Bz-423] (μM ; 0, black; 2.5, purple; 5, blue; 7.5, green; 10, yellow; 15, orange; and 20, red).

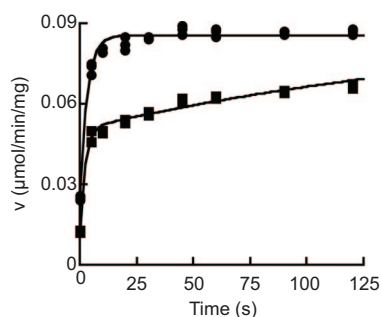


Figure 3. Off-rate kinetics. Bz-423 (●) and oligomycin (■). Control SMPs were incubated with DMSO vehicle control and diluted in a similar fashion to determine the baseline rate of synthesis activity. Control experiments adding Bz-423 and oligomycin to SMPs at the concentrations equivalent to those produced after dilution in off-rate studies (1.4 μM and 3.5 nM, respectively) reduced the activity of the F_1F_0 -ATPase by 10% and 15%, respectively. These data were fit with either a single-exponential equation (Bz-423) or a double-exponential equation (oligomycin) using the measured activity in the diluted concentrations of inhibitor as the endpoint of the reaction.

millimolar levels (25), which is sufficient to prevent competitive inhibition of ATP synthesis by Bz-423. Since these ADP concentrations are sufficient to saturate all three catalytic sites of the F_1F_0 -ATPase (11), uncompetitive inhibition should be optimal and likely accounts for the observed inhibition (8). The therapeutic dosage of Bz-423 yields peak plasma levels similar to $K_{i(\text{ES})}$ and to the EC_{50} for Bz-423-induced apoptosis (2). Although this concentration of Bz-423 has little effect on ATP concentration (5), it is sufficient to trigger a respiratory transition that signals apoptosis. Thus, the interplay of affinity, kinetics, mechanism, and pharmacokinetics prevents the metabolic collapse and resultant toxicities observed with oligomycin.

While drug discovery typically focuses on competitive antagonists, allosteric inhibitors with properties like Bz-423 can afford a superior therapeutic profile (8, 9). This point is exemplified by attempts to develop

N-methyl-D-aspartate (NMDA) receptor antagonists. Initial efforts focused on high-affinity, competitive inhibitors; however, these compounds are toxic due to their slow dissociation rates and consequent accumulation in the receptor. Lower affinity competitive antagonists are not active, since they are readily displaced by endogenous ligand (glutamate). Memantine is an uncompetitive inhibitor with micromolar affinity that rapidly dissociates from the receptor (26). Since memantine is uncompetitive and does not compete with glutamate, high affinity is not required for activity, and like Bz-423 and the F_1F_0 -ATPase, memantine modulates the receptor in a way that preserves normal function.

ATPases are being increasingly recognized as therapeutic targets (24). However, the therapeutic potential of the F_1F_0 -ATPase has yet to be realized, since it has generally been disregarded due to the toxicity of molecules like oligomycin (7). Our data reveal how binding site, kinetics, and mechanism of inhibition combine to afford different biochemical changes in response to F_1F_0 -ATPase inhibition, ultimately resulting in differential efficacy and toxicity (1, 4), and serve as criteria for selecting and optimizing inhibitors of this enzyme with therapeutic potential. As altered cellular bioenergetics are increasingly implicated in the progression of many disease processes (27–30), predictably modulating the F_1F_0 -ATPase is likely to have significant clinical utility.

METHODS

Reagents. ADP, AMP, and P^1, P^5 -di(adenosine)-pentaphosphate ($A_{P_5}A$) were purchased from EMD Biosciences. ATP, NADH, $NADP^+$, pyruvate kinase (PK), lactate dehydrogenase (LDH), hexokinase (HK), glucose-6-phosphate dehydrogenase (G6PDH), and phosphoenolpyruvate (PEP) were from Roche Applied Science. Cow hearts were supplied by Dunbar Meat Packing.

Coupled ATP Hydrolysis Kinetics Assay. SMPs were prepared from bovine mitochondria as previously described (5). ATP hydrolytic activity of SMPs was measured by coupling the production of ADP to oxidation of NADH. Briefly, 125 μL aliquots of SMPs (57 $\mu\text{g mL}^{-1}$) in hydrolysis buffer (100 mM

Tris-HCl, pH 8.0, at 25 $^{\circ}\text{C}$, MgCl_2 (8 mM), KCl (50 mM), and EDTA (0.2 mM) were added to 96-well plates containing 5 μL of 50 \times drug (final) or DMSO vehicle control (1% DMSO, final) and incubated at 30 $^{\circ}\text{C}$ for 5 min. A 125 μL aliquot of substrate-coupling mixture (containing varied ATP (0.1–2.0 mM), NADH (0.2 mM), PEP (1 mM), PK (2 U mL^{-1}), and LDH (2 U mL^{-1}) in hydrolysis buffer) was then added, and the rate of NADH oxidation was monitored for 10 min at 340 nm, 30 $^{\circ}\text{C}$ (ϵ for NADH = 6.22 $\text{mM}^{-1} \text{cm}^{-1}$).

Coupled ATP Synthesis Kinetics Assay. ATP synthetic activity of SMPs was measured by coupling production of ATP to reduction of $NADP^+$. Briefly, a 125 μL aliquot of SMPs (0.16 mg mL^{-1}) in synthesis buffer (HEPES (10 mM), succinate (20 mM), glucose (20 mM), K_2HPO_4 (10 mM; omitted under conditions of varying $[P_i]$), ADP (1.2 mM; omitted under conditions of varying $[\text{ADP}]$), MgCl_2 (6 mM), AMP (11 mM), rotenone (2 μM), and $A_{P_5}A$ (150 μM), pH 8.0) was added to 96-well plates containing 5 μL of 50 \times drug (final) or DMSO vehicle control (1% DMSO, final) and incubated at 30 $^{\circ}\text{C}$ for 5 min. A 125 μL aliquot of substrate-coupling mixture (containing either varied $[\text{ADP}]$ (1.875–1200 μM) or varied $[P_i]$ (0.05–10 mM) and $NADP^+$ (0.75 mM), HK (4 U mL^{-1}) plus G6PDH (2 U mL^{-1}) in synthesis buffer) was added, and the rate of $NADP^+$ reduction was monitored for 15 min at 340 nm, 30 $^{\circ}\text{C}$. The range of substrate concentration encompasses those necessary to achieve physiological rates of catalysis, overlaps with physiologic concentrations observed in living cells, and can be sufficiently described using one K_m value.

Analysis of Kinetic Data. The apparent kinetic parameters at each concentration of inhibitor were determined by fitting the Michaelis–Menten equation to the dependence of initial velocity (v) on substrate concentration. Kinetic parameters were plotted versus concentration of inhibitor, generating secondary plots of the effect of this inhibitor on the hydrolysis and synthesis of kinetic parameters. These secondary plots were fit using eqs 1–4,

$$k_{\text{app}} = k / (1 + ([I]/K_i)) \quad (1)$$

$$V_{\text{max}(\text{app})} = V_{\text{max}} / (1 + ([I]/K_{i(\text{ES})})^{n(\text{ES})}) \quad (2)$$

$$K_{m(\text{app})} = K_m (1 + ([I]/K_{i(\text{ES})})^{n(\text{ES})}) / (1 + ([I]/K_{i(\text{E})})^{n(\text{E})}) \quad (3)$$

$$V_{\text{max}}/K_{m(\text{app})} = (V_{\text{max}}/K_m) / (1 + ([I]/K_{i(\text{E})})^{n(\text{E})}) \quad (4)$$

where k_{app} is the apparent kinetic parameter in the presence of inhibitor, k is the kinetic parameter in the absence of inhibitor, I is the inhibitor, K_i is the inhibition constant, $K_{i(\text{E})}$ and $K_{i(\text{ES})}$ are the inhibition constants describing the competitive and uncompetitive portions of mixed inhibition, respectively, and $n_{(\text{E})}$ and $n_{(\text{ES})}$ are the cooperativity factors for the competitive and uncompetitive inhibition constants, respectively.

To confirm the best model to fit the data, a three-dimensional fit of the dependence of initial velocity (v) on both concentration of substrate (S)

and inhibitor (I) was performed using equations for uncompetitive, noncompetitive, or mixed inhibition as described by eqs 5–7, respectively.

$$v = (V_{\max}[S]) / (K_m + (1 + [I]/K_i)[S]) \quad (5)$$

$$v = (V_{\max}[S]) / ((1 + ([I]/K_i)K_m) + (1 + ([I]/K_i)[S]) \quad (6)$$

$$v = (V_{\max}[S]) / ((1 + ([I]/K_{i(E)})^{n(E)}K_m) + (1 + ([I]/K_{i(ES)})^{n(ES)}[S]) \quad (7)$$

Off-Rate Measurements. The rate of inhibition of ATP synthesis was measured as described above with the following changes. A 20 μL aliquot of SMPs (1 mg mL⁻¹) in synthesis buffer (with 10 mM K₂HPO₄) was added to each well of a 96-well plate containing 2.5 μL of 100 \times drug (14 μM Bz-423 and 35 nM oligomycin, final) or DMSO vehicle control. The drug was diluted by adding 200 μL of synthesis buffer to each well. Substrate-coupling mixture (25 μL with 1 mM ADP) was added to each well at the indicated time point, and the rate of NADP⁺ reduction was monitored over time.

Acknowledgment: Supported by NIH Grant AI-47450 to G.D.G.

REFERENCES

- Boyer, P. D. (1997) The ATP synthase—a splendid molecular machine, *Annu. Rev. Biochem.* 66, 717–749.
- Blatt, N. B., Bednarski, J. J., Warner, R. E., Leonetti, F., Johnson, K. M., Boitano, A., Yung, R., Richardson, B. C., Johnson, K. J., Ellman, J. A., Opiari, A. W., Jr., and Glick, G. D. (2002) Benzodiazepine-induced superoxide signals B cell apoptosis: mechanistic insight and potential therapeutic utility, *J. Clin. Invest.* 110, 1123–1132.
- Leist, M., Single, B., Castoldi, A. F., Kuhnle, S., and Nicoletti, P. (1997) Intracellular adenosine triphosphate (ATP) concentration: a switch in the decision between apoptosis and necrosis, *J. Exp. Med.* 185, 1481–1486.
- Bednarski, J. J., Warner, R. E., Rao, T., Leonetti, F., Yung, R., Richardson, B. C., Johnson, K. J., Ellman, J. A., Opiari, A. W., Jr., and Glick, G. D. (2003) Attenuation of autoimmune disease in Fas-deficient mice by treatment with a cytotoxic benzodiazepine, *Arthritis Rheum.* 48, 757–766.
- Johnson, K. M., Chen, X., Boitano, A., Swenson, L., Opiari, A. W., Jr., and Glick, G. D. (2005) Identification and validation of the mitochondrial F1F₀-ATPase as the molecular target of the immunomodulatory benzodiazepine Bz-423, *Chem. Biol.* 12, 485–496.
- Perl, A., Gergely, P., Jr., Nagy, G., Koncz, A., and Banki, K. (2004) Mitochondrial hyperpolarization: a checkpoint of T-cell life, death and autoimmunity, *Trends Immunol.* 25, 360–367.
- Kramar, R., Hohenegger, M., Srour, A. N., and Khanak, G. (1984) Oligomycin toxicity in intact rats, *Agents Actions* 15, 660–663.
- Swinney, D. C. (2004) Biochemical mechanisms of drug action: what does it take for success? *Nat. Rev. Drug Discovery* 3, 801–808.
- Christopoulos, A. (2002) Allosteric binding sites on cell surface receptors: novel targets for drug discovery, *Nat. Rev. Drug Discovery* 1, 198–210.
- Comish-Bowden, A. (1986) Why is uncompetitive inhibition so rare? A possible explanation, with implications for the design of drugs and pesticides, *FEBS Lett.* 203, 3–6.
- Gao, Y. Q., Yang, W., and Karplus, M. (2005) A structure-based model for the synthesis and hydrolysis of ATP by F₁-ATPase, *Cell* 123, 195–205.
- Hutcheon, M. L., Duncan, T. M., Ngai, H., and Cross, R. L. (2001) Energy-driven subunit rotation at the interface between subunit a and the c oligomer in the F₀ sector of *Escherichia coli* ATP synthase, *Proc. Natl. Acad. Sci. U.S.A.* 98, 8519–8524.
- Menz, R. I., Walker, J. E., and Leslie, A. G. (2001) Structure of bovine mitochondrial F₁-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis, *Cell* 106, 331–341.
- Milgrom, Y. M., and Cross, R. L. (2005) Rapid hydrolysis of ATP by mitochondrial F₁-ATPase correlates with the filling of the second of three catalytic sites, *Proc. Natl. Acad. Sci. U.S.A.* 102, 13831–13836.
- Weber, J., and Senior, A. E. (2000) ATP synthase: what we know about ATP hydrolysis and what we do not know about ATP synthesis, *Biochim. Biophys. Acta* 1458, 300–309.
- Gledhill, J. R., and Walker, J. E. (2005) Inhibition sites in F₁-ATPase from bovine heart mitochondria, *Biochem. J.* 386, 591–598.
- Ramdas, L., Bunnin, B. A., Plunkett, M. J., Sun, G., Ellman, J., Gallick, G., and Budde, R. J. (1999) Benzodiazepine compounds as inhibitors of the src protein tyrosine kinase: screening of a combinatorial library of 1,4-benzodiazepines, *Arch. Biochem. Biophys.* 368, 394–400.
- Salomon, A. R., Voehringer, D. W., Herzenberg, L. A., and Khosla, C. (2000) Understanding and exploiting the mechanistic basis for selectivity of polyketide inhibitors of F₀F₁-ATPase, *Proc. Natl. Acad. Sci. U.S.A.* 97, 14766–14771.
- Esmann, M. (1991) Oligomycin interaction with Na,K-ATPase: oligomycin binding and dissociation are slow processes, *Biochim. Biophys. Acta* 1064, 31–36.
- Morrison, J. F., and Walsh, C. T. (1988) The behavior and significance of slow-binding enzyme inhibitors, *Adv. Enzymol. Relat. Areas Mol. Biol.* 61, 201–301.
- Lardy, H., Reed, P., and Lin, C. H. (1975) Antibiotic inhibitors of mitochondrial ATP synthesis, *Fed. Proc.* 34, 1707–1710.
- Augustin, H. W., and Spengler, V. (1983) Energy balance in rabbit reticulocytes and its control by adenine nucleotides, *Biomed. Biochim. Acta* 42, S223–228.
- Carbajo, R. J., Kellas, F. A., Runswick, M. J., Montgomery, M. G., Walker, J. E., and Neuhau, D. (2005) Structure of the F₁-binding domain of the stator of bovine F₁F₀-ATPase and how it binds an alpha-subunit, *J. Mol. Biol.* 351, 824–838.
- Chene, P. (2002) ATPases as drug targets: learning from their structure, *Nat. Rev. Drug Discovery* 1, 665–673.
- Gellerich, F. N., Laterveer, F. D., Zier, S., and Nicolay, K. (2002) The quantitation of ADP diffusion gradients across the outer membrane of heart mitochondria in the presence of macromolecules, *Biochim. Biophys. Acta* 1554, 48–56.
- Lipton, S. A. (2006) Paradigm shift in neuroprotection by NMDA receptor blockade: memantine and beyond, *Nat. Rev. Drug Discovery* 5, 160–170.
- Curtis, R., Geesaman, B. J., and DiStefano, P. S. (2005) Aging and metabolism: drug discovery opportunities, *Nat. Rev. Drug Discovery* 4, 569–580.
- Fox, C. J., Hammerman, P. S., and Thompson, C. B. (2005) Fuel feeds function: energy metabolism and the T-cell response, *Nat. Rev. Immunol.* 5, 844–852.
- Murray, C. M., Hutchinson, R., Bantick, J. R., Belfield, G. P., Benjamin, A. D., Brazma, D., Bundick, R. V., Cook, I. D., Craggs, R. I., Edwards, S., Evans, L. R., Harrison, R., Holness, E., Jackson, A. P., Jackson, C. G., Kingston, L. P., Perry, M. W., Ross, A. R., Rugman, P. A., Sidhu, S. S., Sullivan, M., Taylor-Fishwick, D. A., Walker, P. C., Whitehead, Y. M., Wilkinson, D. J., Wright, A., and Donald, D. K. (2005) Monocarboxylate transporter MCT1 is a target for immunosuppression, *Nat. Chem. Biol.* 1, 371–376.
- Ramanathan, A., Wang, C., and Schreiber, S. L. (2005) Perturbational profiling of a cell-line model of tumorigenesis by using metabolic measurements, *Proc. Natl. Acad. Sci. U.S.A.* 102, 5992–5997.

A New Approach for Enhancing Differential Selectivity of Drugs to Cancer Cells

Muralikrishna Duvvuri[†], Samidha Konkar[†], Kwon Ho Hong[‡], Brian S. J. Blagg[‡], and Jeffrey P. Krise^{†,*}

Departments of [†]Pharmaceutical Chemistry and [‡]Medicinal Chemistry, The University of Kansas, Lawrence, Kansas 66047

A fundamental goal in the battle against cancer is to design drugs that selectively target tumor cells but have limited effects against the same machinery in nontransformed cells. Many investigators have attempted to direct anticancer agents toward cancer cells by taking advantage of specific receptors or enzymes that are known to be overexpressed on the plasma membrane and thereby increase the concentration of drugs in the vicinity of transformed cells (1). Although conceptually appealing, no such approaches have proven to be highly efficacious. Unfortunately, most conventional anticancer agents have sufficient membrane permeability to penetrate both normal and cancer cells to a similar extent.

We describe herein an innovative strategy by which the selective toxicity of drugs toward cancer cells can be rationally enhanced. We propose that drugs, with optimized physicochemical characteristics, can have considerably different intracellular distribution patterns in normal versus transformed cells and can provide a basis for enhancing selectivity toward malignant cells. Specifically, it is proposed that inhibitors can accumulate differentially in desirable intracellular locations (*i.e.*, a site containing the drug target) depending on the pH status of the cell.

The basis for altered intracellular distribution of inhibitors in normal versus cancer cells relies on differences in intracellular pH gradients that have been previously shown to exist. Normal cells have acidic lysosomes in the pH range of 4–5, whereas the cytosol is approximately neutral. This lysosome-to-cytosol pH gradient provides a driving force for accumulation of weakly basic compounds into the acidic organelles. In the cytosol, many weakly basic drugs (depending on their pK_a value) exist predominantly in a unionized state and can readily permeate lipid bilayers. When such bases cross the lysosomal lipid bilayer, they readily

ABSTRACT The degree to which anticancer agents selectively target cancer cells is a key determinant in successful therapeutic outcomes. Inhibitors of the Hsp90 molecular chaperone represent an important new class of anticancer agents. We propose here a novel mechanism by which physicochemical properties of Hsp90 inhibitors can be optimized to increase selectivity towards cancer cells. The basis for this approach relies on differential intracellular pH gradients that have been shown to exist between normal and transformed cells. Five Hsp90 inhibitors containing basic or neutral properties were evaluated in antiproliferation assays using cells with variable lysosomal pH. Inhibitors with basic functionalities had reduced activity in cells with normal (low) lysosomal pH but showed significantly greater activity in cells with abnormally elevated lysosomal pH (similar to what has been recorded in many types of cancer cells). Conversely, such selectivity enhancement was not observed for neutral inhibitors. The mechanistic basis for the observed selectivity was demonstrated quantitatively by determining the concentration of inhibitors within relevant intracellular compartments. Collectively, these findings suggest that Hsp90 inhibitors with optimal basicity and physicochemical properties have enhanced selectivity toward cancer cells than their neutral counterparts. It is anticipated that these findings may be applicable to other classes of anticancer agents for improvement of differential selectivity.

*To whom correspondence should be addressed.
E-mail: krise@ku.edu.

Received for review March 15, 2006
and accepted May 12, 2006.

Published online June 2, 2006
10.1021/cb6001202 CCC: \$33.50

© 2006 by American Chemical Society

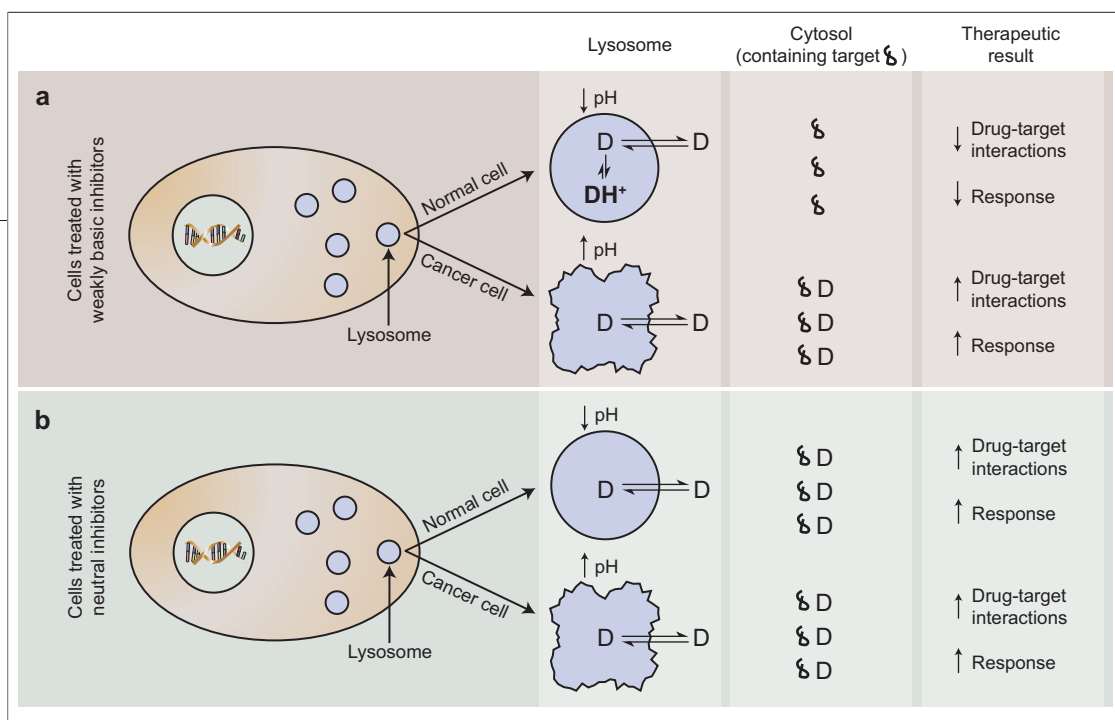


Figure 1. Schematic illustrating the mechanism for improved selectivity of weakly basic inhibitors toward cancer cells proposed in this study. **a)** Cells treated with weakly basic inhibitors have different intracellular distribution in normal cells (having low lysosomal pH) compared to transformed cells (having defective lysosomal acidification and higher lysosomal pH). Weakly basic drugs (represented as D and DH^+ for the free base and ionized species, respectively) are extensively compartmentalized in lysosomes within normal cells according to an ion trapping mechanism. This results in reduced drug available to interact with cytosolic target molecules and therefore reduced response. Alternatively, cancer cells with elevated lysosomal pH have a reduced capacity to sequester drugs in the lysosomes. This increases the availability of the drug to interact with cytosolic targets and leads to a greater potential for therapeutic response. **b)** Cells treated with neutral inhibitors (represented as D) do not have differential intracellular distribution in normal cells versus cancer cells, since these compounds are not subject to ion trapping in lysosomes and are not influenced by differences in lysosomal pH existing between the cell types. Similar amounts of drug are available to interact with cytosolic target molecules in both normal and cancer cells, and therefore, similar therapeutic responses can be expected.

capacity for sequestering weakly basic molecules possessing optimal characteristics. For example, the weakly basic molecule quinacrine has been shown to achieve concentrations inside lysosomes which are in excess of 750 times the concentration contained in the cell culture medium (8). Therefore, extensive lysosomal trapping of drugs can result in dramatically reduced cytosolic concentrations.

ionize into their ammonium salts, which are typically unable to diffuse back into the cytosol. This sequestration process has been referred to as ion trapping, as well as pH-partitioning, and the theoretical basis for this phenomenon has been previously reviewed (2).

Although the precise mechanism is undetermined, it is clear that numerous cancer cell lines have defective acidification of lysosomes (3–5). For example, the lysosomal pH of the human leukemic cell line HL-60 is 6.5, whereas the pH of cytosol is near neutral (6). Consequently, there is very little driving force for the accumulation of weakly basic drugs into lysosomes in these cells. As a result, weakly basic molecules have a greater propensity to accumulate in the cytosol of such cells (similar to how a neutral molecule may distribute).

Methods to quantitate the accumulation of drugs in lysosomes and the cytosol, as well as other intracellular compartments, have recently been reported (7, 8). Specifically, structure–localization relationships that correlate weak-base permeability characteristics and pK_a values with lysosomal sequestration tendency in cells with normally acidified lysosomes have been established (9, 10). Collectively, these evaluations have shown that lysosomes exhibit an extremely high

If the anticancer drug target resides in the cytosol, it is proposed that the difference in lysosomal pH between normal and malignant cells will provide a basis for the enhancement of differential selectivity. Therefore, drugs with optimized physicochemical properties that promote lysosomal sequestration will have lower cytosolic concentrations in normal cells (*i.e.*, those with low lysosomal pH) compared to transformed cells. A schematic illustration of the proposed mechanism for enhanced selectivity towards cancer cells is illustrated (Figure 1). To demonstrate this novel selectivity principle, we chose to focus our investigation on Hsp90 inhibitors. This class of anticancer agents was particularly suitable since (1) there are many inhibitors in the class that have variable physicochemical properties yet similar binding affinity to Hsp90 *in vitro* (11); and (2) Hsp90, the molecular target of these inhibitors, is localized in the cell cytosol (12) and would therefore be sensitive to changes in drug sequestration in lysosomes.

RESULTS AND DISCUSSION

Comparative Cytotoxicity of Hsp90 Inhibitors. The antiproliferative activities of five Hsp90 inhibitors of

varying neutral or weakly basic character were determined. The structures of these compounds are shown (Figure 2). GDA and RAD are neutral molecules, whereas 17-DMAG, 17-DMAP, and 17-AEP are weakly basic (see Table 1 for pK_a values). These compounds were evaluated in the human leukemic cell line HL-60 and its multidrug-resistant (MDR) variant. This pair of cell lines was chosen because the HL-60 strain has a high lysosomal pH (6.5) and the MDR strain mimics cells with normal lysosomal acidification, which was measured to be pH 5.1 (6). The IC_{50} values obtained for these inhibitors are reported (Table 1). To illustrate the selectivity for cells with high versus low lysosomal pH, we divided the IC_{50} value obtained for a given inhibitor in the MDR cell line by the IC_{50} for the inhibitor in the parent cell line. High IC_{50} ratios indicate inhibitors with greater selectivity toward cancer cells, or any cell line with abnormally high lysosomal pH. The ratio of the IC_{50} values is relatively small for neutral inhibitors (Figure 3, panel a), suggesting they have little differential intracellular distribution and selectivity in cells with normal lysosomal pH compared to those that have elevated lysosomal pH. As proposed, weakly basic inhibitors (*i.e.*, 17-DMAG, 17-AEP, and 17-DMAP) had significantly enhanced IC_{50} ratios that result from the differential lysosomal pH observed in these cell lines. Since we are reporting ratios in Figure 3, the observed trend could possibly occur if weakly basic compounds were progressively accumulating to a greater degree in the MDR cell line relative to the drug-sensitive cell line. Although we have not quantitatively evaluated the total cell concentrations for all compounds examined, we have done so for the neutral inhibitor GDA and the basic inhibitor 17-DMAG.

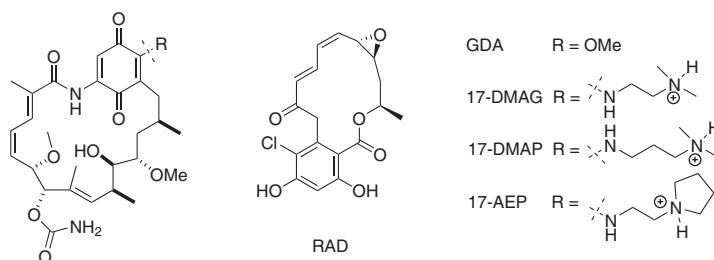


Figure 2. Structures of Hsp90 inhibitors evaluated. GDA and RAD are neutral compounds; the rest (17-DMAG, 17-AEP, and 17-DMAP) are weak bases (pK_a values are reported in Table 1).

We found that the HL-60 cells consistently accumulated about twice the amount of inhibitors relative to the MDR HL-60 cell line (*i.e.*, 1.9- and 1.7-fold increases for 17-DMAG and GDA, respectively). Considering the IC_{50} ratios were not consistently 2-fold, we do not believe this to be the major contributor of the observed trend

Although MDR HL-60 cells have lower lysosomal pH relative to drug-naïve HL-60 cells, it could be argued that the differences observed in these cell lines could be attributed to other factors such as the expression of drug transporters on the plasma membrane of MDR cells that may preferentially efflux weakly basic molecules more readily than neutral molecules (13). However, this is not perceived to be a significant concern with these cells, since we have previously shown that the MDR HL-60 cells do not express transporters typically associated with drug efflux at the plasma membrane (6). Nevertheless, we determined the IC_{50} values for these compounds in a different cell line, namely, RPTEC cells. To mimic the defective acidification that occurs when cells undergo transformation, we pretreated

TABLE 1. Physicochemical and cytotoxicity values for Hsp90 inhibitors

Inhibitor	Classification	pK_a^a	α	IC_{50} values (μM) ^b			
				HL-60		RPTEC	
				HL-60	MDR HL-60	+ NH_4Cl	- NH_4Cl
GDA	neutral	NA ^c	NA ^c	0.9 ± 0.5	4.7 ± 0.2	0.1 ± 0.1	0.05 ± 0.02
RAD	neutral	NA ^c	NA ^c	8.3 ± 0.1	6.0 ± 1.7	1.7 ± 1.0	3.0 ± 0.1
17-DMAG	weak base	7.6	0.010 ± 0.003	0.1 ± 0.02	2.2 ± 0.6	0.1 ± 0.02	0.4 ± 0.2
17-AEP	weak base	8.1	0.005 ± 0.002	0.05 ± 0.01	2.9 ± 0.1	0.04 ± 0.01	1.4 ± 0.7
17-DMAP	weak base	8.1	0.010 ± 0.007	0.08 ± 0.02	8.8 ± 0.3	0.1 ± 0.04	5.5 ± 0.1

^aObtained by curve-fitting the pH-partition plots in Figure 5 (see Methods for procedure). ^bMean ± standard deviation ($n = 3$). ^cNA, not applicable.

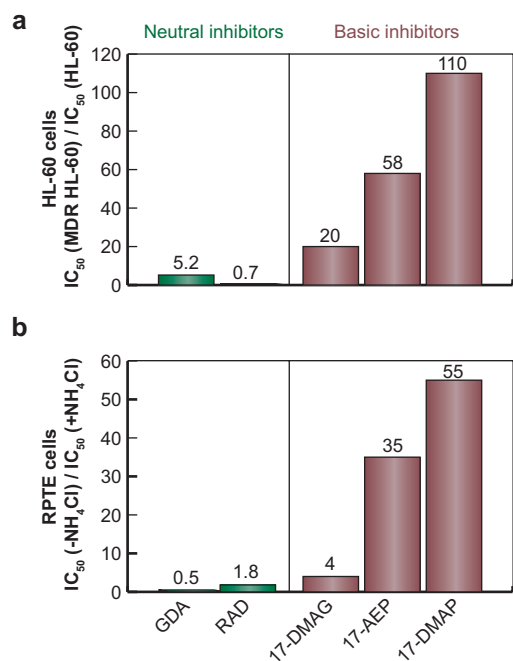


Figure 3. Inhibitors of Hsp90 that have weakly basic functionalities have greater activity in cells with elevated lysosomal pH relative to cells with low lysosomal pH, which is not the case for neutral Hsp90 inhibitors. a) Weakly basic inhibitors have greater differential selectivity (IC₅₀ ratios) in cells with high (HL-60) versus low (MDR HL-60) lysosomal pH. Neutral inhibitors do not show such differential selectivity. b) Weakly basic inhibitors have greater differential selectivity in RPTE cells treated with NH₄Cl (inducing increased lysosomal pH) compared to untreated cells with low and normal lysosomal pH. Neutral inhibitors do not show such differential selectivity. The bars in panels a and b represent the ratio of IC₅₀ values for compounds in cells with low lysosomal pH over the IC₅₀ value in cells with elevated lysosomal pH. Refer to Table 1 for individual IC₅₀ values and associated experimental errors.

treatment of cells with 5 mM NH₄Cl did not influence cell viability during the time course of the experiment (data not shown).

Together, the general trends represented in Figure 3 strongly suggest that weakly basic anticancer agents that have a cytosolic target and physicochemical characteristics that favor sequestration in lysosomes with low pH values (*i.e.*, 17-DMAG, 17-AEP, and 17-DMAP) have preferential activity in cells with higher lysosomal pH (*i.e.*, transformed cells) compared to normal cells with low lysosomal pH.

RPTE cells with 5 mM NH₄Cl, which is a well-known reagent used to elevate lysosomal pH (14). The IC₅₀ values of inhibitors in these cells were determined in the presence and absence of NH₄Cl pretreatment. The results from these studies are presented (Table 1). These results (Figure 3, panel b) parallel those obtained in HL-60 cells (Figure 3, panel a). Once again, inhibitors with weakly basic functional groups had greater activity against cells with high lysosomal pH (*i.e.*, those treated with NH₄Cl). Conversely, raising the lysosomal pH with NH₄Cl had very little effect on differential selectivity of neutral inhibitors. As a control, we had shown that treat-

Intracellular Distribution of Hsp90 Inhibitors. In an effort to provide mechanistic support to the previously stated theoretical basis for differential selectivity, we quantitatively investigated the intracellular distribution differences of two Hsp90 inhibitors in drug-sensitive and MDR HL-60 cells that have differential lysosomal acidification, as previously stated. From the set of five inhibitors previously evaluated for cytotoxicity, we selected GDA and 17-DMAG as representatives of neutral and basic inhibitors, respectively. Lysosome and cytosol concentrations of GDA and 17-DMAG were evaluated in cells according to our previously reported procedure (9). Results from these studies are shown (Figure 4). These data clearly show that neither compound is extensively sequestered into the lysosomes of drug-naïve HL-60 cells, which is consistent with the fact that lysosomal pH is elevated in this cell line to 6.5, which is near cytosolic pH. This translates to relatively low lysosome-to-cytosol concentration ratios for GDA (~1) and 17-DMAG (~16). In contrast, this ratio is ~175 for 17-DMAG and ~2 for the neutral compound, GDA, in MDR HL-60 cells that have low lysosomal pH. These data provide quantitative support for the selective cytotoxicity (Figure 3).

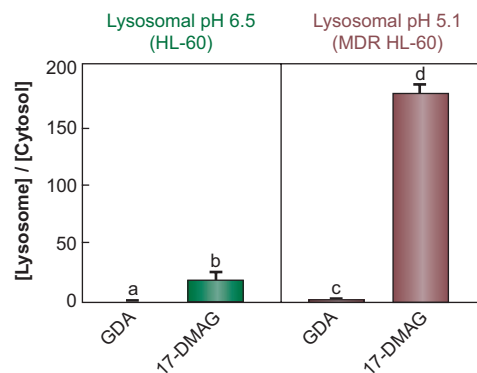


Figure 4. Quantitative evaluations of the intracellular distribution of Hsp90 inhibitors support the observed cytotoxicity differences between neutral and weakly basic inhibitors shown in Figure 3. The weakly basic Hsp90 inhibitor 17-DMAG extensively concentrates in lysosomes in a pH-dependent manner in HL-60 cells. The neutral inhibitor, GDA, does not significantly concentrate in lysosomes regardless of lysosomal pH. Bars represent the experimentally determined concentration of the respective inhibitor obtained from purified lysosomes divided by the concentration in cytosol. Mean \pm standard deviations from three experiments are shown ($a \neq b, P < 0.1$; $c \neq d, P < 0.0001$).

Physicochemical Evaluations of Hsp90 Inhibitors. To optimize the design of inhibitors with enhanced selectivity toward cancer cells based on the lysosome sequestration mechanism presented in this work, it is important to focus on the differential selectivity trends shown for weakly basic inhibitors (Figure 3). Specifically, 17-DMAP produced higher differential selectivity than 17-AEP, which was more selective than 17-DMAG.

In an attempt to rationalize this trend, the physicochemical properties of these weakly basic inhibitors were investigated. Two key parameters appear to be important for lysosomal sequestration. The first has to do with permeability of the base in the ionized state (referred to as α (10)) and the second is the pK_a value of the base (9). Briefly, α is an indicator of the propensity of the weakly basic molecule to diffuse out of the lysosomal lumen in its ionized state. Obviously, the degree to which this occurs will decrease lysosomal retention. Theoretically, α values can range from 0 to 1. When α is near zero, the base is assumed to be completely membrane-impermeable in the ionized state and will therefore significantly accumulate in lysosomes by ion trapping (*i.e.*, it cannot readily reestablish the equilibrium with the ionized inhibitor contained in the cytosol). Alternatively, if the α value is near 1, the base can freely diffuse out of the lysosomes in the ionized state, will behave similar to a neutral molecule, and will not be significantly sequestered into this acidic organelle. The α values were experimentally determined using data generated from octanol/buffer partitioning behavior as a function of pH (Figure 5). Calculated α values from plots shown (Figure 5) are reported (Table 1). The α value for 17-DMAG is low (0.01) and similar to the values obtained for the other weakly basic inhibitors evaluated in this study. Considering these findings, we do not believe that differences in α are playing a significant role in the aforementioned selectivity trend shown (Figure 3).

From the plots shown (Figure 5), we were also able to extrapolate the pK_a value for each base, which, as previously mentioned, is another important factor in lysosomal sequestration. In a previous report, we have quantitatively evaluated the influence of pK_a values for model compounds in cells with low lysosomal pH and have shown that molecules with higher pK_a values accumulated in lysosomes to a greater extent than compounds with lower pK_a (9).

The calculated pK_a values for the bases evaluated in this work are reported (Table 1). The fact that both

17-AEP and 17-DMAP have significantly higher pK_a values relative to 17-DMAG is consistent with greater lysosomal sequestration and is consistent with the observed higher selectivity for these compounds. 17-DMAP appears to have greater differential selectivity than 17-AEP. Although 17-DMAP has a slightly higher pK_a value than 17-AEP, the difference is not significant enough to explain the observed differences in selectivity between these two inhibitors.

The findings presented here suggest that the differential selectivity can be further enhanced by taking into account the physicochemical properties of the inhibitor and the pH status of the cancer cell. Although investigations presented in this work focused on inhibitors of Hsp90, the results presented here are expected to have widespread application in the design, selection, or modification of future anticancer agents with improved differential selectivity. Future studies with additional pairs of cell lines and different classes of drugs will be essential in defining the ultimate scope of the work presented here.

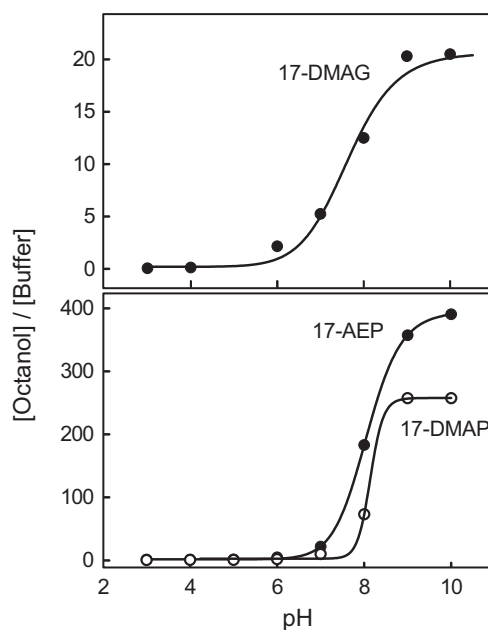


Figure 5. Partition coefficients (octanol/buffer) as a function of pH for weakly basic Hsp90 inhibitors. The above plots are used to determine pK_a and α values, both of which influence lysosomal sequestration tendency and therefore selectivity of the agents toward cancer cells. The α and pK_a values generated from these plots are reported (Table 1).

METHODS

Cell Lines and Reagents. The human acute promyelocytic leukemia cell line HL-60 and the doxorubicin-selected, drug-resistant MDR HL-60 cell lines were kindly provided by Dr. Yueshang Zhang (Arizona Cancer Center, University of Arizona). Renal proximal tubule epithelial (RPTE) cells were purchased from Cambrex Bio Science, and MC/CAR cell line (CRL-8083) was purchased from American Type Culture Collection. Geldanamycin (GDA) was obtained from the Developmental Therapeutics Program (National Institutes of Health). 17-(Dimethylamino)-17-demethoxygeldanamycin (17-DMAG) was synthesized according to a procedure described by Tian and co-workers (15). 17-(Dimethylaminopropylamino)-17-demethoxygeldanamycin (17-DMAP) and 17-[2-(pyrrolidin-1-yl)ethyl]amino-17-demethoxygeldanamycin (17-AEP) were purchased from Invivogen, and radicicol (RAD) was purchased from Sigma-Aldrich.

Determination of Octanol/Buffer Partition Coefficients. The partition coefficients for Hsp90 inhibitors were determined as a function of pH using a previously described shake-flask method (10). For 17-DMAG, 17-DMAP, and 17-AEP, partition coefficients of ionized (PC_i) and unionized base (PC_u) and the pK_a values were obtained by curve-fitting the pH-partition profile with the Hill equation (4 parameters) using SigmaPlot 2001 (SPSS, Inc.). The ratio of PC_i to PC_u was defined as the α value for a given compound.

Cytotoxicity Evaluations. Cell sensitivity to compounds was determined using WST-1 (4-[3-(4-iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolyl]-1,3-benzene disulfonate) as the cell proliferation reagent, and the assays were performed according to the manufacturer's instructions (Roche Applied Science). Briefly, cells were seeded in triplicate in 96-well plates at a density of 5000 cells per well in culture medium and were incubated with increasing concentrations of drug for a period of 72 h (RPTE cells were allowed to adhere to the wells overnight prior to addition of drug). At the end of the incubation period, 10 μ L of a 5 mg mL⁻¹ solution of WST-1 in phosphate-buffered saline (PBS) was added to each well, and the plates were returned to the incubator for an additional 2 h. The absorbance was measured at a wavelength of 450 nm in a Multiskan (model MCC/340) microplate reader (Thermo Electron Corp.). IC₅₀ was defined as the concentration of drug causing 50% inhibition of cell growth as compared with untreated control (in the absence of drug). Designated cells (+NH₄Cl) were incubated with 5 mM ammonium chloride for 1 h prior to co-incubation with drug for 72 h. All assays were done in triplicate.

Quantitating Drug in Subcellular Compartments. Cells (200 × 10⁶) were incubated with 5 μ M compound for 2 h. Subsequently, intact lysosomes were isolated from cells using a magnetic chromatographic approach previously described by our laboratory (10). Isolated lysosomal fraction was acidified with 0.1% formic acid, vortexed with 600 μ L of acetonitrile for 20 s, and centrifuged at 16 000g for 5 min. The supernatant was evaporated to dryness, and the residue was analyzed by HPLC to quantitate the amount of compound extracted. This amount was divided by the total lysosomal volume in the isolated fraction to obtain the lysosomal concentration of the compound. We have previously determined the total volume of isolated lysosomes using this procedure in MDR HL-60 cells to be 32.2 nL (8). For HL-60 and MC/CAR cells, the total volume of isolated lysosomes was estimated based on the total protein content of the isolated lysosomal fraction from each of these cell lines relative to that obtained from MDR HL-60 cells, assuming that the protein content of lysosomes remains constant between cell lines. Using this assumption, we determined the values for total volume of lysosomes in the isolated fraction from HL-60 and MC/CAR cells used in our calculations to be 124.8 and 40.7 nL, respectively.

Drug concentrations in cell cytosol were determined as previously described (7). To evaluate the influence of heat shock on the intracellular drug distribution, cells were incubated at 42 °C for 15 min and allowed to recuperate at 37 °C for 5 h prior to incubation with drug.

HPLC Analysis. An integrated Agilent 1100 series capillary liquid chromatography system comprised of pump, in-line degasser, column thermostat, and autosampler was used. A Waters Xterra MS C₁₈ column (100 mm × 1 mm; 3.5 μ m particle size) was employed and maintained at 30 °C. Injection volumes were 5 μ L for all samples.

The mobile phase was comprised of methanol and 0.1% formic acid. Solvents A and B contained 2% and 95% methanol, respectively. A gradient elution method (25 μ L min⁻¹) was used for 17-DMAG. For the initial 5 min, the mobile phase composition was 100% A, followed by a linear change to 90% A and 10% B in 1 min, then to 10% A and 90% B over 1 min, and maintained at 90% B for 5 min. Subsequently, the mobile phase composition was changed linearly to the initial conditions over 1 min, and the column was equilibrated for 7 min. For GDA, an isocratic elution method (25 μ L min⁻¹) was used with a mobile phase composition of 77% B and 23% A. The duration of runs for GDA and 17-DMAG were 10 and 20 min, respectively.

An API 2000 (Applied Biosciences) triple quadrupole mass spectrometer equipped with a Turbo Ionspray ionization source was employed for detecting 17-DMAG and operated in the positive mode. The transition from 617.1 to 58.1 amu was monitored with a declustering potential of 21 V and collision energy of 53 V. For GDA, an Agilent 1100 series variable wavelength detector was employed, and the column eluent was monitored at 308 nm.

Western Blot Analysis. MC/CAR cells (4 × 10⁷) were cultured in 10 mL media and allowed to incubate for 24 h at 37 °C under 5% CO₂. The cells were heat-shocked at 42 °C for 15 min and then incubated at 37 °C for 5 h, before the cells were harvested by washing with PBS and pelleted (500g for 5 min at 4 °C). PBS was removed, and the pellets were resuspended in lysis buffer (150 μ L, 50 mM Tris, pH 7.5, 1% NP-40, 150 mM NaCl, 2.5 mM Na₃VO₄, 10 mM PMSF, 10 μ M aprotinin, 10 μ M leupeptin, and 10 μ M soybean trypsin inhibitor) and transferred to a 1.5 mL centrifuge tube. Samples were incubated at 4 °C for 1 h with intermittent agitation. Lysates were cleared by centrifugation at 21 000g for 10 min at 4 °C. The supernatants were collected and comprised the experimental samples. The protein concentration of each sample was determined by a BSA assay (Pierce). Equal amount of protein was resolved on a 9% polyacrylamide gel (100 V, 100 mA) and transferred to a nitrocellulose membrane (30 V, 10 mA). Bands were visualized with Ponceau S to confirm protein transfer. Blots were blocked with 5% nonfat milk in PBST (3 × 20 mL), incubated with primary antibodies (Hsp90 α/β (H-114), sc-7947, lot no. J0504, rabbit polyclonal IgG, Santa Cruz Biotechnology, and actin (I-19)-R, sc1616-R, lot no. B1204, rabbit polyclonal IgG, Santa Cruz Biotechnology) for 2.5 h. The blots were washed (3 × 10 mL) with 5% nonfat milk in PBST, incubated with a horseradish peroxidase-conjugated secondary antibody (Amersham; 1 h), and washed with 5% nonfat milk in PBST (1 × 10 mL), followed by PBST (2 × 10 mL). Protein bands were visualized by chemiluminescence using the ECL detection reagents (Amersham).

Acknowledgment: This work was supported by a grant from the National Cancer Institute, Grant No. CA106655 (to J.P.K.).

REFERENCES

1. Jaracz, S., Chen, J., Kuznetsova, L. V., and Ojima, I. (2005) Recent advances in tumor-targeting anticancer drug conjugates, *Bioorg. Med. Chem.* 13, 5043–5054.

2. de Duve, C., de Barse, T., Poole, B., Trouet, A., Tulkens, P., and Van Hoof, F. (1974) Commentary. Lysosomotropic agents, *Biochem. Pharmacol.* **23**, 2495–2531.
3. Altan, N., Chen, Y., Schindler, M., and Simon, S. M. (1998) Defective acidification in human breast tumor cells and implications for chemotherapy, *J. Exp. Med.* **187**, 1583–1598.
4. Kokkonen, N., Rivinoja, A., Kauppila, A., Suokas, M., Kellokumpu, I., and Kellokumpu, S. (2004) Defective acidification of intracellular organelles results in aberrant secretion of cathepsin D in cancer cells, *J. Biol. Chem.* **279**, 39982–39988.
5. Schindler, M., Grabski, S., Hoff, E., and Simon, S. M. (1996) Defective pH regulation of acidic compartments in human breast cancer cells (MCF-7) is normalized in adriamycin-resistant cells (MCF-7adr), *Biochemistry* **35**, 2811–2817.
6. Gong, Y., Duvvuri, M., and Krise, J. P. (2003) Separate roles for the Golgi apparatus and lysosomes in the sequestration of drugs in the multidrug-resistant human leukemic cell line HL-60, *J. Biol. Chem.* **278**, 50234–50239.
7. Duvvuri, M., Feng, W., Mathis, A., and Krise, J. P. (2004) A cell fractionation approach for the quantitative analysis of subcellular drug disposition, *Pharm. Res.* **21**, 26–32.
8. Duvvuri, M., and Krise, J. P. (2005) A novel assay reveals that weakly basic model compounds concentrate in lysosomes to an extent greater than pH-partitioning theory would predict, *Mol. Pharm.* **2**, 440–448.
9. Duvvuri, M., Konkar, S., Funk, R. S., Krise, J. M., and Krise, J. P. (2005) A chemical strategy to manipulate the intracellular localization of drugs in resistant cancer cells, *Biochemistry* **44**, 15743–15749.
10. Duvvuri, M., Gong, Y., Chatterji, D., and Krise, J. P. (2004) Weak base permeability characteristics influence the intracellular sequestration site in the multidrug-resistant human leukemic cell line HL-60, *J. Biol. Chem.* **279**, 32367–32372.
11. Le Brazidec, J. Y., Kamal, A., Busch, D., Thao, L., Zhang, L., Timony, G., Grecko, R., Trent, K., Lough, R., Salazar, T., Khan, S., Burrows, F., and Boehm, M. F. (2004) Synthesis and biological evaluation of a new class of geldanamycin derivatives as potent inhibitors of Hsp90, *J. Med. Chem.* **47**, 3865–3873.
12. Neckers, L. (2002) Hsp90 inhibitors as novel cancer chemotherapeutic agents, *Trends Mol. Med.* **8**, S55–S61.
13. Sharma, R., Awasthi, Y. C., Yang, Y., Sharma, A., Singhal, S. S., and Awasthi, S. (2003) Energy dependent transport of xenobiotics and its relevance to multidrug resistance, *Curr. Cancer Drug Targets* **3**, 89–107.
14. Daniel, W. A., and Wojcikowski, J. (1999) Lysosomal trapping as an important mechanism involved in the cellular distribution of perazine and in pharmacokinetic interaction with antidepressants, *Eur. Neuropsychopharmacol.* **9**, 483–491.
15. Tian, Z. Q., Liu, Y., Zhang, D., Wang, Z., Dong, S. D., Carreras, C. W., Zhou, Y., Rastelli, G., Santi, D. V., and Myles, D. C. (2004) Synthesis and biological activities of novel 17-aminogeldanamycin derivatives, *Bioorg. Med. Chem.* **12**, 5317–5329.

Experimental Tests of Two Proofreading Mechanisms for 5'-Splice Site Selection

Yangming Wang and Scott K. Silverman*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

ABSTRACT Self-splicing group II intron RNAs catalyze a two-step process in which the intron is excised as a lariat by two successive phosphodiester exchange reactions. The reversibility of the first step has been hypothesized to act as a proofreading mechanism for improper 5'-splice site selection. However, without synthetic access to mis-spliced RNAs, this hypothesis could not be tested. Here, we used a deoxyribozyme to synthesize several branched RNAs that are derived from the ai5 γ group II intron and mis-spliced at the 5'-splice site. Unlike the correctly spliced ai5 γ RNAs, the mis-spliced RNAs are observed not to undergo the reverse of the first step. This is well-controlled negative evidence against the hypothesis that first-step reversibility is a proofreading mechanism for 5'-splice site selection. In a reaction equivalent either to the hydrolytic first step of splicing or to the hydrolytic reverse of the second step of splicing, a mis-spliced 5'-exon can be "trimmed" to its proper length by the corresponding mis-spliced intron, and in one case, the trimmed 5'-exon was observed to proceed correctly through the second step of splicing. These findings are the first direct evidence that this second proofreading mechanism can occur with a group II intron RNA that is mis-spliced at the 5'-splice site. On the basis of the likely structural and evolutionary relationship between group II introns and the spliceosome, we suggest that this second proofreading mechanism may be operative in the spliceosome.

The RNA self-splicing pathway catalyzed by group II intron RNAs (1–5) has two reaction steps (Figure 1, panel a). In the first step, an internal adenosine 2'-hydroxyl group from domain 6 attacks the 5'-splice site phosphodiester linkage, forming a lariat–3'-exon intermediate with departure of the 5'-exon. In the second step, the 5'-exon attacks the 3'-splice site, forming the ligated exons and excising the intron as a lariat. Although the first step of splicing is reversible (6), the underlying explanation for this reversibility is unclear. The lack of natural 5'-splice site mis-splicing (7–9) suggests that a proofreading mechanism may exist. One specific hypothesis is that first-step reversibility is itself a proofreading mechanism: any intron that mis-splices at the first step by choosing an improper 5'-splice site can return to the initial unspliced state by the reverse of the first step (6). This would provide the RNA another opportunity to splice correctly, rather than waste the RNA molecule in a dead-end route or lead to improperly spliced exons after subsequently proceeding through the second step of splicing. This first proofreading mechanism requires that mis-spliced introns which have been formed by use of the improper 5'-splice site during the first step will be competent to proceed through the reverse of the first step. However, this mechanism has never been tested experimentally, because the required mis-spliced RNAs could not be synthesized using previously available methods.

Because a mis-spliced 5'-exon that proceeds through the second step would lead to improperly ligated exons, a natural mechanism to repair a mis-spliced 5'-exon would be valuable. A second potential proofreading mechanism for RNAs that are mis-spliced at the 5'-splice site is for the incorrect 5'-exon to be converted to the correct length 5'-exon and subsequently proceed through the second step. When the incorrect 5'-exon has one or more extra nucleotides due to mis-splicing

*To whom correspondence should be addressed.

E-mail: scott@scs.uiuc.edu.

Received for review April 6, 2006
and accepted May 6, 2006.

Published online June 16, 2006

10.1021/cb6001569 CCC: \$33.50

© 2006 by American Chemical Society

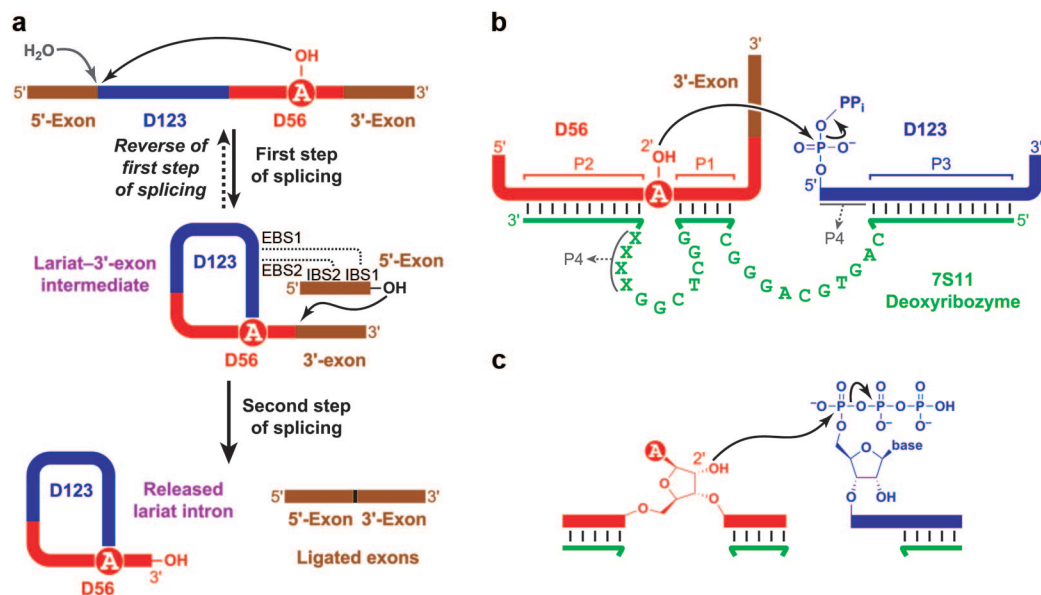


Figure 1. RNA splicing and branched RNA. **a)** The two steps of group II intron splicing. Interactions between the intron binding sites (IBS) and exon binding sites (EBS) are also present during the first step (not depicted). For the ai5 γ intron, D123 is 675 nt and D56 is 77 nt; D4 is not depicted because it is dispensable for catalytic activity (52). First-step reversibility has been observed for correctly spliced RNA (6); whether or not this reversibility applies for mis-spliced RNAs is one focus of this manuscript. Hydrolytic cleavage at the 5'-splice site (4, 10–12) is also depicted (gray water molecule); this leads to linear instead of branched RNA. **b)** Synthesis of 2',5'-branched ai5 γ RNA catalyzed by the 7S11 deoxyribozyme (20, 21). For natural branched RNAs, the branch-site nucleotide is almost always adenosine (circled). For applying 7S11 to synthesize branched RNAs that correspond to selection of an improper 5'-splice site, the DNA sequence of paired region P4 (denoted here with Xs) is chosen to maintain base pairing with the RNA sequence near the 5'-splice site. **c)** The key components of the branch-formation reaction, shown in greater atomic detail.

downstream of the correct 5'-splice site, this requires “trimming” of the incorrect 5'-exon by removal of the extra nucleotide(s), followed by the second step. Such trimming of the 5'-exon could correspond to the well-known hydrolytic first step of splicing, in which water instead of the branch-site adenosine acts as the nucleophile that attacks the 5'-splice site (4, 10–12). Alternatively, the trimming reaction could instead correspond to the hydrolytic reverse of the second step of splicing (11, 13), analogous to spliced exons reopening (SER) (14–16). Hydrolysis has been observed for short RNA oligonucleotides that are analogues of the 5'-exon (7, 8, 13). However, catalysis of oligonucleotide hydrolysis and the subsequent second step have never been observed by mis-spliced intron RNAs, because the required mis-spliced RNAs could not be synthesized. Therefore, a direct test of this second proofreading mechanism for 5'-splice site selection has not been achievable.

Our laboratory has identified many artificial deoxyribozymes (DNA enzymes) (17) for RNA ligation (18, 19). Several of these deoxyribozymes such as 7S11 (20, 21) create 2',5'-branched RNA (Figure 1, panel b) (20–24), where a 2',5'-branch is the key structural element of the lariat RNAs that are the intermediates in biological RNA splicing (25). In reactions catalyzed by group II introns, branched RNA (which lacks the closed loop of a lariat) is as functionally competent as lariat RNA (4, 26). Here, we have used synthetic branched RNAs prepared by the 7S11 deoxyribozyme to allow explicit experimental tests of the two proofreading mechanisms for 5'-splice site selection. Mis-spliced RNAs are observed not to undergo the reverse of the first step of splicing, which provides evidence against the first proofreading mechanism. In contrast, direct evidence is obtained for the operation of the second proofreading mechanism.

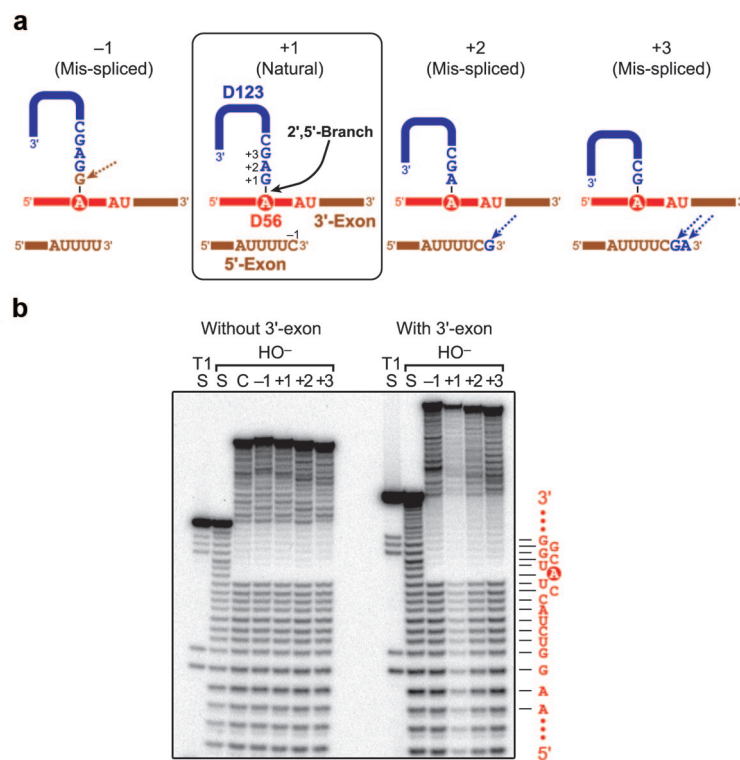


Figure 2. Correctly spliced and mis-spliced ai5 γ branched RNAs, and verification of their structures. **a)** The ai5 γ bipartite D123/D56 system, showing correctly spliced (box) and mis-spliced branch–3'-exon intermediates after the first step of splicing. Before the first step, the 5'-exon is joined at its 3'-end to the 5'-end of D123. The 5'-splice site locations are denoted +1 for the natural splice site (see boxed structure) and either –1, +2, or +3 for an improper splice site. The branch-site adenosine nucleotide is circled. Nucleotides found at the incorrect position within each mis-spliced RNA are marked with dashed arrows. For the –1 mis-spliced RNA, the final nucleotide of the original 5'-exon sequence, which is connected directly to the branch-site adenosine after the first step, has been changed from C to G (see text for details). **b)** Partial alkaline hydrolysis to verify the structures of the ai5 γ branched RNAs synthesized using the 7S11 deoxyribozyme (20% PAGE; T1 denotes G-specific RNase T1 digestion and S = standard for ladder calibration). “C” denotes a branched RNA synthesized using the conventional ai5 γ forward splicing reaction. Prior to alkaline hydrolysis, the core of each branched RNA was excised using two 10–23 deoxyribozymes (38, 39).

RESULTS AND DISCUSSION

Choice of the Splicing System. We used the ai5 γ group II intron RNA (27, 28) as the self-splicing RNA to test the first proofreading mechanism for 5'-splice site selection. The ai5 γ RNA is a common model system for understanding group II intron structure and function, and it has been studied using several biochemical approaches (4). One of the most useful approaches to assemble ai5 γ is with the bipartite D123/D56 system

(Figure 2, panel a, shown schematically after the first step of splicing; see boxed structure for correctly spliced RNA). In this bipartite system, the branch-site adenosine nucleotide of domain 6 is located within the 77-nt D56 RNA that comprises domains 5 and 6, plus the 3'-exon (if any is included; the first splicing step does not require the 3'-exon). The 5'-exon is joined to the 5'-end of domains 1–3 (D123), which encompasses 675 nt. A key advantage of studying ai5 γ splicing is that a lariat topology of D123/D56 is not necessary, because both splicing steps proceed well in the simpler 2',5'-branched RNA (4, 26).

Synthesis of Mis-Spliced ai5 γ Variants by the 7S11 Deoxyribozyme.

Experimentally testing the proofreading mechanisms required synthesis of 2',5'-branched RNAs that correspond to the mis-spliced intermediates after the first step of splicing. The 7S11 deoxyribozyme synthesizes branched RNA from two RNA substrates with few restrictions on their nucleotide sequences (20, 21). To prepare ai5 γ variants using 7S11, the RNA substrate providing the 2'-hydroxyl nucleophile was a T7 polymerase transcript corre-

sponding to D56, and the RNA substrate providing the 5'-portion of the 2',5'-linkage was a D123 transcript (Figure 1, panel b).

Using the 7S11 deoxyribozyme, the ai5 γ branched intermediate with the natural (correct) spliced sequence was prepared along with three sequence variants. Each variant represents a mis-spliced RNA that would be formed if an improper 5'-splice site phosphodiester linkage near the 5'-end of D123 were attacked by the

branch-site 2'-hydroxyl group from D6 (Figure 2, panel a, three unboxed structures). The RNA sequence near the natural 5'-splice site is 5'...CpGAGC...3', where the G is the first nucleotide of the intron (D123 domain) and p denotes the natural 5'-splice site phosphodiester linkage. This natural 5'-splice site is conventionally designated as position +1, with the +1 value pertaining to the nucleotide whose 5'-hydroxyl is attached to the splice site phosphorus atom (*i.e.*, G). The nucleotides one or two positions to the 3'-side of the natural 5'-splice site (here, A and G) are denoted +2 and +3, respectively. The +1, +2, and +3 ai5 γ intermediates were prepared using the 7S11 deoxyribozyme with appropriate RNA substrate sequences that have entirely wild-type nucleotides. The intermediate from splicing at a position one nucleotide to the 5'-side of the natural splice site (*i.e.*, at position -1) was also prepared using 7S11. To accommodate the modest substrate sequence requirements of 7S11, which requires a purine as the branch-site nucleotide (20, 21), the -1 nucleotide was changed from C to G. To compensate, the corresponding base-paired nucleotide within the exon binding site (EBS1) of D123 was changed from G to C. Consistent with expectations (29, 30), this C=G base-pair flip did not substantially affect either the yield or the rate of the forward first step of splicing (Supplementary Figure 1).

We focused on this small window of phosphodiester linkages near the natural 5'-splice site (-1, +1, +2, and +3) because the -1 nucleotide is already within IBS1 that is required for 5'-splice site fidelity (1, 7, 8, 31) and because the +3 nucleotide is involved directly in the ε - ε' tertiary interaction (32). Therefore, it seems unlikely that studying improper 5'-splice site selection outside of this window will be biologically meaningful, because the mis-spliced intron will be missing crucial key components. It should be noted that all six nucleotides at the start of domain 1 (nucleotides +1 through +6) are highly conserved and may contribute to catalytic activity (33, 34).

For all four splice site variants (-1, +1, +2, and +3), the branched RNA intermediate was synthesized either omitting or including a short (6-nt) 3'-exon sequence. Therefore, $4 \times 2 = 8$ branched RNAs were prepared. Only when the 3'-exon is omitted can the first step of splicing be examined in either direction without intervention of the second step. As a positive control (denoted "C"), the natural (+1) branched intermediate without 3'-exon was also synthesized using the

conventional forward splicing reaction of the ai5 γ RNA without using the 7S11 deoxyribozyme. Of course, the control branch that includes the 3'-exon could not be synthesized in more than trace amounts using ai5 γ catalysis, because the second step of splicing is rapid (1, 35-37).

This study is the first in which we report direct biochemical applications of branched RNAs that are synthesized by deoxyribozymes. Therefore, we empirically established the nucleotide connectivities of all branched ai5 γ intermediates using partial alkaline hydrolysis, after initially excising the branched core using two 10-23 deoxyribozymes (38, 39). For all eight synthetic branches as well as for the "C" control, the anticipated branch sites were confirmed unambiguously (Figure 2, panel b).

Assaying Mis-Spliced RNAs for the Reverse of the First Step of Splicing (First Proofreading Mechanism).

When the newly prepared synthetic mis-spliced RNAs were used, it was straightforward to test the first proofreading mechanism, which requires that mis-spliced RNAs undergo the reverse of the first step of splicing. Each of the four internally ^{32}P -radiolabeled synthetic branched RNA variants without the 3'-exon (and the "C" control) was separately incubated with an excess of the matched unradiolabeled 5'-exon oligonucleotide (the 5'-exon is "matched" because reverse splicing would restore the original unspliced RNA sequence). If reverse splicing occurred at all, it was readily detected by a gel shift (Figure 3). To prevent the second step of splicing from occurring, the 3'-exon was not present on D56 in these experiments. When the branched intermediate has the natural 5'-splice site, reverse splicing is observed, with equivalent efficiency for both the control ("C") and synthetic (+1) branched RNA. In contrast, reverse splicing fails to occur (<2%) when the branched intermediate has the incorrect (-1, +2, or +3) 5'-splice site and the matched 5'-exon is provided. In this direct and well-controlled test, the results do not support the first proofreading mechanism.

Assaying Mis-Spliced RNAs for the Second Step of Splicing (Second Proofreading Mechanism).

In the absence of the 3'-exon sequence, the only possible reaction of the branched RNA is the reverse of the first step, as described above. However, when the 3'-exon is included with D56, the second step of splicing can also occur by attack of the 5'-exon at the 3'-splice site of the branched RNA. The availability of synthetic mis-spliced

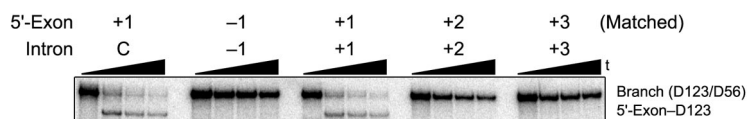


Figure 3. Assays for the reverse of the first step of splicing show no reaction for the mis-spliced RNAs, which is evidence against the first proofreading mechanism for 5'-splice site selection. The internally ^{32}P -radiolabeled branched RNA without 3'-exon was the limiting reagent (unradiolabeled 5'-exon in 200-fold excess; $t = 0, 0.5, 1,$ and 2 h ; 5% PAGE). The control "C" used the +1 splice site and was prepared by forward splicing using the catalytic activity of ai5 γ itself. The other four branched RNAs were prepared using the 7S11 deoxyribozyme (Figure 1, panel b). Consistent results were observed with analogous assays that used either an excess of radiolabeled 5'-exon or an excess of branched RNA (Supplementary Figures 2 and 3). The overall loss of signal intensity with time has been observed numerous times by others (15, 16, 57–60).

branched RNAs allowed us to examine whether the mis-spliced branched RNAs are competent for the second step of splicing, independent of their ability to undergo the reverse of the first step.

We first examined the competence of the mis-spliced branch–3'-exon RNAs to proceed through the second step using the natural (+1) 5'-exon in all cases (Figure 4, panel a). The primary purpose of this assay was to provide baseline information on the catalytic competence of the mis-spliced branches, before testing them with their matched 5'-exons. The –1, +1, and +2 branches were found to catalyze the second step reasonably well, whereas the +3 branch does not. The 10–20% efficiency observed for the natural +1 reaction is similar to that observed for related systems (10, 29).

We then tested each of the four synthetic branch–3'-exon RNAs (–1, +1, +2, and +3) by incubating each RNA with a limiting amount of the matched 5'-exon (Figure 4, panel b). In the case of the natural (+1) branched RNA and matched natural 5'-exon, normal second-step splicing activity was found (as also shown in

Figure 4, panel a). For the +2 mis-spliced RNA, the matched (+2) 5'-exon, which is longer than the natural (+1) 5'-exon at its 3'-end by one nucleotide, was rapidly and efficiently "trimmed" with loss of a single nucleotide to the correct length. This trimmed 5'-exon, now equivalent in length to the +1 5'-exon, then proceeded through the second step of splicing with 2–3% final splicing yield, versus 10–20% yield for the +1 RNA. This indicates that the efficiency of this second proofreading mechanism is on the order of 10–30% relative to formation of correctly spliced exons by the properly branched RNA intermediate. Therefore, for the +2 mis-spliced RNA, the correctly ligated exons can indeed be formed via the second proofreading mechanism. The sequence of the correctly ligated exons was verified by RT-PCR and sequencing of the splice junction (see Supporting Infor-

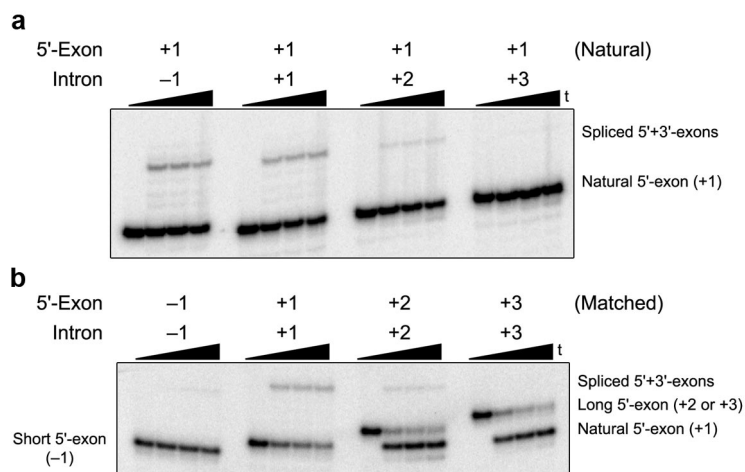


Figure 4. Assays for the second step of splicing directly demonstrate the operation of the second proofreading mechanism for 5'-splice site selection using RNAs mis-spliced at the 5'-splice site. a) Using branch–3'-exon RNA and natural-length (+1) 5'-exon. b) Using branch–3'-exon RNA and matched 5'-exon. In both sets of experiments, the 5'-exon was ^{32}P -radiolabeled, and a 10-fold excess of unradiolabeled branched RNA was added ($t = 0, 0.5, 1,$ and 2 h ; 20% PAGE). When expected, the bands at the top of each gel near the wells showed substantial reverse first-step splicing in competition with the second step (Supplementary Figure 4). The overall loss of signal intensity with time has been observed by others (15, 16, 57–60).

mation). As revealed by data points taken at shorter times (not shown), the second-step splicing product was formed without any detectable induction period relative to formation of the trimmed 5'-exon. This suggests that for the +2 mis-spliced intron RNA, the second step is fast (*i.e.*, not rate-determining), which is also true for the correctly spliced intron RNA (1, 35–37). For the +3 mis-spliced branch, 5'-exon trimming was observed. However, the subsequent second-step efficiency was too low to be detected, as expected on the basis of the analogous experiment with the +1 branch (in Figure 4, panel a; this may be due to the participation of the +3 nucleotide in the ϵ - ϵ' tertiary interaction (32), which in the +3 mis-spliced RNA may be positioned suboptimally). For the -1 mis-spliced 5'-exon, no trimming is possible because the 5'-exon is already too short by one nucleotide, and almost no second-step product is observed.

Implications of the Experimental Data for the Two Proofreading Mechanisms. Using synthetic branched RNAs prepared by the 7S11 deoxyribozyme (Figure 2), we have tested the long-standing hypothesis (5) that reversibility of the first step of group II intron splicing acts as a proofreading mechanism for 5'-splice site selection (6). Because first-step reversibility was not observed for any of the three mis-spliced RNAs that were tested (Figure 3), our data are evidence against this first proofreading mechanism. In contrast, direct evidence was obtained for operation of the alternative, second proofreading mechanism catalyzed by mis-spliced RNA, in which 5'-exon hydrolysis is followed by the second step of splicing (Figure 4). A summary of the experimental observations highlighting both proofreading mechanisms and their relationship is provided (Figure 5).

Our evidence against first-step reversibility is inherently negative evidence; that is, reversibility was not observed (Figure 3). Complicating factors could contribute to this lack of reversibility. For example, a synthetic mis-spliced RNA prepared using a deoxyribozyme might adopt a conformation that does not permit the reverse of the first step, in contrast to a mis-spliced RNA that was formed through the forward operation of the first step using the improper 5'-splice site. It is also possible that protein cofactors which are not required *in vitro* for the forward first step of splicing might nonetheless be required for the reverse of the first step by synthetic mis-spliced RNAs. Such possibilities

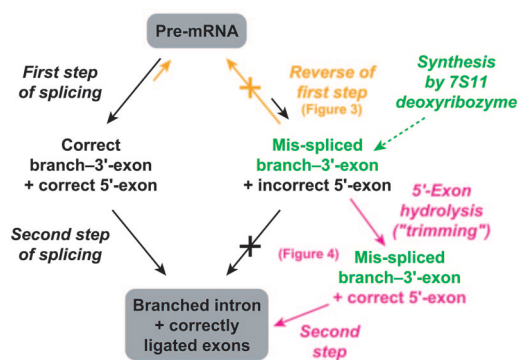


Figure 5. Summary of the experimental observations, highlighting the first and second proofreading mechanisms (orange and magenta, respectively) and synthesis of the key mis-spliced branch-3'-exon RNAs by the 7S11 deoxyribozyme (green).

are intrinsically challenging (if not impossible) to disprove. Instead, we emphasize that our experiments do *not* have the simplest outcome that would be expected if first-step reversibility were to act as a proofreading mechanism for 5'-splice site selection.

In sharp contrast, we have provided positive evidence for the second proofreading mechanism, which is “trimming” of the 5'-exon by hydrolysis (7, 8, 13) followed by the second step. In particular, for the +2 mis-spliced RNA (where the nucleotide one position to the 3'-side of the natural 5'-splice site is the site of branching), both of these reaction steps were directly observed (Figure 4). Therefore, we have provided unambiguous evidence for the *in vitro* operation of this second proofreading mechanism. All previous studies, which could not use mis-spliced RNAs, merely suggested that such a mechanism might be applicable for RNAs that are mis-spliced at the 5'-splice site (7, 8, 13). The efficiency of the second proofreading mechanism is relatively modest (on the order of 10–30%), but this level of proofreading may plausibly be biologically relevant. In addition, protein cofactors could potentially be required *in vivo* for optimal proofreading efficiency.

Natural Roles of First-Step Reversibility and of the Second Proofreading Mechanism. Improper selection of the 5'-splice site in group II introns is rare because the IBS1–EBS1 interaction strictly enforces proper splice-site selection (1, 7, 8, 31). This suggests that an efficient first-step proofreading mechanism may be unnecessary in nature; indeed, our results directly show that first-step reversibility does not contribute to the fidelity of

5'-splice site selection (6). We agree with the previous suggestion (6) that first-step reversibility likely evolved to enable retrotransposition (40, 41), and not for any proofreading function. In contrast to the irreversibility of the first step for the mis-spliced RNAs, our data directly demonstrate the second proofreading mechanism for improper 5'-splice site selection (Figure 4). This second proofreading mechanism may operate in those rare cases for which improper 5'-splice site selection has indeed occurred, at least in cases where mis-splicing has occurred to the 3'-side of the natural 5'-splice site. We do not observe the second step of splicing using a mis-spliced 5'-exon unless 5'-exon hydrolysis (trimming) occurs first. This could be important biologically because the second step of splicing using an untrimmed mis-spliced 5'-exon would give improperly ligated exons and therefore an incorrect transcript. Together, the combination of trimming and the second step constitutes the second proofreading mechanism,

which we have now shown directly to occur for RNAs that are mis-spliced at the 5'-splice site.

Implications for Spliceosomal RNA Processing.

Spliceosomal pre-mRNA processing (25, 42) appears to be mechanistically related to group II intron splicing (2, 16, 43–45). Although reversibility of the first step has not been explicitly demonstrated in the spliceosome, such reversibility is likely on the basis of the relationship of the spliceosome to group II introns. The data presented here therefore suggest that reversibility of the first step is not a proofreading mechanism in pre-mRNA splicing. Proper splice site selection during pre-mRNA splicing requires many protein and RNA factors (46–49), which suggests that proofreading mechanisms would be valuable. In the absence of first-step reversibility, we suggest that the second proofreading mechanism for 5'-splice site selection (involving 5'-exon hydrolysis followed by the second step of splicing) may be operative within the spliceosome.

METHODS

DNA Oligonucleotides and RNA Transcripts. All nucleic acid samples were purified by PAGE as described previously (18, 50). The natural and mis-spliced 5'-exons were prepared by *in vitro* transcription using T7 RNA polymerase with an appropriate double-stranded DNA template (51). The natural 5'-exon was 5'-GCGUGGUGGGACAUUUUC-3', where IBS2 is bold and IBS1 is underlined. Each D123 RNA was transcribed using an Ear-linearized plasmid derived from plasmid pJD20 (52) as the template. Each D56 RNA (with or without 3'-exon) was transcribed using a double-stranded PCR product derived from plasmid pJD20 as the template. A homogeneous 3'-terminus for D56 (with or without the 3'-exon) was provided by intramolecular hepatitis delta virus ribozyme cleavage (53, 54). When the 3'-exon was not included at the 3'-terminus of D56, the 2',3'-cyclic phosphate was removed with T4 polynucleotide kinase in the absence of ATP (55). 5'-³²P-Radiolabeling was performed with γ -[³²P]-ATP and T4 polynucleotide kinase. Internal ³²P-radiolabeling was performed by including in the transcription solution a trace amount of α -[³²P]-CTP along with all four unradiolabeled NTPs, each at 4 mM.

Synthesis of Branched RNA. The branched RNAs were synthesized using the 7S11 deoxyribozyme, with the enzyme region sequence as described previously and the binding arms complementary to appropriate portions of the RNA substrates (20, 21). Disruptor oligonucleotides (56) were required for both D123 and D56 to enable proper binding of 7S11 to the RNA substrates (see Supporting Information for details). The D123 substrate (20 pmol), deoxyribozyme (30 pmol), D56 substrate (with or without 3'-exon; 60 pmol), and disruptor oligonucleotides (200 pmol of D123 disruptor and 600 pmol of D56 disruptor) were combined in 11.8 μ L of 5 mM HEPES (pH 7.5), 15 mM NaCl, and 0.1 mM EDTA. The sample was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The sample volume was raised to 20 μ L with final concentrations of 50 mM EPPS (pH 8.5), 150 mM NaCl, 2 mM KCl, and 40 mM MgCl₂. The

sample was incubated at 37 °C for 1.5 h, then quenched onto 25 μ L of stop solution (80% formamide, 1 \times TB [89 mM each Tris and boric acid, pH 8.3], 50 mM EDTA, and 0.025% each bromophenol blue and xylene cyanol). Samples were purified by 5% denaturing PAGE.

To prepare the control ("C") branched RNA by ai5 γ catalysis to study the reverse of the first step of splicing, 5'-exon-D123 (293 + 675 = 968 nt; 20 pmol) and D56-3'-exon (77 + 180 = 257 nt; 400 pmol) were combined in 64 μ L of 5 mM MOPS (pH 7.0) and 1 mM EDTA and annealed by heating at 95 °C for 1 min and cooling at room temperature for 1 min. The sample volume was then raised to 80 μ L with final concentrations of 40 mM MOPS (pH 7.0), 100 mM MgCl₂, and 500 mM NH₄Cl. The sample was incubated at 45 °C for 1 h, quenched onto 80 μ L of stop solution, and purified by 5% denaturing PAGE.

Verification of Branched RNA Structures. The core of each branched RNA (45 nt for +1 RNA) was prepared by two sequential 10–23 deoxyribozyme cleavage reactions of the large branched RNA (752 nt for +1 RNA). Then, the location of each branch site was determined by partial alkaline hydrolysis of the small branched RNA as previously described (22, 39).

Splicing Assays. For the assay of Figure 3, internally ³²P-radiolabeled branched RNA (0.1 pmol) and unradiolabeled matched 5'-exon (20 pmol) were combined in 16 μ L of 5 mM MOPS (pH 7.0) and 0.1 mM EDTA and annealed by heating at 95 °C for 1 min and cooling at room temperature for 1 min. The sample volume was raised to 20 μ L with final concentrations of 40 mM MOPS (pH 7.0), 100 mM MgCl₂, and 500 mM NH₄Cl. The sample was incubated at 45 °C. Aliquots were withdrawn at desired time points, quenched onto stop solution, and separated by 5% denaturing PAGE. Gel images were acquired using a PhosphorImager. Similar results were observed using an alternative high-salt assay buffer (26) of 40 mM HEPES (pH 7.5), 100 mM MgSO₄, and 500 mM (NH₄)₂SO₄ (data not shown). For the assays of Figure 4, internally ³²P-radiolabeled branched RNA (0.2 pmol) and 5'-³²P-radiolabeled matched 5'-exon

(0.005 pmol) were assayed as described above (20% denaturing PAGE).

Acknowledgment: This research was supported by the Burroughs Wellcome Fund (New Investigator Award in the Basic Pharmacological Sciences), the March of Dimes Birth Defects Foundation (Research Grant No. 5-FY02-271), the National Institutes of Health (GM-65966), the American Chemical Society Petroleum Research Fund (38803-G4), and the UIUC Department of Chemistry (all to S.K.S.). S.K.S. is the recipient of a fellowship from The David and Lucile Packard Foundation. We thank P. Perlman for providing us with the pJD20 (ai5 γ) plasmid. We also thank members of the Silverman lab for discussions.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Michel, F., and Ferat, J. L. (1995) Structure and activities of group II introns, *Annu. Rev. Biochem.* **64**, 435–461.
2. Jacquier, A. (1996) Group II introns: elaborate ribozymes, *Biochimie* **78**, 474–487.
3. Martinez-Abarca, F., and Toro, N. (2000) Group II introns in the bacterial world, *Mol. Microbiol.* **38**, 917–926.
4. Fedorova, O., Su, L. J., and Pyle, A. M. (2002) Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions, *Methods* **28**, 323–335.
5. Lehmann, K., and Schmidt, U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes, *Crit. Rev. Biochem. Mol. Biol.* **38**, 249–303.
6. Chin, K., and Pyle, A. M. (1995) Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection, *RNA* **1**, 391–406.
7. Xiang, Q., Qin, P. Z., Michels, W. J., Freeland, K., and Pyle, A. M. (1998) Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets, *Biochemistry* **37**, 3839–3849.
8. Su, L. J., Qin, P. Z., Michels, W. J., and Pyle, A. M. (2001) Guiding ribozyme cleavage through motif recognition: the mechanism of cleavage site selection by a group II intron ribozyme, *J. Mol. Biol.* **306**, 655–668.
9. Soller, M. (2006) Pre-messenger RNA processing and its regulation: a genomic perspective, *Cell. Mol. Life Sci.* **63**, 796–819.
10. Jacquier, A., and Rosbash, M. (1986) Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon-intron interaction, *Science* **234**, 1099–1104.
11. Podar, M., Perlman, P. S., and Padgett, R. A. (1995) Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions, *Mol. Cell. Biol.* **15**, 4466–4478.
12. Vogel, J., and Borner, T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing, *EMBO J.* **21**, 3794–3803.
13. Michels, W. J., Jr., and Pyle, A. M. (1995) Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships, *Biochemistry* **34**, 2965–2977.
14. Jarrell, K. A., Peebles, C. L., Dietrich, R. C., Romiti, S. L., and Perlman, P. S. (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products, *J. Biol. Chem.* **263**, 3432–3439.
15. Koch, J. L., Boulanger, S. C., Dib-Hajji, S. D., Hebbar, S. K., and Perlman, P. S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity, *Mol. Cell. Biol.* **12**, 1950–1958.
16. Sontheimer, E. J., Gordon, P. M., and Piccirilli, J. A. (1999) Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome, *Genes Dev.* **13**, 1729–1741.
17. Silverman, S. K. (2004) Deoxyribozymes: DNA catalysts for bioorganic chemistry, *Org. Biomol. Chem.* **2**, 2701–2706.
18. Flynn-Charlebois, A., Wang, Y., Prior, T. K., Rashid, I., Hoadley, K. A., Coppins, R. L., Wolf, A. C., and Silverman, S. K. (2003) Deoxyribozymes with 2'-5' RNA ligase activity, *J. Am. Chem. Soc.* **125**, 2444–2454.
19. Purtha, W. E., Coppins, R. L., Smalley, M. K., and Silverman, S. K. (2005) General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages, *J. Am. Chem. Soc.* **127**, 13124–13125.
20. Coppins, R. L., and Silverman, S. K. (2004) A DNA enzyme that mimics the first step of RNA splicing, *Nat. Struct. Mol. Biol.* **11**, 270–274.
21. Coppins, R. L., and Silverman, S. K. (2005) A deoxyribozyme that forms a three-helix-junction complex with its RNA substrates and has general RNA branch-forming activity, *J. Am. Chem. Soc.* **127**, 2900–2907.
22. Wang, Y., and Silverman, S. K. (2003) Deoxyribozymes that synthesize branched and lariat RNA, *J. Am. Chem. Soc.* **125**, 6880–6881.
23. Pratico, E. D., Wang, Y., and Silverman, S. K. (2005) A deoxyribozyme that synthesizes 2',5'-branched RNA with any branch-site nucleotide, *Nucleic Acids Res.* **33**, 3503–3512.
24. Wang, Y., and Silverman, S. K. (2005) Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme, *Angew. Chem., Int. Ed.* **44**, 5863–5866.
25. Burge, C. B., Tuschl, T., and Sharp, P. A. (1999) In *The RNA World*, 2nd ed. (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 525–560, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Perlman, P. S., and Podar, M. (1996) Reactions catalyzed by group II introns in vitro, *Methods Enzymol.* **264**, 66–86.
27. Peebles, C. L., Perlman, P. S., Mecklenburg, K. L., Petrillo, M. L., Tabor, J. H., Jarrell, K. A., and Cheng, H. L. (1986) A self-splicing RNA excises an intron lariat, *Cell* **44**, 213–223.
28. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A., and Macino, G. (1980) Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase, *J. Biol. Chem.* **255**, 11927–11941.
29. Jacquier, A., and Michel, F. (1987) Multiple exon-binding sites in class II self-splicing introns, *Cell* **50**, 17–29.
30. Michel, F., and Jacquier, A. (1987) Long-range intron-exon and intron-intron pairings involved in self-splicing of class II catalytic introns, *Cold Spring Harbor Symp. Quant. Biol.* **52**, 201–212.
31. Jacquier, A., and Jacquesson-Breuleux, N. (1991) Splice site selection and role of the lariat in a group II intron, *J. Mol. Biol.* **219**, 415–428.
32. Jacquier, A., and Michel, F. (1990) Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns, *J. Mol. Biol.* **213**, 437–447.
33. Peebles, C. L., Belcher, S. M., Zhang, M., Dietrich, R. C., and Perlman, P. S. (1993) Mutation of the conserved first nucleotide of a group II intron from yeast mitochondrial DNA reduces the rate but allows accurate splicing, *J. Biol. Chem.* **268**, 11929–11938.
34. Boudvillain, M., and Pyle, A. M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis, *EMBO J.* **17**, 7091–7104.
35. Mikheeva, S., Murray, H. L., Zhou, H., Turczyk, B. M., and Jarrell, K. A. (2000) Deletion of a conserved dinucleotide inhibits the second step of group II intron splicing, *RNA* **6**, 1509–1515.
36. Bar-Shalom, A., and Moore, M. J. (2000) Tri-partite assay for studying exon ligation by the ai5 γ group II intron, *Biochemistry* **39**, 10207–10218.
37. Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000) Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2'-OH in catalysis, *Biochemistry* **39**, 12939–12952.
38. Santoro, S. W., and Joyce, G. F. (1997) A general purpose RNA-cleaving DNA enzyme, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4262–4266.

39. Pyle, A. M., Chu, V. T., Jankowsky, E., and Boudvillain, M. (2000) Using DNAzymes to cut, process, and map RNA molecules for structural studies or modification, *Methods Enzymol.* **317**, 140–146.
40. Lambowitz, A. M., and Belfort, M. (1993) Introns as mobile genetic elements, *Annu. Rev. Biochem.* **62**, 587–622.
41. Lambowitz, A. M., and Zimmerly, S. (2004) Mobile group II introns, *Annu. Rev. Genet.* **38**, 1–35.
42. Collins, C. A., and Guthrie, C. (2000) The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.* **7**, 850–854.
43. Weiner, A. M. (1993) mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? *Cell* **72**, 161–164.
44. Gaur, R. K., McLaughlin, L. W., and Green, M. R. (1997) Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron, *RNA* **3**, 861–869.
45. Newman, A. (1997) RNA splicing: out of the loop, *Curr. Biol.* **7**, R418–420.
46. Krämer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing, *Annu. Rev. Biochem.* **65**, 367–409.
47. Reed, R. (2000) Mechanisms of fidelity in pre-mRNA splicing, *Curr. Opin. Cell Biol.* **12**, 340–345.
48. Brow, D. A. (2002) Allosteric cascade of spliceosome activation, *Annu. Rev. Genet.* **36**, 333–360.
49. Jurica, M. S., and Moore, M. J. (2003) Pre-mRNA splicing: awash in a sea of proteins, *Mol. Cell* **12**, 5–14.
50. Wang, Y., and Silverman, S. K. (2003) Characterization of deoxyribozymes that synthesize branched RNA, *Biochemistry* **42**, 15252–15263.
51. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res.* **15**, 8783–8798.
52. Jarrell, K. A., Dietrich, R. C., and Perlman, P. S. (1988) Group II intron domain 5 facilitates a trans-splicing reaction, *Mol. Cell. Biol.* **8**, 2361–2366.
53. Grosshans, C. A., and Cech, T. R. (1991) A hammerhead ribozyme allows synthesis of a new form of the *Tetrahymena* ribozyme homogeneous in length with a 3' end blocked for transesterification, *Nucleic Acids Res.* **19**, 3875–3880.
54. Ferré-D'Amaré, A. R., and Doudna, J. A. (1996) Use of *cis*- and *trans*-ribozymes to remove 5' and 3' heterogeneities from milligrams of in vitro transcribed RNA, *Nucleic Acids Res.* **24**, 977–978.
55. Miduturu, C. V., and Silverman, S. K. (2005) DNA constraints allow rational control of macromolecular conformation, *J. Am. Chem. Soc.* **127**, 10144–10145.
56. Wang, Y., and Silverman, S. K. (2006) A general two-step strategy to synthesize lariat RNAs, *RNA* **12**, 313–321.
57. Kwakman, J. H., Konings, D., Pel, H. J., and Grivell, L. A. (1989) Structure-function relationships in a self-splicing group II intron: a large part of domain II of the mitochondrial intron *al5* is not essential for self-splicing, *Nucleic Acids Res.* **17**, 4205–4216.
58. Schmidt, U., Riederer, B., Morl, M., Schmelzer, C., and Stahl, U. (1990) Self-splicing of the mobile group II intron of the filamentous fungus *Podospora anserina* (COI 1) in vitro, *EMBO J.* **9**, 2289–2298.
59. Hebbar, S. K., Belcher, S. M., and Perlman, P. S. (1992) A maturase-encoding group IIA intron of yeast mitochondria self-splices in vitro, *Nucleic Acids Res.* **20**, 1747–1754.
60. Fedorova, O., Mitros, T., and Pyle, A. M. (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site, *J. Mol. Biol.* **330**, 197–209.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

H. Ulrich Stiltz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Experimenting with Peer Review

The current peer-review system began in the late 1800s. More than a century later, we at *ACS Chemical Biology* (*ACS CB*) and many of our colleagues in the publishing world are discussing how the system will evolve. The process of peer review is much like the ongoing discussions that take place in small communities, such as journal clubs, and among a larger group of scientists at conferences. In both cases, data are presented, questions are asked, and points are clarified with the expectation that the discourse will improve the science. In effect, peer review is an ongoing conversation between reviewers and authors. The emergence of the Internet and electronic publishing is changing the way we disseminate, discuss, and sanction scientific content. Can the new electronic tools be used to improve the quality of the review process and provide a more fruitful interaction among reviewers, authors, and readers? Answers might come from recent experiments with alternative open peer-review systems.

Single-blind peer review (SBPR), in which the reviewer knows the identity of the author but not *vice versa*, took hold in the 1840s (1). The editorial board at *Philosophical Transaction*, a journal of the Royal Society of London, was asking scientists to read and comment on submitted papers. This system was only used when the editorial board lacked the expertise to evaluate the manuscript. In 1918, Arthur Lamb, the chief editor of the *Journal of the American Chemical Society* (*JACS*), instituted a system that he hoped would remove perceived unfairness in the review process. Instead of using a select group of editorial board members to evaluate submitted manuscripts, Lamb sent papers to experts in the field for anonymous review. This peer-review system persevered at *JACS* and is now the norm for >20,000 scientific journals.

What is the purpose of peer review? It depends on whom you ask. It provides feedback and sometimes discussion among scientists. It also adds incentive for the authors of the paper to improve their manuscript, because their peers suggest they do so. Other scientists note that peer review also serves as a gatekeeper: it ensures that the published work meets certain criteria. In the current system, the anonymity of the reviewer seems necessary: it allows him or her the freedom to be honest and comment on whether an article should even be published. Lastly, some researchers contend that the current peer-review system can show university administrators that the work of a scientist merits his or her receiving tenure and promotion. Most scientists agree that SBPR is imperfect but that it works most of the time.

If peer review is meant to be a conversation within the scientific community in which reviewers engage in a debate with authors, then would the review process benefit from more back-and-forth discussion among reviewers, between reviewers and authors, or even among disparate peers in an open public-review system? Extensive discourse will no doubt improve the paper. On the other hand, lengthy conversations can have diminishing returns (and delay publication). The publishing world requires more data before we can begin to improve the current peer-review process.

In the scientific spirit of experimentation, some journals (*Biology Direct*, *Atmospheric Chemistry and Physics*, *Electronic Transactions in Artificial Intelligence*, to name a few) have adopted open peer-review systems in which the identities of authors and reviewers are disclosed to all participants. *Nature* is trying a variation on this theme by running a side-by-

side experiment with traditional peer review and open peer commentary on a select number of papers. Authors can choose to participate in the experiment or have their paper reviewed through traditional SBPR. Later this year, the Public Library of Science (PLOS) will launch *PLOS ONE*, its experiment in open peer review. Will scientists be willing to commit time to engage in a discourse on the web versus writing a single set of comments? Will technically deficient papers be weeded out quickly so that reviewers' time won't be wasted? If the paper is posted on a public website before publication, what measures will be taken to prevent plagiarism or outright theft of intellectual property? If an author excludes an individual as a reviewer, how will the journal honor this request? The data from these open peer-review experiments will be telling.

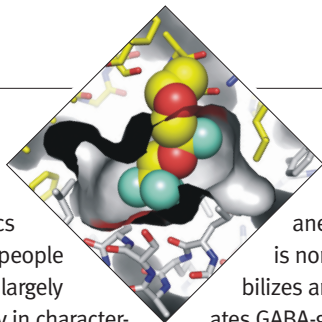
Here at *ACS CB*, we've been conducting a behind-the-scenes experiment with open editor discussions. As our readers can see, we receive papers that run the gamut of organic chemistry to cell biology. How do the editors decide what is appropriate for the journal? We use the collective expertise of the seven members of our Board of Editors and the staff editors to prescreen manuscripts. This process is an online discussion that takes 1–3 days and normally involves at least three editors. It is not a formal review but simply a prescreen to establish the suitability of a paper for the journal, determine whether it should be reviewed, and identify appropriate reviewers. We have considered whether a similar discussion forum among the chosen reviewers would enhance our review process. After the paper returns from review (usually 2–3 weeks), the editors revisit these discussions, this time with reviewer comments in hand. We integrate the expertise of reviewers and the editors to make a decision within 2–4 days. This synthesis during the review process would have been impractical before the Internet but is now relatively straightforward.

The jury is still out on how peer review will evolve, but it is clear that it will. Please email us with your comments and suggestions. Should *ACS CB* change its peer-review process? If so, how? Go to our WIKI (free and open to all scientists), and add your opinion to the *Experimenting with Peer Review* discussion. Thanks for your input.



Evelyn Jabri
Executive Editor

1. Kronick, D. A. (1990) Peer review in 18th-century scientific journalism, *JAMA, J. Am. Med. Assoc.* 263, 1321–1322.



Knocked Out!

The ability of anesthetics to reversibly knock out people is not well understood, largely because of the difficulty in characterizing the protein targets of these low-affinity, volatile compounds. All current clinical anesthetics are haloethers, and though some data are available on the mechanisms of older anesthetics, new tools are needed to probe the molecular mechanisms behind the haloethers. To this end, Xi *et al.* (p 377) have synthesized a photoactivatable haloether, H-diaziflurane, and solved the crystal structure of this compound bound to the protein apoferritin.

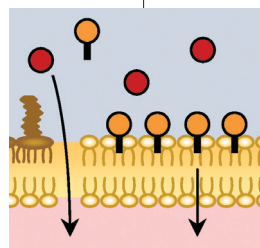
Functional characterization of H-diaziflurane confirmed that it was an adequate photolabel mimic of haloether anesthetics. Like other isoflurane-based

anesthetics, H-diaziflurane is nontoxic, reversibly immobilizes amphibians, and potentiates GABA-gated chloride channels in primary cultures of hippocampal neurons. Structural characterization using MS and X-ray crystallography of ultraviolet-exposed apoferritin/H-diaziflurane mixtures revealed that H-diaziflurane likely binds in a previously identified isoflurane-binding cavity in the ferritin L-chain. In addition, evidence was seen of adducts at two residues, the carbonyl oxygen of Leu24 and the side-chain oxygen of Ser27. Future studies with this compound and other isoflurane derivatives will facilitate identification and characterization of additional molecular targets of haloether anesthetics.

A Chemical Condition

Membrane biogenesis, the process by which the composition of cell membranes is determined, is a critical mechanism cells use to discriminate between the passage of essential nutrients and toxic small molecules. The outer membrane (OM) of Gram-negative bacteria has a unique composition, with phospholipids in the inner leaflet, glycolipids in the outer leaflet, and various integral OM β -barrel proteins. The factors involved in building the OM have been notoriously difficult to discern, in part because of the complexity of the environment and the components involved. Now, Ruiz *et al.* (p 385) use chemical conditionality, a genetic technique, to uncover how the OM β -barrel protein YaeT contributes to the discriminatory ability of the OM.

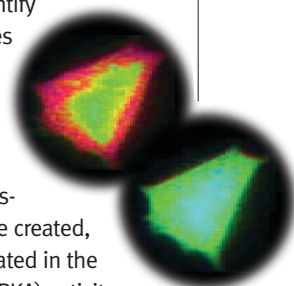
Chemical conditionality works by searching for mutations that alter the entry of toxic small molecules into cells. *imp4213*, a bacterial strain that is susceptible to passage of toxic molecules in part because of increased levels of phospholipids in the outer leaflet, is used to screen for suppressors of the permeability defect. A strain containing a mutation in the *yaeT* gene, *imp4213yaeT6*, was found to be resistant to two bile salts, sodium deoxycholate and sodium chenodeoxycholate, but the strain is not resistant to the more hydrophilic bile salt sodium cholate. Additional experiments led the researchers to propose that this mutation changes the protein landscape of the OM such that the phospholipid content of the outer leaflet is reduced. This leads to altered sensitivity to molecules on the basis of their hydrophobicity. A variety of techniques, including chemical conditionality, can be used to gain further understanding of how membrane components work together to modulate small-molecule entry and will help expand this model and provide additional details about the mechanism of membrane biogenesis.



Don't FRET over the Kinome

Kinases are key signaling proteins and an important class of drug targets, and methods to evaluate their activity are critical in understanding cellular signaling events and in drug-discovery efforts. One method to monitor kinase activity involves dynamic tracking using genetically encoded reporters, wherein a phosphorylation event induces a change in conformation that triggers a FRET response. These reporters provide sensitive ratiometric fluorescence readouts and enable precise molecular targeting and high spatiotemporal resolution, but thus far, these tools have only been utilized in single-cell imaging experiments. Now, Allen *et al.* (p 371) have adapted this reporter system for high-throughput screening of live cells; this approach can potentially be used to identify the modulators of most kinases in the cell.

Improved versions of two reporter systems, A-Kinase Activity Reporter and Indicator for cyclic adenosine monophosphate (cAMP) Using Epac, were created, and their utility was demonstrated in the detection of protein kinase A (PKA) activity and cAMP levels, respectively. In addition, a collection of 160 FDA-approved drugs and other clinically relevant compounds was screened, and three known PKA agonists, a known β -adrenergic antagonist, and a potential new cAMP modulator were identified. This reporter assay design has several attractive features, including parallel evaluation of key targets in the same signaling pathway and potential identification of both agonists and antagonists from one primary screen. Reporter systems can be created for any known kinase enabling mechanistic and kinetic studies of signaling pathways as well as discovery of new signaling modulators and drug candidates for essentially the entire kinome.



Published online July 21, 2006.

10.1021/cb600278v CCC: \$33.50

© 2006 by American Chemical Society

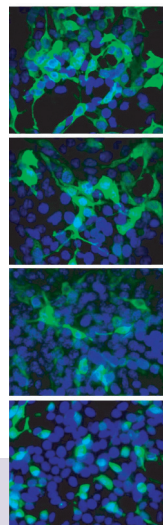
Spotlight

Evolving Inteins

An intein is a segment of a protein that can excise itself and ligate the remaining portions to create an intact protein. The presence of an intein within a target protein renders the macromolecule nonfunctional, but upon excision of the intein, the target can assume its native conformation. In yeast, directed evolution has been used to create inteins whose self-splicing is dependent on the small molecule 4-hydroxytamoxifen (4-HT). This method of controlling protein function has distinct advantages over other approaches, such as the use of non-native promoters or heterologous control elements, because it requires minimal disruption of the cell's regulatory networks. Now, Yuen *et al.* (*J. Am. Chem. Soc.* 2006, 128, 8939–8946) expand the application of evolved inteins to the manipulation of protein function in mammalian cells.

First, 4-HT-dependent intein splicing was characterized in mammalian cells transfected with a green fluorescent protein (GFP) intein variant. The researchers demonstrated that spliced, functional GFP was generated only in the presence of 4-HT in a dose- and time-dependent manner.

Next, evolved intein variants inserted into the murine transcription factors Gli1 and Gli3T were used to modulate transcription-factor activity. Transcription-factor activity was dependent on the addition of 4-HT, and spliced Gli1 and Gli3T localized to their native subcellular compartments. This finding provides additional evidence that spliced proteins behave similarly to their native counterparts. Finally, the researchers demonstrated that the evolved intein system



Reprinted with permission from *Journal of the American Chemical Society*

is capable of producing protein levels sufficient to fulfill the role of the native protein in the cell. Endogenous Gli1 can induce differentiation of mouse embryonic cells into osteoblasts. Remarkably, in the presence of 4-HT, mouse embryonic cells transfected with a Gli1-intein variant differentiated into osteoblasts. Further application of evolved, small-molecule-dependent inteins harbors great potential for the exploration of protein function in cells. **EG**

Manipulating Destruction

Cells maintain precise mechanisms to control their protein levels, including regulating protein expression, function, and degradation. The ability to manipulate these processes helps us understand how they operate and enables the design of engineered protein circuits. While many methods are available to control protein expression, protein degradation has proven

more difficult to manipulate. Now, McGinness *et al.* (*Mol. Cell* 2006, 22, 701–707) present a system to control destruction by exploiting properties of the adaptor protein SspB and the *ssrA*-tag, which mark proteins for degradation by the ClpXP protease.

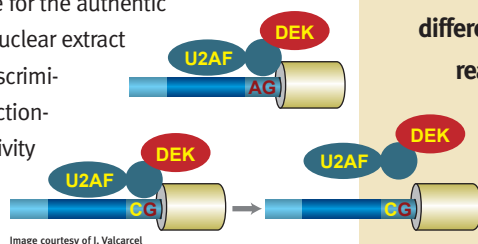
In *Escherichia coli*, proteins that contain the *ssrA* peptide sequence are targeted for destruction by the ClpXP protease, and

this process is facilitated by SspB. To engineer a controllable system, the researchers cleverly designed a modified *ssrA*-tag, termed DAS+4, that has a similar affinity for SspB but a reduced affinity for the protease. It was first demonstrated *in vitro* that proteins carrying the modified tag are degraded poorly by ClpXP in the absence of SspB but are rapidly degraded when SspB is

(continued on page 330)

Introns Draw One from the DEK

The spliceosome is a cellular machine composed of RNA and >100 proteins. In mammals, the complexity of the spliceosome is dwarfed only by the essential task that it performs. It must recognize specific signals at the ends of noncoding regions known as introns and catalyze their removal. The majority of introns have ends marked by dinucleotides, GU at the 5' end and AG at the 3' end. Because these signals are often thousands of nucleotides apart, a puzzling question has long remained unanswered: How does the spliceosome pick the correct dinucleotides among so many possible candidates? Now, Mendes-Soares *et al.* (*Science* 2006, 312, 1961–1965) have uncovered a new activity that helps explain the spliceosome's selective nature. They monitored the binding of an early factor in 3' splice site recognition known as the U2-auxiliary factor (U2AF) to a segment of RNA that resembles an authentic 3' splice site or a mutant. An isolated human U2AF complex was not highly selective for the authentic site, but an activity in nuclear extract made U2AF far more discriminating. Biochemical fractionations followed this activity until mass spectrometry identified the responsible protein,



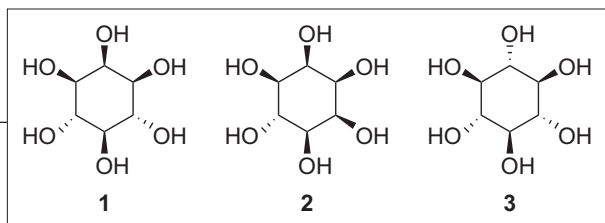
DEK, a factor formerly associated with the chromatin coats on DNA rather than RNA processing. In this study, DEK was shown to interact with U2AF and act as a proctor during that factor's choice of 3' dinucleotide. DEK's watchdog activity appears to be more complex than just enforcing the right binding event. Depletion of the protein from splicing-competent extracts resulted in a loss of splicing catalysis but no change in the assembly of the spliceosome onto the intron. This implies that DEK may function during the mysterious "catalytic activation" stage of the spliceosome cycle. This is especially interesting given its previous connection to chromatin regulation. Perhaps DEK acts as one of the nuclear coordinators to mitigate cross talk between transcriptional regulation and the splicing apparatus. Also of interest is that an overabundance or mutation of human DEK has been linked with certain cancers. **JU**

TIPping the Target Identification Scale

Target identification is a key step and often a major hurdle in chemical genetics experiments and drug-discovery efforts. One method to identify targets of small molecules is photoaffinity labeling, in which UV irradiation is used to covalently cross-link a photoactivatable derivative of the small molecule to its target. However, purification and identification of photoaffinity-labeled proteins are often complicated by the presence of contaminating proteins. Now, Lamos *et al.* (*Angew. Chem., Int. Ed.* 2006, 45, 4329–4333) present the synthesis and application of target identification probe (TIP) reagents that exploit the use of isotopic labels to facilitate target identification.

TIP reagents are cleverly designed multifunctional molecules that contain a mixed isotope photoaffinity label and an affinity tag. The isotope label is composed of a benzophenone moiety containing 11 deuterium atoms; this is a substantial mass difference compared with its unlabeled counterpart. The TIP reagent is used as a 1:1 mixture of labeled to unlabeled molecules, and when incubated in a protein mixture and exposed to UV light, both labeled and unlabeled reagents are covalently attached to target proteins. Avidin affinity chromatography is used to isolate the protein–TIP conjugates from the mixture, and proteolysis and subsequent mass spectrometry (MS) analysis of unlabeled peptides provide a list of candidate binding proteins. Of those peptide fragments, only those that also possess the unique MS signature obtained from the isotopic label are derived from the target of interest. As proof of principle, a TIP reagent containing the immunosuppressive drug cyclosporine was synthesized and successfully used to isolate and identify its binding protein cyclophilin A from a mixture of four proteins.

This approach paves the way for accelerated target identification for many applications. For example, TIP reagents can be used in live cells, and the large mass difference between labeled and unlabeled proteins enables resolution of intact proteins in top-down proteomics experiments. **EG**

Reprinted with permission from *Nature Medicine*

Defeating Dementia

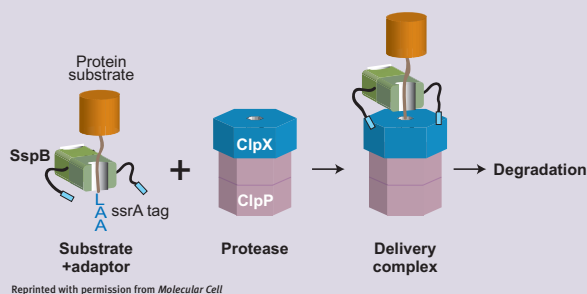
Alzheimer's disease (AD) is the most common form of dementia, and the accumulation of amyloid β -peptide ($A\beta$) aggregates in the brain is thought to be a critical element in the pathogenesis of the disease. It has been suggested that inhibiting $A\beta$ aggregation could be an effective therapeutic strategy for AD, but the complexity of the disease and the challenge of creating appropriate disease models have hindered effective discovery and testing of inhibitors. McLaurin *et al.* (*Nat. Med.* 2006, 12, 801–808) now report the evaluation and efficacy of cyclohexanehexol (CHH) inhibitors of $A\beta$ fibril assembly in a robust transgenic mouse model of AD.

Phosphatidylinositol lipids have been previously shown to strongly facilitate $A\beta$ oligomerization. It has been hypothesized that inhibition of $A\beta$ fibril assembly by CHHs *in vitro* is due to their ability to directly compete with phosphatidylinositol binding to $A\beta$. Two different treatment models were used

to evaluate two CHH stereoisomers, scyllo-CHH and epi-CHH, in mice expressing a mutant protein that causes AD in humans. In the first model, mice were treated prophylactically beginning at 6 weeks of age, before symptoms of AD surface, until either 4 or 6 months of age. Treatment with either inhibitor resulted in significant improvements in cognitive function and other AD-like phenotypes, including reductions in brain $A\beta$ levels and mortality, although the scyllo-CHH stereoisomer was the more effective inhibitor. In the second treatment strategy, mice were given the inhibitors at 5 months of age, after the AD phenotype is established. Remarkably, mice treated with scyllo-CHH, but not epi-CHH, exhibited behavioral improvements and biochemical evidence for reduced $A\beta$ plaques. The authors propose that CHH inhibitors directly prevent and possibly reverse $A\beta$ oligomeric assembly in the brain: clearing of soluble $A\beta$ through normal mechanisms and consequent prevention of disease may result. **EG**

Manipulating Destruction, *continued*

present. The researchers next tested the system *in vivo* in *sspB+*, *clpX+*, *sspB-*, and *clpX-* strains by tagging the protein titin with DAS+4 and placing it under the control of an inducible promoter. Upon induction, very low levels of titin-DAS+4 were found in *sspB+* *clpX+* strains, but substantially higher levels were present in the *sspB-* and *clpX-* strains, an indication that

Reprinted with permission from *Molecular Cell*

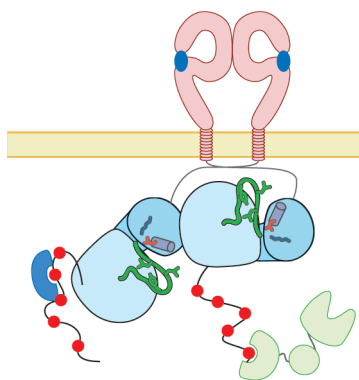
both SspB and ClpXP are necessary for efficient degradation of titin-DAS+4. Taking the system one step further, the researchers placed SspB under an inducible

promoter in a Δ *sspB* strain expressing titin-DAS+4. Rapid degradation of titin-DAS+4 was observed only upon induction of SspB expression. The researchers further demonstrate that controlling degradation of proteins tagged with DAS+4 can be used to manipulate other cellular properties, including antibiotic resistance and enzyme production. **EG**

Asymmetry and Autoinhibition

The epidermal growth factor receptor (EGFR) family of proteins plays critical roles in cell proliferation, differentiation, and motility and is activated in a variety of human cancers, including malignancies of the lung, head, and neck. EGFR is a transmembrane receptor tyrosine kinase that is activated upon dimerization induced by ligand binding to the extracellular domain. Homodimerization of the protein results in the activation of the kinase and phosphorylation of tyrosine residues in the C-terminal domain. These phosphotyrosines serve as docking sites for other signaling proteins, leading to the activation of numerous downstream pathways. Previous research has shown the structural basis for the dimerization of the EGFR extracellular domain. However, the mechanism by which EGFR kinase activity is induced upon ligand-binding has remained unclear. Now Zhang *et al.* (*Cell* 2006, 125, 1137–1149) propose an allosteric mechanism for kinase activation.

Comparing the structures of the prototypical tyrosine kinase Src and cyclin-dependent kinase (CDK) to the structures of EGFR, the authors hypothesized that EGFR is likely activated through asymmetric interaction of two protein molecules. They first showed that an EGFR mutation frequently found in



Reprinted with permission from *Cell*

human lung cancer activates EGFR. This mutation (mutating a neutral leucine to a charged arginine) interrupts several hydrophobic interactions necessary for maintaining the Src/CDK-like inactive state. This result led the authors to hypothesize that EGFR naturally resides

in an autoinhibited state. Further examination of previously described crystal structures of EGFR in activated conformations revealed a potentially important structure in which the kinase domains of two molecules interacted in an asymmetric manner that is similar to the binding of cyclin to CDK. This cyclin binding to CDK leads to activation of the CDK. Mutation of an EGFR residue at the cyclin/CDK-like interface produced crystals of EGFR that resembled an inactive kinase. In addition, the authors showed that mutations at the cyclin/CDK-like interface prevent EGFR activation. Taken together, these results suggest that EGFR is normally autoinhibited and that ligand-induced dimerization leads to an asymmetric interaction of the kinase domains that aids in the activation of the molecule. Given the serious consequences of spurious activation of EGFR, this study provides a glimpse at an elegant solution evolved to precisely regulate EGFR signaling. **ST**

UPCOMING CONFERENCES

Translational Control, CSHL

September 6–10, 2006
Cold Spring Harbor, NY

2006 ACS Fall National Meeting

September 10–14, 2006
San Francisco, CA

Dynamic Organization of Nuclear Function, CSHL

September 27–October 1, 2006
Cold Spring Harbor, NY

RNA Chemistry Meets Biology

September 29–30, 2006
Lund, Sweden

Neuroscience 2006

October 14–18, 2006
Atlanta, GA

HUPO World Congress

October 28–November 1, 2006
Long Beach, CA

Spotlights written by Eva Gordon, Sarah Tegen, and Jason Underwood

Frontiers in Bioimaging

Masaki Takeuchi and Takeaki Ozawa*

Department of Molecular Structure, Institute for Molecular Science, 38 Nishigonaka, Myodaiji, Okazaki, 444-8585 Japan

Under the international collaboration between the National Institute for Basic Biology (NIBB) and the European Molecular Biology Laboratory (EMBL), the second Frontiers in Bioimaging symposium was held March 22–23, 2006, in Okazaki, Japan. This symposium highlighted emerging and innovative technologies for bioimaging and their practical applications in biology. Distinguished scientists offered 26 oral presentations and 20 poster sessions on the state of the art in various biological fields. More than 150 participants took the opportunity to exchange views and share experiences. The meeting covered cutting-edge topics such as the development of fluorescent and bioluminescent probes, analytical techniques with novel concepts, new microscopic systems for 3D imaging of living subjects, and biological progress with imaging techniques. Newly developed probes, analytical techniques, and microscopic systems were the main focus of the symposium.

Existing fluorescent proteins are powerful tools for multicolor labeling of different proteins in a single living cell. However, the difficulty of aligning several lasers and the emission crossover between the fluorophores make multicolor imaging a challenge. Atsushi Miyawaki (RIKEN) introduced newly developed fluorescent proteins with various spectral properties (1). For example, the protein Keima experiences a large shift between excitation and emission spectra: the fluorescent protein is excited at 440 nm and emits light at 620 nm. Combining Keima with cyan fluorescent protein allows quantification and imaging of protein–pro-

tein interactions by fluorescence cross-correlation spectroscopy with a single 458-nm laser line. Labeling a target protein with synthetic chemical probes is also important for detection, purification, and functional studies. For practical applications, the labeling reaction must be highly specific. Kai Johnsson of the Swiss Federal Institute for Technology (Ecole Polytechnique Fédérale de Lausanne) talked about new methods for the covalent and specific labeling of fusion proteins with chemically diverse compounds. He demonstrated several applications, including labeling with spectroscopic and caged probes, selectively immobilizing protein microarrays, and manipulating membrane proteins. These new fluorescent and chemical probes will offer innovative answers to biological questions that traditional approaches have been unable to address.

Another topic of interest included the latest developments of analytical techniques for examining protein–protein interactions, protein modification, and protein transport. In the past few decades, genetic and biochemical approaches have led to the identification of thousands of potential protein interactions. But the cell specificity and the subcellular localization of most of these interactions remain unknown. Tom Kerppola (University of Michigan) a talk



Image courtesy of Corbis

*Corresponding author,
ozawa@ims.ac.jp.

Published online July 21, 2006
10.1021/cb600266h CCC: \$33.50

© 2006 by American Chemical Society

about protein-interaction analyses by complementation of split fragments of fluorescent proteins. Transcriptional protein complexes and protein modification by ubiquitin-family peptides were visualized by the complementation technique. Tarik Issad (Centre National de la Recherche Scientifique) discussed protein-interaction analyses based on bioluminescence-resonance energy transfer. The technique makes it possible to visualize the interactions that occur in the insulin-signaling cascade of an insulin receptor and its intracellular adapters. Spatiotemporal visualization of RNA is also important for understanding the complex function of these molecules. Robert Singer (Albert Einstein College of Medicine) presented an analysis of intracellular messenger RNA trafficking by the fluorescence recovery after photobleaching technique. It was demonstrated that nuclear RNA mobility was not directed by other macromolecules but rather governed by rules of simple diffusion. These new analytical techniques will help us design new experimental systems for a variety of purposes and obtain novel biological data about living cells.

The third topic of interest was discussed in presentations about microscopes based on new concepts. The development of microscopes with highly spatiotemporal resolution is now one of the most important and challenging areas of the bioimaging field. Ernst Stelzer (EMBL) reported on selective plane illumination microscopy (SPIM), a new technique that allows the observation of large (up to a few millimeters) and even living specimens. The illumination of a light sheet instead of a single laser line, along with the rotation of a sample embedded in agarose, has enabled the rapid capture of 3D images. Because excitation by the light sheet minimizes photobleaching, 4D imaging of moving specimens such as a beating heart of a fish was also made possible. John Sedat (California, San Francisco) talked about OMX, an advanced microscopic system. It is the first practical imple-

mentation of structured illumination (SI), in which the grid is superimposed onto the sample (fringe projection) by the insertion of a grid structure into the plane diaphragm. After acquisition of several images with different grid positions and combination of those raw images, the grid lines and the image that are out of focus become invisible, and the contrast and image sharpness are much improved. The system was designed to record rapid 3D multiwavelength imaging on live samples. Stelzer also reported that this SI technology was incorporated into the SPIM to improve resolution at a wide range of magnifications. In addition to new chemical and proteinaceous probes, the development of novel hardware for microscopes appears to be a key challenge for the study of complex biological phenomena.

Such technological advances in bioimaging contribute significantly to the growing body of biological information. Thus, the bioimaging field is in a position to compile and integrate the data to elucidate new biological information. Jan Ellenberg (EMBL) discussed the MitoCheck project (www.mitocheck.org), which is a genome-wide screening analysis of mitotic genes and is run by various European research groups. Transfected-cell arrays, automated time-lapse fluorescence imaging, and computational phenotype analysis of a chromosome-segregation assay were used to develop a fully automated method for microscopy-based small interfering RNA screening. The automated imaging system is powerful enough to analyze complex spatiotemporal processes of target proteins in living cells. Scientists can obtain new quantitative and qualitative insights into the dynamic intracellular events of eukaryotic cells. In the near future, many researchers will apply such systematic analyses to the study of the biologically complex eukaryotic cells. In addition to the above-mentioned presentations, other speakers reported on impres-

sive recent results and perspectives from various fields in which bioimaging is used.

We agree with many participants that this symposium was a fruitful exchange of cutting-edge information about the rapid developments in imaging science, particularly new probes, analytical techniques, and microscopic systems. We would like to express our gratitude to the organizing committee for this successful meeting. In 3 years, NIBB plans to hold the next symposium, which promises even more remarkable advances.

REFERENCE

1. Kogure, T., Karasawa, S., Araki, T., Saito, K., Kinjo, M., and Miyawaki, A. (2006) *Nat. Biotechnol.* 24, 577–581.

RNAs Regulate Biology

Rachel Green^{†,*} and Jennifer A. Doudna^{†,§,*}

[†]Howard Hughes Medical Institute and [‡]Department of Molecular Biology and Genetics, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, and [§]Department of Molecular and Cell Biology and Department of Chemistry, University of California, Berkeley, California 94720

We both entered the RNA field in the late 1980s, when catalytic RNAs and *in vitro* selection approaches were an exciting new frontier. The first scientific meeting one of us (J.A.D.) attended was the 52nd Cold Spring Harbor Symposium (CSHS) on the Evolution of Catalysts in the spring of 1987. It was a spectacular experience: attendees at the meeting were abuzz with the excitement and novelty of ribozymes and the mechanisms of chemical reactions critical for life. Ample time to mingle with colleagues fostered many conversations about the details and unanticipated connections among the systems we were each studying.

Almost 20 years later, the world of functional RNAs is no less exciting. This year's 71st CSHS on Regulatory RNAs, held May 31–June 5, 2006, in Cold Spring Harbor, NY, discussed novel aspects of RNA biology that we are only beginning to understand. Focused largely, though not exclusively, on RNA interference (RNAi), the meeting brought together a diverse group of scientists broadly interested in understanding how, where, when, and why RNA molecules have evolved to regulate gene expression in a wide variety of ways in cells and viruses. Although the pathways and molecular players involved in RNA-mediated gene regulation are being elucidated at a rapid pace, the chemical and mechanistic underpinnings remain to be worked out. The underlying molecular mechanisms, and the possibilities for tapping these processes for therapeutic purposes, fall squarely into the realm of chemical biology.

Rather than attempt to summarize every topic covered at the meeting (a nearly impossible task!), this review focuses on several of the key themes that emerged from the many presentations and informal discussions that occurred during the course of the symposium.

A central topic was the discovery of new noncoding RNAs (ncRNAs). What's out there? The resounding answer: a lot! ncRNAs are abundant in all three kingdoms of life, as revealed by a myriad of approaches, including direct cloning and sequencing of cellular RNA, computational prediction, and microarray analysis. Several research groups are using sequencing methodology, referred to as 454 (www.454.com), for the "deep sequencing" of many thousands of small RNAs (sRNAs). This technique has enabled the rapid compilation of large numbers of sequences that can be sorted by size, sequence, evolutionary conservation, and location within a genome.

As a result, new classes of small ncRNAs are rapidly being identified, although the functions for these molecules remain mysterious. What is clear so far is that size matters. In plants, for example, David Baulcombe described how four distinct variants of the RNA-cleaving enzyme Dicer each produce double-stranded RNA (dsRNA) products (micro- or small interfering RNAs) of a characteristic length that confer on them the ability to enter particular gene-regulating pathways. Steve Jacobsen and Richard Jorgensen presented genetic studies that similarly emphasized the diversity of RNAi-mediated silencing pathways

*Corresponding authors,
ragreen@jhmi.edu,
doudna@berkeley.edu.

Published online July 21, 2006
10.1021/cb600277m CCC: \$33.50
© 2006 by American Chemical Society

that appear to function in different plants. Greg Hannon (1) presented a large family of ~30-mer RNAs, found exclusively in murine testes, that have been dubbed piRNAs because of their propensity to bind Piwi-domain-containing proteins in these germ cells. These 30-mers, which usually have a 5' uridine residue, are typically encoded in



© Miriam Chua, 1999

intergenic regions clustered in the genome and are not conserved at the sequence level in other organisms. Phil Zamore presented data suggesting that a similar class of sRNAs, the rasiRNAs (repeat-associated silencing RNAs), might be generated through a previously undescribed pathway. In the nematode *Caenorhabditis elegans*, David Bartel has found that almost one-tenth of the sequences identified so far by the 454 method corresponds to so-called 21U-RNAs. These are 21-mer RNAs that always contain a 5' uridine residue and are derived from thousands of loci in two broad regions of chromosome IV, dispersed between protein-coding genes and within their introns. No precise function for piRNAs or 21U-RNAs has yet been determined.

ncRNAs are not always small. RNAs produced by RNA polymerase II, the same enzyme that transcribes precursor messenger RNAs (mRNAs) in the nucleus, are sometimes retained in the nucleus rather than exported to sites of protein synthesis in the cytoplasm. Two fascinating examples discussed by David Spector include the 9-kb CTN-RNA and another 7-kb RNA, found in murine and human cells. The CTN-RNA appears to be a long version of an mRNA encoding an amino acid transporter, raising the possibility that CTN-RNA is a storage form of the message that is on hand for rapid processing and export in the event of amino acid starvation. The 7-kb-long RNA is also found in the nucleus, in this case in neurons at sites of preliminary mRNA processing, and its abundance intriguingly correlates with the numbers of neural synapses

formed in cell culture. Other intriguing connections among RNA editing, transport, and RNAi were touched upon in talks by Brenda Bass and Gordon Carmichael, highlighting the extensive interplay among these seemingly distinct processes. Tom Gingeras and Mike Snyder described results from the ENCODE (Encyclopedia of DNA Elements) project aimed at characterizing all the transcripts produced from 1% of the human genome. Findings from their labs and others suggest that much more of the genome is transcribed than previously known, at least at a low level, and more than half of all coding genes have very distal, alternative, tissue-specific transcription start sites. Nick Proudfoot presented evidence that aberrant, intergenic, and genic globin locus transcripts are subjected to RNAi mechanisms, whereas regular globin transcripts are made from genes arranged in transcription-dependent loop structures.

ncRNAs are not unique to eukaryotes: bacteria also contain small regulatory RNAs (so far at least 80 have been identified in *Escherichia coli*), and use structured RNAs called riboswitches (as described by Wade Winkler, Tina Henkin, and Eduardo Groisman) to control gene expression in response to a wide array of small molecules, including magnesium ions (2). Gigi Storz shared new data indicating that small bacterial RNAs can mediate their effects by binding to the 3' untranslated region of the gene, reminiscent of many examples of translational control in eukaryotes (3). Molecular structures of riboswitches are appearing at a

rapid clip, including recent crystal structures of the guanine (4) and S-adenosylmethionine-binding riboswitches (5) presented by Robert Batey and a structure of the *E. coli* thiamine pyrophosphate-binding riboswitch presented on a poster from Dinshaw Patel's lab (6). The remarkable T-box RNA described by Tina Henkin recog-

nizes specific transfer RNAs (tRNAs) and changes structure depending on whether the tRNA is charged with its cognate amino acids; thus, expression of bacterial genes in response to cells' nutritional status is controlled (7). One wonders whether such relatively small RNAs, linked together, could function as a primitive ribosome if supplied with an mRNA template, food for thought about the evolutionary origins of RNA-catalyzed protein biosynthesis.

Many of the identified sRNAs whose functions are known are microRNAs (miRNAs), and hence, numerous labs have used genetic and biochemical approaches to set about finding their molecular partners in different organisms. Scientists from the Ruvkun and Carthew labs described negative regulators of RNAi in nematodes and fruit flies; that cells might control RNAi in response to environmental stimuli was suggested. Craig Hunter presented evidence for a transmembrane transporter of dsRNA required for the spreading of RNA silencing between cells in nematodes (8, 9).

Crystallographic structure determinations of RNAi pathway components, presented by the Barford (10), Joshua-Tor (11), and Doudna (12) and Patel labs (13, 14), show how siRNAs are likely to bind within complexes containing Dicer and the mRNA target-cleaving endonuclease Argonaute 2. These structures are guiding biochemical investigation of the molecular mechanisms that underlie the individual steps in RNAi.

How miRNAs work *in vivo* remains a subject of much discussion and debate. Although evidence in some experimental

systems and from computational approaches suggests that miRNAs can fine-tune gene regulation, data from the Hobert lab show that a miRNA in the nematode nervous system is a clear switch that determines left-versus right-brain functional asymmetry (15, 16). These studies are consistent with early genetic observations in the field and suggest the possibility that such on-off control of gene expression by miRNAs will be more generally observed. Although miRNA tissue localization can be exquisitely specific, as shown in the zebrafish by Ron Plasterk, just a quarter of the miRNAs for which chemically stable antisense oligonucleotides were used to block function had a discernible phenotype. That said, Frank Slack shared compelling evidence for the involvement of the let-7 miRNA family in regulating the expression of important oncogenes implicated in lung cancer. Alex Schier showed an example of another way to affect gene expression, whereby a miRNA promotes deadenylation of maternal mRNAs during early embryogenesis (17, 18). In a related story, Jim Dahlberg and Elsebet Lund reported that miRNA biogenesis is controlled during *Xenopus laevis* egg maturation and that miRNA-dependent deadenylation can occur in the absence of translation.

One particularly vexing question is how miRNAs are able to down-regulate protein synthesis without affecting mRNA levels, a process called translational repression. Several speakers, including Witold Filipowicz, Phil Sharp, and Tim Nilsen, presented biochemical and cell biological studies that attempted to decipher the mechanism(s) of gene regulation by miRNAs (19). And, although little consensus exists about what stage of translation (or some other process) is being controlled by miRNA interactions, all would agree that determining where the components are localized would provide important insight and that processing bodies (P bodies) and stress granules figure prominently in these discussions. An interesting talk by Elisa Izaurralde brought some

of these points home. She showed that multiple pathways exist for regulation and that simply knowing that a gene is controlled by a miRNA does not reveal how it is controlled. Peter Sarnow's talk about a hepatic miRNA established that miRNAs can function as activators rather than as repressors (20). The take-home message is to keep an open mind in examining every new case: the rules are far from established.

Discussed at the meeting was transcriptional silencing, another broad topic that incorporates both small and large RNAs and their cellular effects. Danesh Moazed and Shiv Grewal presented complementary stories on the idea of a self-reinforcing loop of processes responsible for sRNA-mediated gene silencing: transcription makes RNAs, RNA-dependent RNA polymerase amplifies them, Dicer processes, Argonautes bind, and assembled RNA-induced transcriptional silencing complexes act in *cis* at the locus to degrade nascent transcripts (21–24). Although initially surprising, transcriptional silencing depends on transcription of the locus. Edith Heard spoke about how specific subnuclear zones appear to be established for gene silencing by Xist RNA, and Jeannie Lee discussed how direct physical pairing between the X chromosomes is important for establishing chromosome inactivation (25). As we have learned for miRNA-mediated gene regulation, localization is a big factor in explaining how these processes take place. As biochemists, we should note that biology depends on localization; although we often may be able to mimic such effects *in vitro* with high concentrations, we may not always be successful.

With so much of the meeting focused on RNA, it was refreshing to hear several presentations about RNPs, ribonucleoprotein complexes that facilitate the functional association of RNAs with their *in vivo* protein partners. In a session focused on telomerase and cancer, Liz Blackburn presented her lab's recent discovery of "t-stumps", very short telomeres found on the ends of chro-

mosomes in certain cancerous human cells. Data from the Cech lab suggest that yeast telomerase RNA, which includes the template sequence for extending the telomeric stretches at the ends of the chromosomes, is a flexible scaffold for assembling telomerase proteins whose copy number is important for proper telomere maintenance *in vivo*. Carol Greider presented a genetic story indicating that half the amount of telomerase RNA over multiple generations results in a heritable phenotype associated with shortened telomeres (26, 27). An exciting talk on alternative splicing from Bob Darnell provided new insight into how RNP complexes regulate which parts of a pre-mRNA are stitched together to produce tissue-specific messages. Gideon Dreyfuss's dissection of the SMN (survival of motor neurons) complex is beginning to reveal some of the biochemical rules for splicing complex assembly (28, 29).

After a dizzying week thinking about mechanisms in a seemingly overwhelming sea of connections in RNA biology, we will all remember the striking stories that we heard that bring home just how central these processes are to understanding life. Surely one of the most memorable talks about biology was by Michel Georges, who studies why Texel sheep are "exceptionally meaty"; the answer appears to be that these sheep carry a mutation in the untranslated region of the myostatin gene that creates a binding site for an already expressed miRNA (30). We are only beginning to appreciate how much of known phenotypic variation can be explained by these novel classes of regulators, the sRNAs.

REFERENCES

1. Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins, *Nature*, published online June 4, <http://dx.doi.org/10.1038/nature04917>.
2. Cromie, M. J., Shi, Y., Latifi, T., and Groisman, E. A. (2006) An RNA sensor for intracellular Mg(2+), *Cell* 125, 71–84.

3. Opdyke, J. A., Kang, J. G., and Storz, G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*, *J. Bacteriol.* **186**, 6698–6705.
4. Batey, R. T., Gilbert, S. D., and Montange, R. K. (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine, *Nature* **432**, 411–415.
5. Montange, R. K., and Batey, R. T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element, *Nature* **441**, 1172–1175.
6. Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R., and Patel, D. J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch, *Nature* **441**, 1167–1171.
7. Yousef, M. R., Grundy, F. J., and Henkin, T. M. (2005) Structural transitions induced by the interaction between tRNA(Gly) and the *Bacillus subtilis* glyQS T box leader RNA, *J. Mol. Biol.* **349**, 273–287.
8. Feinberg, E. H., and Hunter, C. P. (2003) Transport of dsRNA into cells by the transmembrane protein SID-1. *Science* **301**, 1545–1547.
9. Winston, W. M., Molodowitch, C., and Hunter, C. P. (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1, *Science* **295**, 2456–2459.
10. Parker, J. S., Roe, S. M., and Barford, D. (2005) Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex, *Nature* **434**, 663–666.
11. Song, J. J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity, *Science* **305**, 1434–1437.
12. Macrae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P. D., and Doudna, J. A. (2006) Structural basis for double-stranded RNA processing by Dicer, *Science* **311**, 195–198.
13. Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T., and Patel, D. J. (2005) Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein, *Nature* **434**, 666–670.
14. Yuan, Y. R., Pei, Y., Ma, J. B., Kuryaviy, V., Zhadina, M., Meister, G., Chen, H. Y., Dauter, Z., Tuschl, T., and Patel, D. J. (2005) Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage, *Mol. Cell* **19**, 405–419.
15. Johnston, R. J., Jr., Chang, S., Etchberger, J. F., Ortiz, C. O., and Hobert, O. (2005) MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12449–12454.
16. Johnston, R. J., and Hobert, O. (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*, *Nature* **426**, 845–849.
17. Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P., and Schier, A. F. (2005) MicroRNAs regulate brain morphogenesis in zebrafish, *Science* **308**, 833–838.
18. Van Dongen, S., Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Inoue, K., Enright, A. J., and Schier, A. F. (2006) Zebrafish miR-430 promotes deadenylation and clearance of maternal mRNAs, *Science* **312**, 75–79.
19. Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., and Filipowicz, W. (2006) Relief of microRNA-mediated translational repression in human cells subjected to stress, *Cell* **125**, 1111–1124.
20. Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M., and Samow, P. (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA, *Science* **309**, 1577–1581.
21. Buhler, M., Verdel, A., and Moazed, D. (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing, *Cell* **125**, 873–886.
22. Cam, H. P., Sugiyama, T., Chen, E. S., Chen, X., Fitzgerald, P. C., and Grewal, S. I. (2005) Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome, *Nat. Genet.* **37**, 809–819.
23. Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S. I. (2004) RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing, *Nat. Genet.* **36**, 1174–1180.
24. Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S. I. (2005) RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 152–157.
25. Xu, N., Tsai, C. L., and Lee, J. T. (2006) Transient homologous chromosome pairing marks the onset of X inactivation, *Science* **311**, 1149–1152.
26. Armanios, M., Chen, J. L., Chang, Y. P., Brodsky, R. A., Hawkins, A., Griffin, C. A., Eshleman, J. R., Cohen, A. R., Chakravarti, A., Hamosh, A., and Greider, C. W. (2005) Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenital, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15960–15964.
27. Hao, L. Y., Armanios, M., Strong, M. A., Karim, B., Feldser, D. M., Huso, D., and Greider, C. W. (2005) Short telomeres, even in the presence of telomerase, limit tissue renewal capacity, *Cell* **123**, 1121–1131.
28. Battle, D. J., Lau, C.-K., Wan, L., Deng, H., Lotti, F., and Dreyfuss, G. (2006) The Gemin5 protein of the SMN complex identifies snRNAs, *Mol. Cell*, in press.
29. Golembe, T. J., Yong, J., and Dreyfuss, G. (2005) Specific sequence features, recognized by the SMN complex, identify snRNAs and determine their fate as snRNPs, *Mol. Cell Biol.* **25**, 10989–11004.
30. Clop, A., Marq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J. M., Eychehenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C., and Georges, M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep, *Nat. Genet.* **38**, 813–818.

Crossing a Biological Velvet Rope

Dewey G. McCafferty*

Department of Chemistry, B219 Levine Science Research Center, Box 90354, Duke University, Durham, North Carolina 27708-0354

In Gram-negative bacteria, the outer membrane (OM) serves as a selective permeability barrier, governing the influx and efflux of solutes essential for sustaining life while protecting against deleterious agents in the extracellular milieu. Recently, using chemical genetic methods, Silhavy and Kahne (1, 2) discovered that toxic small molecules can be used in selections employing strains with OM permeability defects to create particular chemical conditions that demand specific suppressor mutations to restore OM function in *Escherichia coli*. This “chemical conditionality” approach (2) was used to identify a multi-protein complex that is required for OM biogenesis. On page 385 of this issue of *ACS Chemical Biology* (ACS CB), Silhavy and Kahne (3) use chemical conditionality to identify YaeT as part of the OM assembly complex. They also report that the suppressor mutation in *yaeT* confers resistance to a specific structural subset of bile acids, thus demonstrating that structurally diverse toxic small molecules select different and specific genetic solutions for correcting permeability defects. This novel application of chemical genetics points to a molecular basis for OM barrier restoration by the OM assembly complex, provides molecular-level identification of potential targets for antimicrobial chemotherapy, and offers a mechanism for identifying factors involved in the assembly of other organelles.

Gram-negative bacteria possess a dual-membrane architecture consisting of a phospholipid and protein inner membrane (IM), a periplasmic space containing peptidoglycan

and soluble proteins, and an unusual OM largely composed of OM proteins (OMPs), lipopolysaccharides (LPSs), and phospholipids. The OM bilayer is an unusual asymmetric structure with the outer leaflet composed largely of highly compact LPS and the inner leaflet made of phospholipids (4). In addition, integral OMPs such as β -barrel proteins span the bilayer, and lipoproteins are attached to the inner leaflet through covalent lipid modifications. Also, other components such as LPSs can be produced as an additional extracellular layer. The OM functions as a protective barrier to toxic materials, yet it is selectively permissive for solute import and waste disposal required for sustaining life in varying environments. The critical importance of OM function for Gram-negative pathogens also makes its assembly an antimicrobial target.

The molecular components of the OM are biosynthesized in the cytoplasm or inner leaflet of the IM and exported across the periplasmic space and into the inner or outer leaflets of the OM (5). OM lipoproteins, biosynthesized and post-translationally modified in the IM and transported via the ATP-binding cassette (ABC) transporter LolCDE, are escorted through the periplasm by the LolA chaperone en route to the OM. Similar chaperone-assisted mechanisms are suggested for transport of Pili proteins to the OM. Integral β -barrel OMPs are produced in the cytoplasm, targeted to the IM, and secreted into the periplasm in a secretory protein-dependent mechanism, where they interact with chaperones and protein folding factors. However, once OMPs

ABSTRACT In contrast to our understanding of the composition of the outer membrane (OM) of Gram-negative bacteria, the biogenesis of this organelle has remained elusive. This is in part because factors involved in OM assembly have been refractive to chemical and biological analyses. A recent study shows how small molecules and chemical conditionality can be used to probe the biogenesis of the OM at the molecular level and suggests that similar techniques can be used to identify factors involved in the assembly of other organelles.

*Corresponding author,
dewey@duke.edu.

Published online July 21, 2006
10.1021/cb6002948 CCC: \$33.50
© 2006 by American Chemical Society

Structurally different toxic small molecules select different and specific genetic solutions for correcting permeability defects.

are in the periplasm, less is known about their trafficking to the OM, bilayer insertion, penultimate organization within the OM, and impact on permeability.

Analysis of the biogenesis of the OM has been hampered for decades because factors involved in OM assembly have not been amenable to identification and genetic analyses. Although “leaky” OM-defective mutants have been generated, generally these mutants have nonspecific permeability defects, precluding identification of factors responsible for OM biogenesis. However, several discoveries now may pave the way to a molecular-level understanding of OM barrier function. First is the discovery of the Cpx and σ^E cell stress responses that led to the subsequent identification of potential OM-biogenesis candidates (6). The latter stress response is extracytoplasmic and specifically activated by misfolded OMPs and disruption of LPS structure. Identification of this regulon suggested possible involvement of a family of candidate OM factors. A second is the identification of OM-defective *E. coli* mutants of a candidate OM-biogenesis gene, *imp*. Genetic and biochemical analyses subsequently confirmed an essential role for *imp* in LPS assembly in the OM.

Silhavy and Kahne (3) added an important chapter to our understanding of OM biogenesis by demonstrating an additional capability of forward chemical genetics. Armed with an *E. coli* strain carrying a defective *imp* allele (*imp4213*) and thus a leaky OM barrier susceptible to the effects of toxic molecules, Silhavy and Kahne established a forward genetic screen designed to identify mutants of *imp4213* that restored OM barrier function. Standard genetic techniques were used to map the suppressor mutations that restored total or partial OM function. Unlike the typical forward chemical genetic screens which are designed to reveal the interaction between a molecule and its target or to illuminate downstream events caused by this interaction, this novel screen was designed to reveal mutants that governed the entry of

the small molecule into the cell. This application has the potential to elucidate many aspects of organelle biogenesis. Silhavy and Kahne identified an interesting continuum of suppressor mutation behavior. For example, intragenic suppressors carrying an additional mutation in the *imp4213* restored near-wild-type levels of OM barrier function and afforded protection from all small-molecule toxins. At the other continuum, some of the suppressors obtained by selection afforded protection only for bile salts and yet were susceptible to antibiotics. In the middle of this continuum was the *yfgL* suppressor, which restored the barrier function for bile salts and chlorobiphenyl vancomycin (CBP-V) but not for erythromycin or vancomycin. Intriguingly, *yfgL* suppressors exhibit remarkable chemical specificity: they remain sensitive to vancomycin but not the closely related CBP-V. Subsequently, the researchers determined that the *yfgL* suppressor gains its protective effects by altering the OM permeability barrier. Lipoprotein YfgL, a component of the cellular machinery that assembles β -barrel OMPs in Gram-negative bacteria, is part of a β -barrel assembly protein complex that includes the integral membrane protein YaeT and two other lipoproteins, YfiO and NlpB.

In this issue of *ACS CB*, Silhavy and Kahne (3) have expanded the continuum of OM barrier function analysis, and a new OM barrier selectivity function has been revealed by another protein component of the OMP β -barrel protein assembly complex. The authors demonstrate that structurally different toxic small molecules select different and specific genetic solutions for correcting permeability defects. Most intriguingly, the *imp4213* suppressor mutation in *yaeT* results in partial restoration of OM barrier function and interestingly confers resistance to a specific structural subset of dihydroxylated bile acid regioisomers sodium deoxycholate and chenodeoxycholate vs the related trihydroxylated bile acid sodium cholate.

Significant steps toward a plausible mechanism for the observed suppressor selectivity

by components of the OMP β -barrel assembly complex are presented. Different suppressors obtained *via* chemical conditionality exhibit chemical specificity, because the entry of each chemical into the cell is determined by its physicochemical properties. The most obvious difference between the structures of these three bile acids as well as the vancomycin/CBP-V derivatives is hydrophobicity. The authors suggest that *yfgL*⁻ and *yaeT6* mutations may cause a reduction in the phospholipid content of the outer leaflet of the *imp4213* cells. The result is the reduction of the surface area of phospholipid bilayer patches and thereby the local concentration at the cell surface of hydrophobic and amphipathic compounds that partition into these patches. That *yfgL*⁻ and *yaeT6* mutations may affect OM phospholipids transport is possible as well.

Thus chemical conditionality has been used to examine OM biogenesis and to discover a novel membrane protein complex that plays an important role in OM barrier assembly and function. This approach is clearly a powerful tool to identify factors involved in organelle and membrane biogenesis and to deconstruct assembly and regulatory events in such seemingly intracellular biological environments.

REFERENCES

1. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., and Kahne, D. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235–245.
2. Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T. J. (2005) Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* **121**, 307–317.
3. Ruiz, N., Wu, T., Kahne, D., and Silhavy, T. J. (2006) Probing the Barrier Functionality of the Outer Membrane with Chemical Conditionality, *ACS Chem. Biol.* **1**, 385–395.
4. Ruiz, N., Kahne, D., and Silhavy, T. J. (2006) Advances in understanding bacterial outer-membrane biogenesis, *Nat. Rev. Microbiol.* **4**, 57–66.
5. Tokuda, H., and Matsuyama, S. (2004) Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim. Biophys. Acta* **1694**, 1N1–9.
6. Ruiz, N., and Silhavy, T. J. (2005) Sensing external stress: watchdogs of the *Escherichia coli* cell envelope. *Curr. Opin. Microbiol.* **8**, 122–126.

Flipping Off the Riboswitch: RNA Structures That Control Gene Expression

Dipali G. Sashital and Samuel E. Butcher*

Department of Biochemistry, University of Wisconsin–Madison, 433 Babcock Drive, Madison, Wisconsin 53706

One of the overriding themes in biology is the need for exquisite control and precision in the regulation of gene expression for the normal growth and development of organisms. Regulated gene expression has been discovered at every level, from transcription of the genetic code to post-translational protein modification. In its simplest terms, genetic regulation is the process by which the cell recognizes its metabolic needs and acts to modulate the levels of certain gene products on the basis of these requirements. Traditionally, these functions were thought to be controlled exclusively by proteins. However, it is now known that >2% of bacterial genes are regulated by metabolite-sensing RNAs without the assistance of proteins (1–3). These “riboswitches” form within the leader sequences of messenger RNAs (mRNAs) and bind their target metabolites with the affinity and specificity required for the precise regulation of gene expression. Metabolite binding to the riboswitch effects an allosteric conformational change that modulates gene expression at the transcriptional or translational level. In the past, our knowledge of how riboswitches could transfer the binding energy of specific metabolites into optimal genetic control was limited by a shortage of 3D structural data. A new flood of riboswitch crystal structures from the laboratories of Batey (4), Ban (5), and Patel (6) now presents a clearer picture of the structural complexity required for these functions (Figure 1).

Similar to RNA aptamers selected *in vitro* to bind small molecules (7, 8), riboswitches have evolved *in vivo* to bind a remarkably wide range of molecules with extremely high affinity, including several diverse metabolites and coenzymes such as guanine and adenine (9–12), L-lysine (13, 14), thiamine pyrophosphate (TPP) (15), and S-adenosylmethionine (SAM) (16–18) (for review, see refs 1–3). This observation is not wholly surprising, given the numerous types of artificial RNA aptamers that had been previously isolated *in vitro* (7, 8, 19). One of the truly fascinating and unique features of riboswitches is the elegant way in which they utilize a variety of ligand-induced structural rearrangements to modulate gene expression. This mechanism depends on the riboswitch’s intricate architecture, which uses both secondary and tertiary structures with far greater complexity than is typically seen in aptamers selected *in vitro* (Figure 1) (20). Within a riboswitch, the ligand-binding aptamer domain is coupled to an “expression platform” whose conformation controls gene expression through a variety of methods. Transcription can be turned off or on through either the formation or the preclusion of a terminator stem (Figure 2, panel a), translation can be inhibited by the sequestration of the Shine–Dalgarno (SD) sequence within a structure (Figure 2, panel b), or the mRNA can be cleaved through enzymatic activity of the riboswitch itself (1–3, 21).

Riboswitches, like their metabolite-sensing protein counterparts, must be able

ABSTRACT Riboswitches are metabolite-sensing RNA structures that have been discovered in regulatory regions of messenger RNA (mRNA). They have the remarkable ability to shut off the transcription or translation of their own mRNAs in response to binding a specific metabolite. In other words, riboswitches regulate their own genes using RNA instead of protein. Three new crystal structures reveal how S-adenosylmethionine and thiamine pyrophosphate riboswitches accomplish this task.

*Corresponding author,
butcher@nmrfam.wisc.edu.

Published online July 21, 2006
10.1021/cb002465 CCC: \$33.50
© 2006 by American Chemical Society

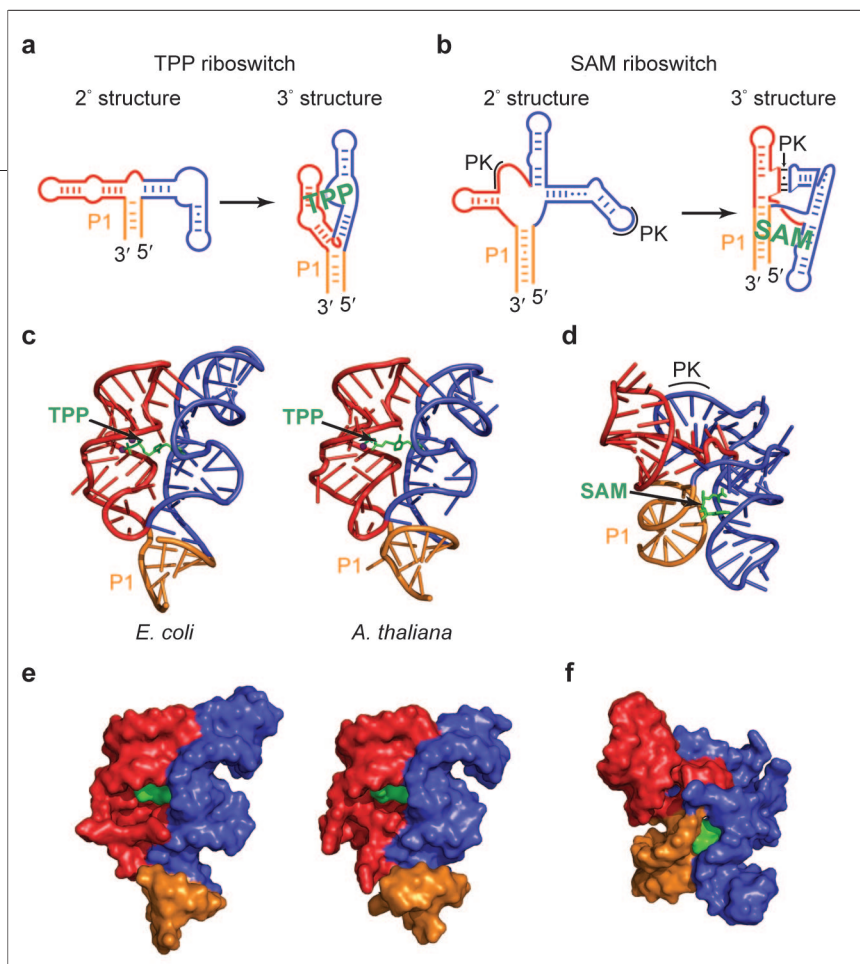


Figure 1. Structures of TPP-sensing riboswitches from *E. coli* and *A. thaliana* and a SAM-sensing riboswitch. Overall folding schemes for a) TPP and b) SAM riboswitches. Crystal structures of *E. coli* and *A. thaliana* c) TPP riboswitches and d) a SAM riboswitch. Ligands (green) are bound at the interface of parallel helices in all three structures. Helix P1, the ligand, and the PK are labeled. Surface representation of e) TPP and f) SAM riboswitches reveals the significant burial of the ligand within the globular RNA fold.

to discriminate against other small molecules in order to be viable gene regulators. Accordingly, riboswitches are exceptionally specific and can reject even subtle variations of the natural ligand. For example, riboswitches distinguish between molecules on the basis of the presence or absence of functional groups (15, 18, 22), atomic charge (23), and stereochemistry (14, 24). This ligand selectivity has led to several seemingly paradoxical observations about the types of molecules that are bound. In several cases, including the TPP riboswitch, negatively charged phosphate groups on the ligand are required for optimal binding to the polyanionic RNA (15, 21, 22); this reveals another rather puzzling facet to the extraordinary ligand specificity displayed by riboswitches.

So how do riboswitches translate a binding event into a conformational rearrangement dramatic enough to turn off a gene? And how is a binding pocket created that can comfortably house negatively charged molecules that would normally be repelled by RNA? New crystal structures of the TPP- and SAM-sensing riboswitches offer insights into these questions (4–6). The TPP-responsive riboswitch is located in the 5' untranslated region of mRNAs involved in thiamine biosynthesis, controlling the expression of these genes by inhibiting translation, in the case of the crystallized TPP riboswitch, or transcription, as has been observed for other known TPP riboswitches (15). Interestingly, this riboswitch is one of the most ubiquitous of the known classes, with representatives identified in all

three domains of life (25). Appropriately, the structures of both a prokaryotic (*Escherichia coli*) and a eukaryotic (*Arabidopsis thaliana*) TPP riboswitch are revealed in publications from the Patel and Ban laboratories, respectively; this allows for an evolutionary comparison (Figure 1, panels a, c, and e) (5, 6). The Batey laboratory unveiled the structure of a SAM-responsive riboswitch, a common bacterial regulator that modulates the expression of genes involved in sulfur metabolism and methionine biosynthesis through transcription termination (Figure 1, panels b, d, and f) (4, 16–18).

One of the most striking features of the riboswitches is their elaborate tertiary structures, in which the RNAs adopt globular folds that surround and bury their ligands within their cores (Figure 1, panels c–f). Complex RNA structures are made up of pre-formed secondary structural elements, including helices, loops, and junction regions that organize into tertiary folds through RNA–RNA stacking or hydrogen-bonding interactions (Figure 1, panels a and b) (26). Intriguingly, in the riboswitch structures, the ligands appear to be the catalysts for the formation of many of the tertiary contacts. Binding of the ligand in each structure occurs at the interface between two parallel helices, creating intricate hydrogen-bonding and electrostatic networks between the RNA, bound Mg^{2+} ions, and the small molecule (Figure 3). These interactions drive the induced-fit binding mechanism of the riboswitch, with the ligand serving to juxtapose and tether separate domains that form tertiary contacts only upon ligand binding. The resulting compaction is instrumental in the formation of a crucial helix (P1) that turns off both riboswitches through the sequestration of the SD sequence (Figure 2, panel a) or the formation of a terminator stem (Figure 2, panel b).

Along with helical packing, ligand binding triggers the organization of single-stranded regions of the RNA that would likely be dis-

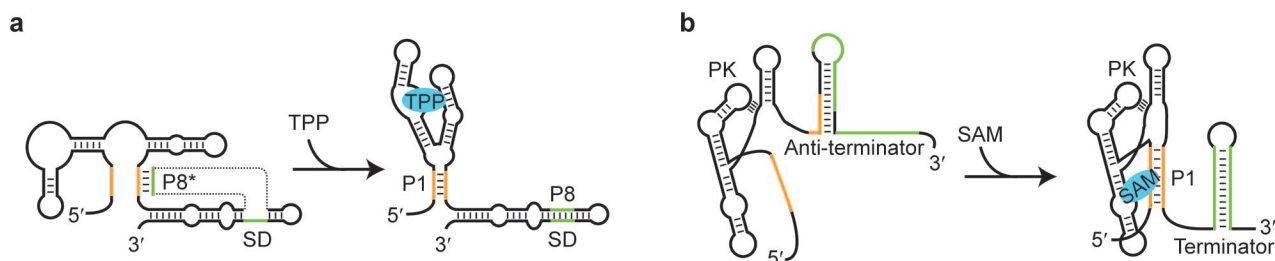


Figure 2. Genetic control by riboswitches is achieved by the coupling of aptamer and expression platform domains. a) Translational and b) transcriptional methods for genetic regulation are shown, as they apply to the TPP and SAM riboswitches, respectively. a) In the absence of TPP, the riboswitch contains little tertiary structure, and helix P8* forms, leaving the SD sequence free to interact with the ribosome. Upon binding of TPP, the aptamer domain folds into a compact, globular structure, forming helix P1. Helix P8 forms within the expression platform, sequestering the SD, thus inhibiting translation. b) The SAM riboswitch forms a PK interaction in the absence of ligand, forming a partial binding pocket for the ligand. An anti-terminator stem forms, portions of which are mutually exclusive with helix P1. Ligand binding stabilizes the formation of P1, thus disrupting the anti-terminator and allowing a terminator stem to form.

ordered in the absence of the small molecule. Many of these unpaired strands are located at the junctions between helices and must be folded in order for helical stacking to occur (Figure 1, panels a and b). This is especially true in the TPP-sensing riboswitch, in which the pyrimidine ring of the ligand interacts extensively with a single-stranded junction between two helices, resulting in a helical stack that forms one-half of the parallel helical motif that dominates the structure (Figure 1, panel c).

In general, the globular structure of the TPP riboswitch appears to form only upon ligand binding. Structure-probing experiments presented by Serganov *et al.* (6) suggest that the tertiary contacts observed in the structure only occur when the ligand is bound. In contrast, the tertiary structure in the SAM-sensing riboswitch is partially preformed (18, 23) through a pseudoknot (PK) interaction (27) that occurs between a loop region of one stem-loop and the junction region between two other helices (Figure 1, panels b and d). In the SAM riboswitch, the PK preserves the global architecture of the structure, including a helical stack and a kink-turn motif (28) (Figure 1, panel b, and Figure 2, panel b). The partial formation of the SAM riboswitch structure likely aids in ligand recognition (4).

One of the most surprising aspects of the TPP riboswitch is its ability to select for negatively charged phosphate groups. Previous studies of the TPP riboswitch revealed that the aptamer also binds thiamine and thiamine phosphate, albeit with decreasing affinity; this suggests that the presence and the length of the phosphate functional group are instrumental in binding specificity (15). The structures of the TPP riboswitch confirm this hypothesis, because the RNA forms two binding pockets: one for the pyrimidine ring present in all three compounds and one for the pyrophosphate found solely in TPP (Figure 3, panels a and b). In the first pocket, the pyrimidine ring of thiamine-containing compounds forms several hydrogen bonds with nucleotides within the riboswitch. Additionally, the pyrimidine ring is sandwiched between two purines, creating a very snug fit for the ligand (Figure 4, panel b). However, for the two helical domains of the riboswitch to be bridged, the pyrophosphate-binding pocket must also be occupied.

Binding of the TPP pyrophosphate moiety occurs in a large pocket formed at a junction between two helices, which also coordinates Mg^{2+} ions that shield the negative charge of the phosphate groups (Figure 3, panels a and b). Binding of the TPP pyrophosphate moiety occurs in a large pocket

formed at a junction between two helices. The helices also coordinate the Mg^{2+} ions that shield the negative charge of the phosphate groups (Figure 3, panels a and b). The *E. coli* structure contains two partially hydrated Mg^{2+} ions, which directly coordinate to the pyrophosphate and form water-mediated hydrogen bonds between the pyrophosphate and the RNA (Figure 3, panel a). The *A. thaliana* structure also shows an identically bound Mg^{2+} ion (Figure 3, panel b). The second Mg^{2+} ion was not observed in the *A. thaliana* structure, possibly because of differences in the data resolution or the occupancy of the second ion. In both structures, the nonbridging oxygen atoms of the pyrophosphates hydrogen-bond to nucleotides within the junction. The structures highlight the importance of Mg^{2+} in neutralizing repulsive forces between the ligand and RNA; this process allows the pyrophosphate to hydrogen-bond within the binding pocket.

Like the TPP riboswitch, the SAM riboswitch integrates the polar functionalities of the ligand into an extensive hydrogen-bonding network (Figure 3, panel c). The methionine amino acid group of SAM stacks on top of its own adenine ring, and both the amino acid and the adenosyl moieties of SAM are recognized by three or four hydro-

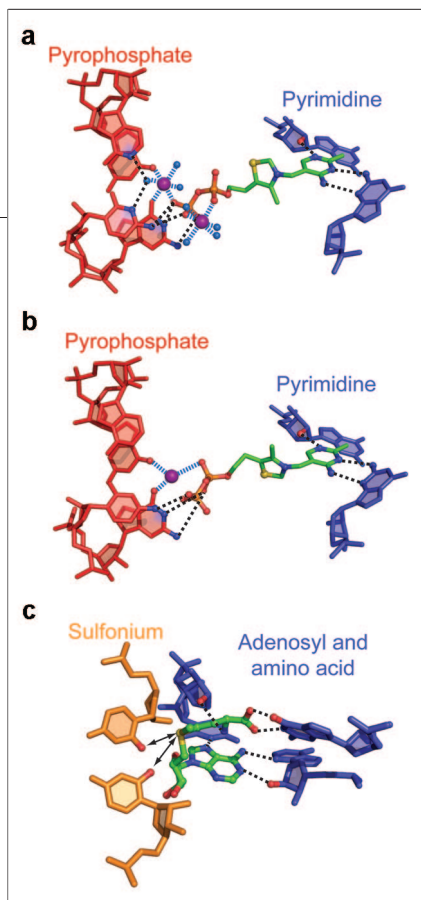


Figure 3. Close-up views of ligand-binding pockets reveal extensive hydrogen-bonding and electrostatic interactions. **a)** and **b)** The TPP riboswitch forms two binding pockets for the pyrimidine ring and pyrophosphate groups. The pyrimidine ring forms three hydrogen bonds with two nucleotides in both TPP riboswitch structures. **a)** The *E. coli* pyrophosphate-binding pocket contains two Mg^{2+} ions (purple spheres) that coordinate (blue dashed lines) the RNA, the ligand, and the water molecules (blue spheres). Several water-mediated and direct ligand–RNA hydrogen bonds (black dashed lines) also form to facilitate pyrophosphate binding. **b)** The *A. thaliana* structure only contains one Mg^{2+} ion in the pyrophosphate-binding region. The hydrogen bonds occur directly between the ligand and the RNA. **c)** The adenosyl and amino acid moieties of SAM hydrogen-bond extensively with the RNA, securing the ligand in a compact conformation. Two carbonyl oxygens in the RNA mediate selectivity for SAM through favorable electrostatic interactions with the positively charged sulfonium ion (indicated by arrows).

gen bonds with adjacent residues within the nearby helix. As a result of these RNA–ligand interactions, the molecule is held in a compact conformation in which its positively charged sulfonium ion is juxtaposed with car-

bonyl oxygens in the minor groove of helix P1 (Figure 3, panel c). This arrangement of the ligand creates a unique electrostatic basis for recognition of SAM over the analogous molecule *S*-adenosylhomocysteine, which lacks the positively charged sulfonium ion (18).

The TPP and SAM riboswitches reveal a variety of strategies with which RNA can selectively bind small molecules. However, information about features that are not recognized by riboswitches may also be informative. For example, the TPP riboswitch does not recognize the TPP thiazole ring through any specific contacts; this may explain why pyriithiamine pyrophosphate (PTPP) is an effective antimicrobial compound (5, 6, 29). In PTPP, a pyridine ring replaces the thiazole ring, but the compound binds the riboswitch with similar affinity to TPP and subsequently can turn off gene expression in a vital metabolic pathway. This example elegantly demonstrates how understanding the structures of bacterial-riboswitch-binding pockets can contribute significantly to our search for novel antimicrobial agents.

Given the ubiquity of riboswitches in bacteria, it is somewhat surprising how few have been found in eukaryotes. The one

exception is the TPP riboswitch, which has been identified in fungi and plants and appears to retain many of the characteristics of the prokaryotic riboswitch (25). The structures of the *E. coli* and *A. thaliana* riboswitches from the Patel and Ban groups confirm this conservation, revealing very few differences between the prokaryotic and eukaryotic RNAs (Figure 4). The divergences occur mainly in regions that do not contact the ligand (Figure 4, panel a). Therefore, ligand binding and specificity are achieved nearly identically in the two structures (Figure 4, panels b and c). These observations support the hypothesis that the TPP riboswitch is a relic of an ancient “RNA world”, in which RNA controlled all the processes of life. The known riboswitches appear to have survived because of the economy that they provide the cell. Given that this efficiency is inherent in a self-regulating mRNA, it is tempting to speculate that modern riboswitches may have evolved in eukaryotes that are not present in prokaryotes. The multitude of additional mRNA processing steps unique to eukaryotes affords a variety of targets for RNA-based genetic control. Further research into potential eukaryotic riboswitches will help to shed

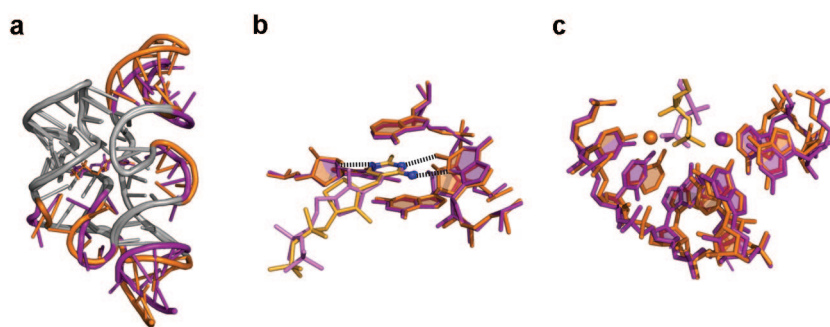


Figure 4. A comparison of prokaryotic and eukaryotic riboswitch structures. **a)** Overall folding of the *E. coli* (orange) and *A. thaliana* (purple) riboswitches. Regions with minimal divergence are highlighted in gray. **b)** A close-up of the TPP pyrimidine-ring-binding pocket reveals a high degree of similarity between the two structures. **c)** A close-up of the pyrophosphate-binding pocket. The RNA does not diverge significantly, although the positioning of the pyrophosphate moiety does. The Mg^{2+} ion observed in the *A. thaliana* structure occupies the same space as one of the Mg^{2+} ions in the *E. coli* structure.

further light onto these fascinating RNAs and their structures.

REFERENCES

- Nudler, E., and Mironov, A. S. (2004) The riboswitch control of bacterial metabolism, *Trends Biochem. Sci.* **29**, 11–17.
- Soukup, J. K., and Soukup, G. A. (2004) Riboswitches exert genetic control through metabolite-induced conformational change, *Curr. Opin. Struct. Biol.* **14**, 344–349.
- Winkler, W. C., and Breaker, R. R. (2003) Genetic control by metabolite-binding riboswitches, *ChemBioChem* **4**, 1024–1032.
- Montange, R. K., and Batey, R. T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element, *Nature* **441**, 1172–1175.
- Thore, S., Leibundgut, M., and Ban, N. (2006) Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand, *Science* **312**, 1208–1211.
- Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R., and Patel, D. J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch, *Nature* **441**, 1167–1171.
- Feigon, J., Dieckmann, T., and Smith, F. W. (1996) Aptamer structures from A to zeta, *Chem. Biol.* **3**, 611–617.
- Hermann, T., and Patel, D. J. (2000) Adaptive recognition by nucleic acid aptamers, *Science* **287**, 820–825.
- Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C., and Breaker, R. R. (2003) Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria, *Cell* **113**, 577–586.
- Mandal, M., and Breaker, R. R. (2004) Adenine riboswitches and gene activation by disruption of a transcription terminator, *Nat. Struct. Mol. Biol.* **11**, 29–35.
- Batey, R. T., Gilbert, S. D., and Montange, R. K. (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine, *Nature* **432**, 411–415.
- Serganov, A., Yuan, Y. R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R., and Patel, D. J. (2004) Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs, *Chem. Biol.* **11**, 1729–1741.
- Grundy, F. J., Lehman, S. C., and Henkin, T. M. (2003) The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12057–12062.
- Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S., and Breaker, R. R. (2003) An mRNA structure in bacteria that controls gene expression by binding lysine, *Genes Dev.* **17**, 2688–2697.
- Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, *Nature* **419**, 952–956.
- Epshtein, V., Mironov, A. S., and Nudler, E. (2003) The riboswitch-mediated control of sulfur metabolism in bacteria, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5052–5056.
- McDaniel, B. A., Grundy, F. J., Artsimovitch, I., and Henkin, T. M. (2003) Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3083–3088.
- Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) An mRNA structure that controls gene expression by binding S-adenosylmethionine, *Nat. Struct. Biol.* **10**, 701–707.
- Wilson, D. S., and Szostak, J. W. (1999) In vitro selection of functional nucleic acids, *Annu. Rev. Biochem.* **68**, 611–647.
- Gilbert, S. D., and Batey, R. T. (2005) Riboswitches: natural SELEXion, *Cell. Mol. Life Sci.* **62**, 2401–2404.
- Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme, *Nature* **428**, 281–286.
- Winkler, W. C., Cohen-Chalamish, S., and Breaker, R. R. (2002) An mRNA structure that controls gene expression by binding FMN, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15908–15913.
- Lim, J., Winkler, W. C., Nakamura, S., Scott, V., and Breaker, R. R. (2006) Molecular-recognition characteristics of SAM-binding riboswitches, *Angew. Chem., Int. Ed.* **45**, 964–968.
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA, *Chem. Biol.* **9**, 1043
- Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes, *RNA* **9**, 644–647.
- Leontis, N. B., and Westhof, E. (2003) Analysis of RNA motifs, *Curr. Opin. Struct. Biol.* **13**, 300–308.
- Staple, D. W., and Butcher, S. E. (2005) Pseudoknots: RNA structures with diverse functions, *PLoS Biol.* **3**, e213
- Winkler, W. C., Grundy, F. J., Murphy, B. A., and Henkin, T. M. (2001) The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs, *RNA* **7**, 1165–1172.
- Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G. M., and Breaker, R. R. (2005) Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine, *Chem. Biol.* **12**, 1325–1335.

Identification of a Smad Phosphatase

Caroline S. Hill*

Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London, WC2A 3PX, United Kingdom

ABSTRACT Activation of Smad signaling pathways downstream of TGF- β superfamily ligands via receptor-mediated Smad phosphorylation is well understood, but little is known about the phosphatases that turn off Smad activity. Now in *Cell*, Feng and colleagues (Lin, X., *et al.* (2006) *Cell* 125, 915–928) report their discovery that PPM1A acts as a Smad phosphatase to terminate TGF- β signaling.

Members of the transforming growth factor β (TGF- β) superfamily of growth and differentiation factors control many key biological processes in embryonic development and in the adult organism, and aberrant signaling by these ligands is involved in several serious human diseases, including fibrosis and cancer. Signals are transduced from the serine (Ser)/threonine (Thr) kinase receptors at the plasma membrane to the nucleus by the Smad family of signal transducers. A crucial step in this process is phosphorylation of the receptor-regulated Smads (R-Smads) at their extreme C-termini by the receptor complex. However, although the details of Smad phosphorylation have been known for 10 years (1, 2), the phosphatase(s) that remove these phosphates have remained elusive. Such phosphatases would be expected to be key regulatory molecules in these important signal-transduction pathways. Now, in a recent paper in *Cell*, Xin-Hua Feng and colleagues (3) have identified PPM1A as the phosphatase that removes the C-terminal phosphates from the R-Smads to terminate signaling by TGF- β -superfamily ligands.

The original concept of the TGF- β signaling pathways was that they were linear and unidirectional from the plasma membrane to the nucleus. Ligands of the TGF- β superfamily activate their receptors by bringing together type I and II receptors in a heterotetrameric complex (4). In this complex, the type II receptor phosphorylates and activates the type I receptor, which in turn phosphorylates R-Smads at two serines in an SM/VS motif at their extreme C-termini.

Smad2 and Smad3 are activated by receptors for TGF- β /Nodal/Activin ligands, whereas Smad1, Smad5, and Smad8 are predominantly activated by receptors for the bone morphogenetic protein (BMP) and growth and differentiation factor subfamilies of ligands (4). R-Smad phosphorylation induces the formation of homomeric and heteromeric complexes with Smad4, which accumulate in the nucleus. There, the complexes are directly involved in both transcriptional activation and repression of target genes in cooperation with other transcription factors. It was deduced that signaling was then terminated by degradation of phosphorylated Smads in the nucleus in a ubiquitin-regulated process (5).

More recently, however, it has become clear that these pathways are more complex and that the Smads are actually continuously shuttling between the cytoplasm and nucleus in unstimulated cells (6) and in the presence of TGF- β -superfamily ligands (7). When TGF- β receptor activity was blocked by a specific small-molecule inhibitor, SB-431542, Smad2 became rapidly dephosphorylated, and Smad2 and Smad4 reaccumulated back in the cytoplasm (7). Moreover, *in vitro* studies of Smad2 export in isolated permeabilized nuclei from TGF- β -induced HeLa cells showed that the Smad2 that was exported from the nucleus was dephosphorylated (6). These and other fluorescence-based approaches (8) have led to a model in which the nucleocytoplasmic shuttling of the Smads during signaling allows them to continuously monitor receptor activity; the duration of receptor activation is thus reflected in the time that the acti-

*Corresponding author,
caroline.hill@cancer.org.uk.

Published online July 21, 2006
10.1021/cb6002702 CCC: \$33.50

© 2006 by American Chemical Society

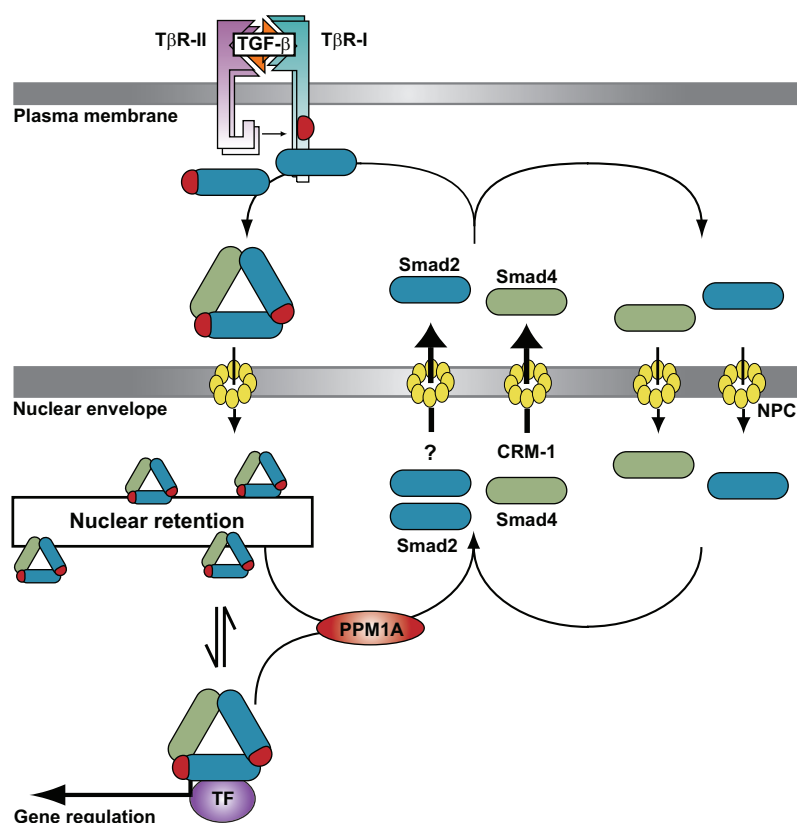


Figure 1. A schematic representation of the TGF- β signaling pathway showing PPM1A as the nuclear Smad phosphatase. For details, see text. The red oval denotes phosphoserines and/or phosphothreonines. CRM-1, chromosome maintenance region 1; NPC, nuclear pore complex; T β R1, TGF- β type I receptor; T β R2, TGF- β type II receptor; TF, transcription factor. Adapted from ref 7.

vated Smad complexes remain nuclear (7). This nucleocytoplasmic shuttling of the Smads requires cycles of phosphorylation and dephosphorylation. A central player in this model is a nuclear R-Smad phosphatase that would remove the C-terminal phosphates from activated R-Smads. This would disrupt Smad complexes and lead to Smad export from the nucleus. If the receptors are still active, these R-Smads would be rapidly rephosphorylated, form complexes with Smad4, and reaccumulate in the nucleus. If the receptors are inactivated as a result of signal termination, then the R-Smads would accumulate in their dephosphorylated basal state in the cytoplasm. Thus the Smad phos-

phatase would have a critical role both in the termination of signaling and in its regulation.

Despite an intense search for Smad phosphatases, their identity has remained elusive until very recently. The first Smad phosphatase reported was pyruvate dehydrogenase phosphatase, which dephosphorylates Smad1 in *Drosophila* in the decapentaplegic pathway and in mammalian cells in the BMP pathway (9). However, this still left the Smad2 and Smad3 phosphatase(s) to be discovered. Now Feng and colleagues (3) describe PPM1A as such a phosphatase. PPM1A, which is well conserved during evolution, is a member of the PPM

family of monomeric, metal-ion-dependent protein phosphatases. These researchers performed an overexpression screen to search in a relatively unbiased way for the Smad2 and Smad3 C-terminal phosphatase. They tested 39 catalytic subunits of Ser/Thr protein phosphatases and dual-specificity phosphatases. The assay involved determining which of these phosphatases could dephosphorylate Smad2 and Smad3, which had already been phosphorylated by constitutively active TGF- β type I receptors in 293T cells. Overexpression of PPM1A clearly decreased C-terminal phosphorylation of Smad2 and Smad3, whereas the other phosphatases had no effect. It was obviously important to address whether the Smad2 and Smad3 dephosphorylation by overexpression of PPM1A was direct, especially given that the receptors that phosphorylate Smad2 and Smad3 are themselves activated by phosphorylation. Therefore, a phosphatase that targets the receptors and deactivates them would have the same effect on Smad2 and Smad3 phosphorylation as that of a phosphatase that directly removes phosphates from the Smad. Feng and colleagues (3) confirmed that the Smad C-terminal phosphoserines were indeed a substrate for PPM1A by demonstrating that PPM1A can dephosphorylate phospho-Smad2 *in vitro*. Interestingly, PPM1A can also dephosphorylate Smad1; this suggests that PPM1A is a general Smad phosphatase. The authors demonstrate an interaction between PPM1A and either Smad2 or Smad3 and show that PPM1A is predominantly nuclear; this was predicted for the Smad phosphatase (6, 7). A current model for the TGF- β signaling pathway in which PPM1A has now been positioned as the nuclear phosphatase that dephosphorylates Smad2 and Smad3 is shown (Figure 1).

Feng and colleagues (3) also investigate the effects on the TGF- β pathway of overexpressing or knocking down PPM1A. They show in tissue-culture cells that overexpression of PPM1A reduces the ability of TGF- β

to cause growth arrest or to activate target genes and, conversely, that knocking down PPM1A promotes TGF- β signaling and enhances TGF- β responses. Finally, they investigate whether the actions of PPM1A are evolutionarily conserved by testing the effects of overexpressing zebrafish PPM1A in early zebrafish embryos. In this case, the pathway that would be expected to be affected is the Nodal signaling pathway, which signals via Smad2 and Smad3 (10). The authors demonstrate that overexpression of PPM1A causes a fusion of the eyes and a thinner posterior notochord; these are consistent with a reduction of Nodal signaling. PPM1A also rescued the dorsalizing effects of the overexpression of the C-terminal domain of Smad2; this suggests that PPM1A acts downstream of Smad2 and antagonizes its activity.

The identification of this long-sought Smad phosphatase now raises some interesting questions. Understanding the regulation of PPM1A activity will be important, particularly to determine whether its activity is altered in any disease states known to be influenced by TGF- β signaling pathways. In addition, is PPM1A the only C-terminal phosphatase of Smad2 and Smad3, or are there others? If PPM1A is a phosphatase that dephosphorylates all R-Smads, both those downstream of TGF- β /Activin/Nodal receptors and those downstream of BMP receptors, then what consequences does this have for coordinated regulation of these pathways? Could its activity be rate-limiting for these pathways? PPM1A also has other substrates, including p38 kinase, Cdk2, PI3 kinase, and Axin. It will be essential to determine whether its activity in signaling pathways involving these molecules could influence its activity in the TGF- β -superfamily signaling pathways. Finally, the Smads are clearly phosphorylated at others sites in addition to those at the C-termini, and the receptors are also highly phosphorylated. Other phosphatases are thus likely to fine-tune the strength and duration of TGF- β sig-

naling pathways, and it will be important to identify them and understand their role in TGF- β signaling.

REFERENCES

- Hoodless, P. A., Haery, T., Abdollah, S., Stapleton, M., O'Connor, M. B., Attisano, L., and Wrana, J. L. (1996) MADR1, a MAD-related protein that functions in BMP2 signaling pathways, *Cell* **85**, 489–500.
- Macías-Silva, M., Abdollah, S., Hoodless, P. A., Pironi, R., Attisano, L., and Wrana, J. L. (1996) MADR2 is a substrate of the TGF β receptor and its phosphorylation is required for nuclear accumulation and signalling, *Cell* **87**, 1215–1224.
- Lin, X., Duan, X., Liang, Y.-Y., Su, Y., Wrighton, K. H., Long, J., Hu, M., Davis, C. M., Wang, J., Brunnicardi, F. C., Shi, Y., Chen, Y.-G., Meng, A., Feng, X.-H. (2006) PPM1A functions as a Smad phosphatase to terminate TGF β signaling, *Cell* **125**, 915–928.
- Massagué, J., Seoane, J., and Wotton, D. (2005) Smad transcription factors, *Genes Dev.* **19**, 2783–2810.
- Lo, R. S., and Massagué, J. (1999) Ubiquitin-dependent degradation of TGF- β -activated Smad2, *Nat. Cell Biol.* **1**, 472–478.
- Xu, L., Kang, Y., Col, S., and Massagué, J. (2002) Smad2 nucleocytoplasmic shuttling by nucleoporins CAN/Nup214 and Nup153 feeds TGF β signaling complexes in the cytoplasm and nucleus, *Mol. Cell* **10**, 271–282.
- Inman, G. J., Nicolás, F. J., and Hill, C. S. (2002) Nucleocytoplasmic shuttling of Smads 2, 3 and 4 permits sensing of TGF- β receptor activity, *Mol. Cell* **10**, 283–294.
- Schmierer, B. and Hill, C. S. (2005) Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor β -dependent nuclear accumulation of Smads, *Mol. Cell Biol.* **25**, 9845–9858.
- Chen, H. B., Shen, J., Ip, Y. T., and Xu, L. (2006) Identification of phosphatases for Smad in the BMP/DPP pathway, *Genes Dev.* **20**, 648–653.
- Whitman, M. (2001) Nodal signaling in early vertebrate embryos. Themes and variations, *Dev. Cell* **1**, 605–617.

Yersinia Inhibits Host Signaling by Acetylating MAPK Kinases

James B. Bliska*

Department of Molecular Genetics and Microbiology, Center for Infectious Diseases, SUNY Stony Brook, Stony Brook, New York 11794-5222

Type III secretion systems (TTSSs) are specialized protein export pathways present in several Gram-negative bacteria that are pathogenic for animals or plants (1). The *Yersinia* spp. *Yersinia pestis*, *Yersinia pseudotuberculosis*, and *Yersinia enterocolitica* are causative agents of plague or enteric infections in humans. These pathogens use a TTSS to counteract development of innate and adaptive immune responses during infection (2). The *Yersinia* TTSS delivers into host cells a set of six effectors known as Yops (2). One of these effectors, YopJ (known as YopP in *Y. enterocolitica*) has been the subject of intense study. YopJ functions biologically to induce apoptosis in macrophages (3, 4) and to inhibit cytokine production in infected host cells (5). However, until recently, the molecular function of YopJ had remained mysterious. In a recent issue of *Science*, Mukherjee *et al.* (6) report that YopJ has acetyltransferase (ATF) activity. This unexpected finding may have important implications for understanding how signaling pathways are normally regulated in eukaryotic cells.

Previously, it had been established that induction of apoptosis and inhibition of cytokine production by YopJ was a consequence of its ability to deactivate or inhibit the mitogen-activated protein kinase (MAPK) and nuclear factor κ B (NF κ B) signaling pathways (5, 7, 8). Subsequently, Orth *et al.* (9) showed that YopJ binds to members of the MAPK kinase (MKK) superfamily and blocks their ability to be activated by phos-

phorylation (9). The MKKs are activated by phosphorylation of two serine or threonine residues within an activation loop. An additional clue to the function of YopJ came from the discovery that YopJ is predicted to share secondary structure similarity with the CE clan of cysteine proteases, which includes the ubiquitin (Ub)-like protease family (10). These enzymes contain a triad of residues (His, Asp/Glu, and Cys) required for catalysis, and mutational analysis of YopJ indicated that these residues were also required for biological function (10). YopJ activity was associated with decreased global levels of protein modification by the Ub-like protein small Ub-related modifier (SUMO) (10). It was thus proposed that YopJ removed Ub-like modifications from signaling proteins to inactivate MAPK and NF κ B response pathways (10). In line with this idea, it was demonstrated that YopJ activity prevented ubiquitination of I κ B, an MKK family member that is integral to the NF κ B pathway (11). Furthermore, Zhou *et al.* (12) found evidence for a broad deubiquitinating activity in YopJ and reported *in vitro* activity on an artificial Ub substrate. YopJ activity was associated with decreased deubiquitinating of several proteins within the NF κ B signaling pathway, including tumor necrosis factor receptor-associated factor-6 (TRAF6) and inhibitor of NF κ B α .

Given the prevailing belief that YopJ acts as a protease on Ub-like substrates, it was quite surprising when Mukharjee *et al.* (6) published evidence that YopJ inactivates

ABSTRACT Pathogenic *Yersinia* spp. secrete the effector YopJ (YopP) into host cells to counteract cytokine production and to induce programmed cell death (apoptosis). YopJ achieves these aims by inactivating mitogen-activated protein kinase (MAPK) and nuclear factor κ B signaling pathways. YopJ was shown to bind to members of the MAPK kinase (MKK) family and was predicted to have protease activity toward ubiquitin (Ub)-like proteins. In a recent report, YopJ was demonstrated to inactivate MKKs *via* acetylation of critical serine or threonine residues. The ramifications of these exciting results are discussed in the context of other studies implicating YopJ as a Ub-like protease.

*Corresponding author,
jbliska@ms.cc.sunysb.edu.

Published online July 21, 2006
10.1021/cb600261k CCC: \$33.50
© 2006 by American Chemical Society

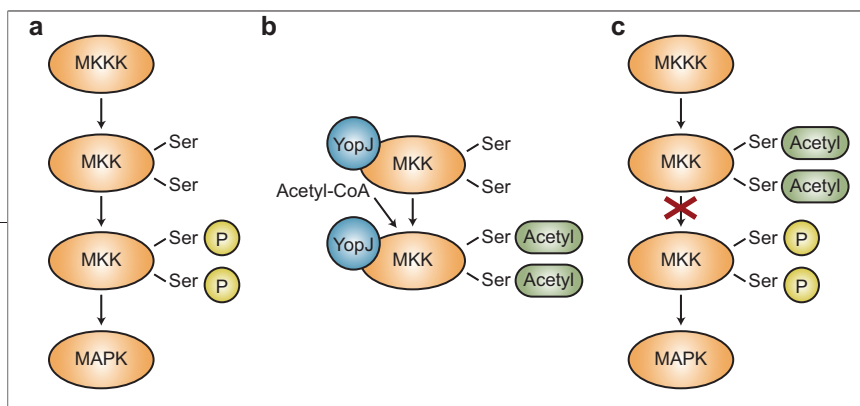


Figure 1. Model of MAPK signaling pathway and inactivation by YopJ ATF activity. a) Shown is a simplified pathway of MAPK signaling through sequential MKKK, MKK, and MAPK modules. MKK is activated by phosphorylation (P) of serine or threonine residues in an activation loop. b) Acetylation of MKK by YopJ. YopJ binds through protein–protein interactions to MKK and, using acetyl-CoA as a cofactor, acetylates serine and threonine residues in an activation loop. c) Mechanism of MKK inactivation by acetylation. Acetylation of Serine and Threonine residues in an activation loop of MKK prevents phosphorylation by upstream MKKK. As a result, the pathway is blocked (red X).

MKKs by acetylation. A prototypical MKK, MKK6, was acetylated on Ser207 and Thr211 when co-expressed with YopJ in *Escherichia coli*. Ser207 and Thr211 are the residues that are normally phosphorylated to activate MKK6. The authors also show that purified YopJ acetylates purified MKK6 or MEK1, another MKK, in a reaction that requires acetyl-CoA. Finally, in a cell-free signaling system, acetylation of MKK6 by YopJ was associated with decreased phosphorylation of this kinase by upstream activators. Together, these data strongly support the idea that YopJ acetylates activation loop serine or threonine residues in MKKs; this prevents phosphorylation of these residues and activation by upstream kinases (Figure 1). The fact that YopJ is specifically targeted to MKKs through protein–protein interactions (9) is an especially attractive feature of the model (6).

These recent findings raise the interesting question of whether YopJ is both an ATF and a deubiquitinating enzyme (Figure 2). In this context, it is important to note that the latter activity has been revealed under conditions in which YopJ and its putative substrates are overexpressed together in cell lines. For example, experiments in which proteins such as TRAF6 and I κ B α were shown to be deubiquitinated in the presence of YopJ utilized transfection vectors to overproduce YopJ and an epitope-tagged form of Ub in cell lines (12). It now appears that such experiments can give misleading results because YopJ can inhibit protein expression from certain commonly used

transfection vectors. In fact, Orth *et al.* (10) observed that the expression of unconjugated epitope-tagged SUMO was decreased in the presence of YopJ; this could explain the global decrease in SUMOylated proteins observed under these conditions. Other observations also argue against the possibility that the deubiquitinating activity linked to YopJ in cell lines is biologically important. For example, I κ B α could be ruled out as an important target of YopJ because earlier experiments indicated that YopJ acts at, or above, the level of IKK β to block the NF κ B pathway (13, 14).

It is possible that under certain conditions YopJ does exhibit protease activity toward Ub. This could explain why Zhou *et al.* (12) observed cleavage of a Ub substrate by YopJ in an *in vitro* assay. In this case, the substrate was Ub-conjugated at its C-terminus to 7-amino-4-methylcoumarin, which is a sensitive substrate for Ub hydrolases. In this context, it is interesting that Mukherjee *et al.* (6) observed acetylation of Lys210 of MKK6 by YopJ in *E. coli*. If YopJ encountered this residue with an attached Ub, might it first remove the Ub before the acetyl group is added?

There is another observation in the field that is difficult to reconcile with the idea that YopJ specifically targets MKKs and inactivates these enzymes by acetylation. Haase *et al.* (15) have obtained evidence that YopP inhibits the NF κ B pathway upstream of IKK β , at the level of the TAK1 (Figure 2). Although Mukherjee *et al.* (6) showed in a cell-free signaling system that YopJ could

block phosphorylation of IKK β , it has not been demonstrated that YopJ directly acetylates IKK β . Could YopJ acetylate TAK1 instead of IKK β ? Alternatively, could acetylation of IKK β by YopJ somehow affect the activation of TAK1? These important questions must be resolved in the future.

The YopJ proteins encoded by *Y. pestis* and *Y. pseudotuberculosis* are highly conserved at the primary sequence level (99–100% identical over 288 residues). Greater sequence divergence occurs between the YopJ and YopP proteins (~94% identity). Interestingly, the YopP proteins from two different serogroups of *Y. enterocolitica* (O:8 and O:9) also show sequence divergence (94% identity) with 17 amino acid differences between the proteins. Because all YopJ/YopP proteins contain the same catalytic triad (His, Glu, and Cys), investigators in the field generally consider them interchangeable. However, evidence already exists that sequence differences in YopP proteins can have important biological consequences. Ruckdeschel *et al.* (16) have shown that YopP proteins with an Arg at position 143 (serogroup O:8) have greater activity toward inhibiting the NF κ B pathway compared with YopP proteins with a serine at this position (serogroup O:9). In future experiments designed to investigate whether these proteins function only as ATFs, or as ATFs and deubiquitinating enzymes, researchers must keep in mind that sequence differences among the YopJ/YopP proteins could impact the outcome.

Mukherjee *et al.* (6) use their knowledge of the chemistry of cysteine protease reactions to suggest a model for the acetylation reaction catalyzed by YopJ. In this model, a thioester bond is formed between the essential Cys172 of YopJ and the acetyl group derived from acetyl-CoA. Subsequently, the thioester bond is attacked by a hydroxyl group on a serine or threonine residue within the activation loop of an MKK member. A well-studied family of arylamine N-ATFs contains a catalytic triad of Cys, His,

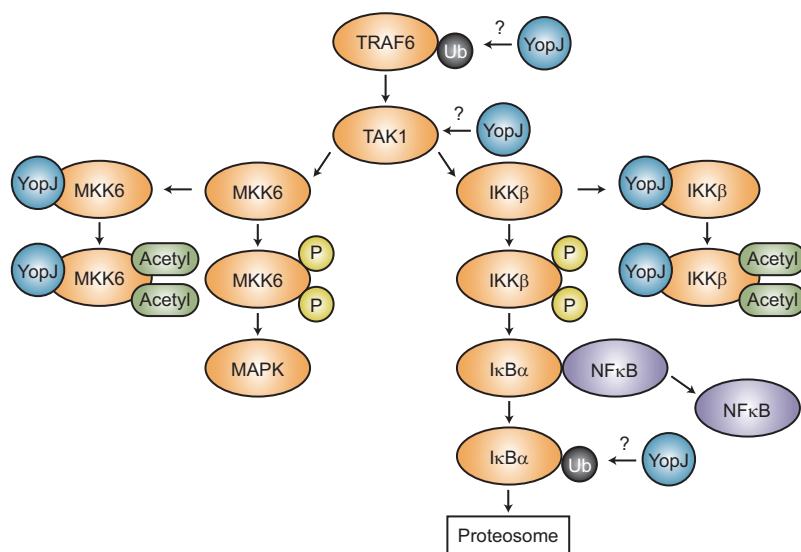


Figure 2. Model of MAPK and NF κ B signaling pathways and mechanisms of inactivation by YopJ. Shown is a simplified pathway of MAPK signaling through sequential steps involving TRAF6, TAK1, MKK6, and MAPK. A parallel pathway involving TRAF6, TAK1, IKK β , and I κ B α and resulting in activation of NF κ B is shown. MKK6 and IKK β are activated by phosphorylation (P) of serine or threonine residues in their activation loops. Also shown are proposed steps at which YopJ inactivates these pathways. Acting as an ATF (6), YopJ binds specifically to MKK6 and IKK β . YopJ then acetylates regulatory serine or threonine residues in each activation loop, and the result is molecules that cannot be phosphorylated by TAK1. Acting as a broad-specificity deubiquitinase (12), YopJ removes Ub from proteins, including TRAF6 and I κ B α . Also shown is the possibility that YopJ inactivates TAK1 by an unknown mechanism (15).

and Asp residues and may utilize a similar mechanism (17). These enzymes are responsible for acetylation of several arylamine compounds, including the antimycobacterial drug isoniazid. Histone ATFs modify protein function by acetylation of His residues, and this activity plays a key role in transcriptional regulation in eukaryotes (18). However, YopJ appears to be the first enzyme identified that acetylates serine and threonine residues on proteins. Mukharjee *et al.* (6) raise the provocative idea that YopJ mimics a previously unidentified class of eukaryotic enzymes that regulate signaling via acetylation of Ser or Thr residues. If true, this would once again show how the study of bacterial virulence factors leads to important new insights into the biology of the host cell.

Acknowledgment: I thank Yue Zhang for helpful suggestions to improve this article. Work in my laboratory on the biology of Yop proteins is sup-

ported by a grant from the National Institutes of Health (R01 AI 433890).

REFERENCES

- Galán, J. E., and Collmer, A. (1999) Type III secretion machines: bacterial devices for protein delivery into host cells, *Science* 284, 1322–1328.
- Viboud, G. I., and Bliska, J. B. (2005) *Yersinia* outer proteins: role in modulation of host cell signaling responses and pathogenesis, *Annu. Rev. Microbiol.* 59, 69–89.
- Monack, D. M., Mecsas, J., Ghori, N., and Falkow, S. (1997) *Yersinia* signals macrophages to undergo apoptosis and YopJ is necessary for this cell death, *Proc. Natl. Acad. Sci. U.S.A.* 94, 10385–10390.
- Mills, S. D., Boland, A., Sory, M. P., van der Smitsen, P., Kerbouch, C., Finlay, B. B., and Cornelis, G. R. (1997) *Yersinia enterocolitica* induces apoptosis in macrophages by a process requiring functional type III secretion and translocation mechanisms and involving YopP, presumably acting as an effector protein, *Proc. Natl. Acad. Sci. U.S.A.* 94, 12638–12643.
- Palmer, L. E., Hobbie, S., Galán, J. E., and Bliska, J. B. (1998) YopJ of *Yersinia pseudotuberculosis* is required for the inhibition of macrophage TNF α production and the downregulation of the MAP kinases p38 and JNK, *Mol. Microbiol.* 27, 953–965.

- Mukherjee, S., Keitany, G., Li, Y., Wang, Y., Ball, H. L., Goldsmith, E. J., and Orth, K. (2006) *Yersinia* YopJ acetylates and inhibits kinase activation by blocking phosphorylation, *Science* 312, 1211–1214.
- Schesser, K., Spiik, A. K., Dukuzumuremyi, J. M., Neurath, M. F., Pettersson, S., and Wolf-Watz, H. (1998) The *yopJ* locus is required for *Yersinia*-mediated inhibition of NF- κ B activation and cytokine expression: YopJ contains a eukaryotic SH2-like domain that is essential for its repressive activity, *Mol. Microbiol.* 28, 1067–1079.
- Ruckdeschel, K., Harb, S., Roggenkamp, A., Hornef, M., Zumbihl, R., Kohler, S., Heesemann, J., and Rouot, B. (1998) *Yersinia enterocolitica* impairs activation of transcription factor NF- κ B: involvement in the induction of programmed cell death and in the suppression of the macrophage tumor necrosis factor α production, *J. Exp. Med.* 187, 1069–1079.
- Orth, K., Palmer, L. E., Bao, Z. Q., Stewart, S., Rudolph, A. E., Bliska, J. B., and Dixon, J. E. (1999) Inhibition of the mitogen-activated protein kinase superfamily by a *Yersinia* effector, *Science* 285, 1920–1923.
- Orth, K., Xu, Z., Mudgett, M. B., Bao, Z. Q., Palmer, L. E., Bliska, J. B., Mangel, W. F., Staskawicz, B., and Dixon, J. E. (2000) Disruption of signaling by *Yersinia* effector YopJ, a ubiquitin-like protein protease, *Science* 290, 1594–1597.
- Carter, R. S., Pennington, K. N., Ungurait, B. J., Arrate, P., and Ballard, D. W. (2003) Signal-induced ubiquitination of I κ B Kinase- β , *J. Biol. Chem.* 278, 48903–48906.
- Zhou, H., Monack, D. M., Kayagaki, N., Wertz, I., Yin, J., Wolf, B., and Dixit, V. M. (2005) *Yersinia* virulence factor YopJ acts as a deubiquitinase to inhibit NF- κ B activation, *J. Exp. Med.* 202, 1327–1332.
- Zhang, Y., and Bliska, J. B. (2003) Role of Toll-like receptor signaling in the apoptotic response of macrophages to *Yersinia* infection, *Infect. Immun.* 71, 1513–1519.
- Ruckdeschel, K., Mannel, O., Richter, K., Jacobi, C. A., Trulzsch, K., Rouot, B., and Heesemann, J. (2001) *Yersinia* outer protein P of *Yersinia enterocolitica* simultaneously blocks the nuclear factor- κ B pathway and exploits lipopolysaccharide signaling to trigger apoptosis in macrophages, *J. Immunol.* 166, 1823–1831.
- Haase, R., Richter, K., Pfaffinger, G., Courtois, G., and Ruckdeschel, K. (2005) *Yersinia* outer protein P suppresses TGF- β -activated kinase-1 activity to impair innate immune signaling in *Yersinia enterocolitica*-infected cells, *J. Immunol.* 175, 8209–8217.
- Ruckdeschel, K., Richter, K., Mannel, O., and Heesemann, J. (2001) Arginine-143 of *Yersinia enterocolitica* YopP crucially determines isotype-related NF- κ B suppression and apoptosis induction in macrophages, *Infect. Immun.* 69, 7652–7662.
- Sandy, J., Mushtaq, A., Holton, S. J., Schartau, P., Noble, M. E., and Sim, E. (2005) Investigation of the catalytic triad of arylamine N-acetyltransferases: essential residues required for acetyl transfer to arylamines, *Biochem. J.* 390, 115–123.
- Roth, S. Y., Denu, J. M., and Allis, C. D. (2001) Histone acetyltransferases, *Annu. Rev. Biochem.* 70, 81–120.

Energy Transfer between Biological Membranes

Volkmar Braun*

Microbiology/Membranephysiology, University of Tuebingen, Tuebingen, Germany

ABSTRACT The crystal structures of two transport proteins associated with a fragment of the TonB protein have been reported for the first time in two recent papers. TonB is implicated in transferring energy from the cytoplasmic membrane to the outer membrane of Gram-negative bacteria. The very similar structures of the protein complexes define the mode of interactions of TonB with active transport proteins.

The outer membrane (OM) of Gram-negative bacteria is ideal for studying mechanisms of substrate transport through membranes. The OM contains porins through which substrates pass by diffusion, specialized porins that recognize substrates entering cells by facilitated diffusion, and energy-coupled transporters. These membrane proteins are abundant when needed under appropriate growth conditions, and relatively large amounts of OMs can be isolated. The OM has unique properties. Its outer leaflet is mainly formed by the fatty acids bound to glycolipids exposed at the cell surface, and its inner leaflet is composed of a lipid covalently bound to a lipoprotein and phospholipids, mainly phosphatidylethanolamine. Proteins form β -barrels in OMs. The high abundance of the proteins, β -barrel structure, and unique lipid environment contribute to the isolation of these proteins in amounts and purity that allow protein crystals to be obtained. These crystals are of such a high quality that they diffract to a resolution permitting structures to be elucidated. In fact, several crystal structures of all three kinds of porins and transporters have been determined.

The energy-coupled transporters are of particular interest because the OM does not contain an energy source. Energy is transferred from the cytoplasmic membrane (CM) into the OM. The proton motive force (PMF) of the CM drives substrate transport across the OM. Three proteins are involved in energy transfer from the CM into the OM:

TonB, with its N-terminus inserted in the CM and its C-proximal regions interacting with OM transporters; ExbB, which spans the CM three times, with most of the protein in the cytoplasm; and ExbD, arranged similarly as TonB. These three proteins are required for energy-coupled transport across the OM, but not for transport across the CM (1). Three principal questions have emerged: How does the TonB–ExbB–ExbD complex (Ton complex) respond to the PMF of the CM? How is the energy transferred from the CM into the OM? How do the transporters respond to the energy input?

The crystal structures of five OM transport proteins have been determined. Four of these proteins transport Fe^{3+} siderophores, and one transports vitamin B_{12} . Siderophores are secreted by bacteria and fungi and complex Fe^{3+} , which otherwise forms a virtually insoluble hydroxide precipitate. The very low concentrations of these substrates and their large size exclude uptake by diffusion through porins at sufficient rates. The substrates bind to the transporters with K_d values in the nanomolar range. Extraction of the substrates from the medium and their concentration on the cell surface guarantee their availability in growth-promoting amounts.

The crystal structures of all five proteins reveal the same basic transporter structures. The proteins are composed of 22 antiparallel β -strands that form a β -barrel (residues 161–714 in FhuA). The pore inside the β -barrel is completely occluded by a globular domain, which has been called a cork, a

*Corresponding author,
volkmar.braun@mikrobio.uni-
tuebingen.de.

Published online July 21, 2006
10.1021/cb00256k CCC: \$33.50

© 2006 by American Chemical Society

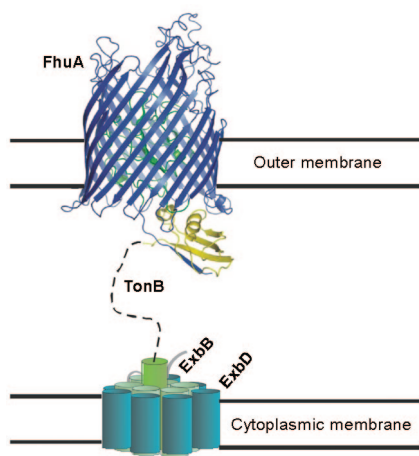


Figure 1. Crystal structure of FhuA–TonB(158–235) incorporated into the OM (3). The structure of the TonB–ExbB–ExbD complex is not known, and only the transmembrane portions are drawn in the CM. The structure of the TonB segment that connects the C-terminal crystal form to the transmembrane portion is not known and is depicted as a dashed line.

plug, or a hatch. The latter term is the most appropriate because it implies an active participation of the globular domain in substrate transport. The β -strands are connected by rather large loops at the cell surface and short turns in the periplasm. The substrate-binding sites are formed by amino acid side chains of the loops and the hatch. The sites are located within the proteins well above the cell surface (Figure 1). The loops are flexible and move upon binding of the substrates; they partially or completely occlude the access of the binding site, depending on the transporter. This is most obvious in the FecA protein. Binding of the substrate diferric dicitrate induces movement of loop 7 by 11 Å and loop 8 by 15 Å, and closure of the substrate entry site results. The substrate can then no longer escape into the medium; it can only move into the periplasm. Substrate binding occurs independent of the Ton complex, but translocation across the OM through the pore of the β -barrel requires energy medi-

ated by the Ton complex. Energy is also required for the release of the substrates tightly bound by ~ 10 amino acid side chains. The stereochemistry of the side chains must be altered to weaken binding. In addition, the hatch must move within or out of the β -barrel so that a pore is formed through which the substrates can move into the periplasm.

In the June 2, 2006, issue of *Science*, the crystal structures of periplasmic TonB fragments bound to BtuB (2) and to FhuA (3) are described. FhuA transports the Fe^{3+} siderophore ferrichrome, and BtuB transports vitamin B_{12} . These long-awaited crystal structures provide much-needed structural information on energy-coupled transport across the OM of Gram-negative bacteria. Both structures delineate the interaction of TonB with the transporters. Of particular importance is the so-called Ton box in the transporters. The Ton box is located close to the N-terminus of the transporters and interacts with TonB (Figure 1). Previous results have demonstrated that certain single amino acid replacements in the Ton box of FhuA and BtuB inactivate the transporters. Amino acid replacements in Gln160 of TonB partially restore the activities of the FhuA and BtuB Ton box mutants. This has been taken as evidence that these two regions interact. This conclusion is supported by spontaneous *in vivo* formation of disulfide bridges between cysteine residues introduced in the Ton box of BtuB and FecA and cysteine residues introduced in region 160 of TonB.

In the earlier FhuA crystal structure, the Ton box (residues 6–13) is not seen because it is flexible. In the new crystal structure, in which the TonB fragment (residues 33–239) were used for crystallization, but only residues 158–235 are observed) associates with FhuA, the FhuA Ton box forms a parallel interaction with the β 3-strand of TonB. The β 3-strand is part of a three-stranded β -sheet that also includes the β 1- and β 2-strands (Figure 2). The NMR

structure of TonB(152–239) reveals a fourth β -strand, in antiparallel orientation, which is replaced by the β -strand of the FhuA TonB box (3). The crystal structure of another TonB fragment, TonB(148–239), shows a fold similar to that of TonB(152–239), in which one β -strand of the three-stranded β -sheet of one monomer forms a fourth antiparallel β -strand with the three-stranded β -sheet of another monomer (4). Gln160 of TonB is not seen in the FhuA–TonB(158–235) structure but can be oriented such that it forms a hydrogen bond with Thr12. In the BtuB–TonB(153–233) structure, Gln160 is seen and interacts with Asp6, Leu8, and Val10 of the BtuB Ton box (residues 6–12). The new crystal structures confirm the earlier results of genetic suppressor analyses and cysteine cross-linking experiments. Numerous additional interactions occur between the TonB fragments and the transporters. In both structures, the TonB fragments occupy approximately half of the periplasmic surface area of the transporters.

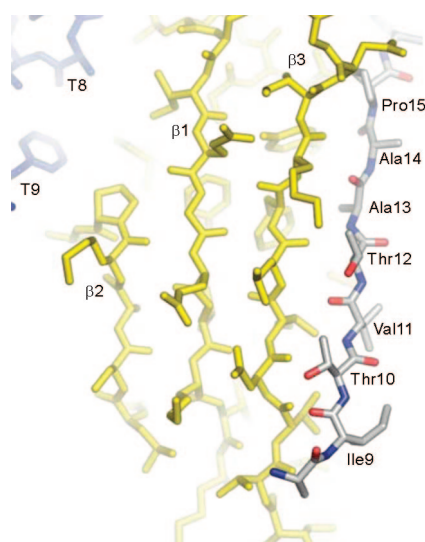


Figure 2. Interface between the FhuA Ton box and the TonB β -sheet. The amino acids (carbon, white; nitrogen, blue; oxygen, red) of the FhuA Ton box are numbered (9–15). The TonB β -sheet consists of three β -strands, β 1, β 2, and β 3. T8 and T9 denote turn 8 and turn 9 in FhuA (3).

Of particular interest in terms of function is the position of the TonB α 1-helix close to the hatch domain; this location allows the interaction of TonB Arg166 with Glu56 of the FhuA hatch. This interaction might be important for the dislocation of the hatch to open the pore. In BtuB–TonB(153–233), Arg158 of TonB interacts with Asp6 of the BtuB Ton box.

Crystal structures are static and show only one form of a protein or protein complex. They do not reveal functional dynamics but may suggest experiments aimed at helping scientists understand the way proteins function. The determined structures of TonB fragments bound to BtuB and to FhuA do not immediately provide solutions to the energy transfer from the CM into the OM and to the conformational changes that must occur for substrate transport. In fact, binding of the TonB fragments to the transporters causes only small changes in the structures of the transporters, except for the fixation of the Ton box in a defined position. The crystal structures are likely to reflect an inactive form of the transporters, given that the substrates still occupy their binding sites and the hatches still close the β -barrels. Dimeric crystal structures have been determined (4, 5), and dimeric full-length TonB has been found *in vivo* (6). Do they represent intermediary forms in the reaction cycle underlying energy-coupled transport? Is TonB permanently associated with transporters, or does it dissociate from the transporters and associate with other transporters? These questions are important, given that the OM can have several-hundred-thousand copies of transporters yet only a few-hundred copies of TonB. The transporters also fulfill receptor functions that may or may not depend on energy input via TonB. For example, FecA transports ferric citrate but is also essential for induction of transcription of the transport genes (7). FhuA binding of bacteriophages T1 and Φ 80 requires TonB, whereas binding of phage T5 is TonB-independent. Because

mutants of phages T1 and Φ 80 that infect *tonB* mutants can be isolated, phage DNA uptake into cells does not require TonB; rather, binding to FhuA requires TonB. This was actually the first evidence of a TonB-dependent conformational change in FhuA at the cell surface. In addition, sensitivity of cells to a protein toxin (colicin M) and a toxic peptide (microcin J25) requires FhuA and TonB. Whether FecA and FhuA fulfill such activities through the same structural changes is not known. Which roles the ExbB and ExbD proteins play in energy harvesting in the CM is also not known. The TonB–ExbB–ExbD complex must be isolated to determine whether the relative protein molar amounts of 1:7:2 found in cells is the stoichiometry of the complex. The complex must be reconstituted in artificial lipid membranes to determine whether it conducts protons. Such experiments should also involve artificially applied transmembrane potentials. A reconstituted system with wild-type and inactive mutant proteins also would provide access to physicochemical investigations, which are necessary to unravel the dynamics of energy harvesting and transfer. Finding conditions for solubilization and purification of the complex and its crystallization should be possible, because crystal structures of membrane–protein complexes, such as cytochrome oxidase, have been determined. Certainly, understanding energy-coupled transport across the OM is a long way off, but it will be well worth the effort.

The benefits of any future effort will not be limited to the elucidation of how the Ton system works. The Ton system resembles the Tol system, which serves for the uptake of certain phages and bacterial toxins (colicins) and is involved in the assembly of the OM. The membrane-spanning TolA portion can functionally replace the membrane-spanning TonB portion. The TolR and TolQ proteins can partially substitute for the ExbD and ExbB proteins, and *vice versa*. These proteins also have sequence similari-

ties to flagellar motor proteins (MotA/ExbB/TolQ and MotB/ExbD/TolR). For example, replacement of a conserved aspartate residue in the transmembrane region of ExbD, TolR, and MotB inactivates the Ton, Tol, and Mot functions (8). Figuring out how the Ton system functions will also help us understand how the Tol and Mot systems work, and *vice versa*.

REFERENCES

1. Postle, K., and Kadner, R. J. (2003) Touch and go: tying TonB to transport, *Mol. Microbiol.* 49, 869–882.
2. Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396–1399.
3. Pawelek, P. D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C. M., Molseeva, N., Allaire, M., and Coulton, J. W. (2006) Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor, *Science* 312, 1399–1402.
4. Peacock, R. S., Weljie, A. M., Howard, S. P., Price, F. D., Vogel, H. (2004) The solution structure of the C-terminal domain of TonB and interaction studies with TonB box peptides, *J. Mol. Biol.* 345, 1185–1197.
5. Ködding, J., Killig, F., Polzer, P., Howard, S. P., Diederichs, K., and Welte, W. (2005) Crystal structure of a 92-residue C-terminal fragment of TonB from *Escherichia coli* reveals significant conformational changes compared to structures of smaller TonB fragments, *J. Biol. Chem.* 280, 3022–3028.
6. Sauter, A., Howard, S. P., and Braun, V. (2003) *In vivo* evidence for TonB dimerization, *J. Bacteriol.* 185, 5742–5754.
7. Braun, V. (1995) Energy-coupled transport and signal transduction through the gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins, *FEMS Microbiol. Rev.* 16, 295–307.
8. Braun, V., and Herrmann, C. (2004) Point mutations in transmembrane helices 2 and 3 of ExbB and TolQ affect their activities in *Escherichia coli* K-12, *J. Bacteriol.* 186, 4402–4406.

Nitrosative Stress in the ER: A New Role for S-Nitrosylation in Neurodegenerative Diseases

Michael T. Forrester[†], Moran Benhar[‡], and Jonathan S. Stamler^{†,*,*}

[†]Department of Biochemistry and [‡]Department of Medicine, Duke University Medical Center, Box 2612, Durham, North Carolina 27710

Nitric oxide (NO) mediates cellular signaling pathways that regulate a plethora of physiological processes. In the brain, NO has been implicated in neurotransmission, neuromodulation, and synaptic plasticity (1). Excessive generation of NO and NO-derived reactive nitrogen species (RNS), however, has also been implicated in the pathogenesis of neurodegenerative disorders (NDDs), including Alzheimer's disease (AD) and Parkinson's disease (PD) (2). Subsequent findings have demonstrated that apoptosis is a critical process underlying these disorders (3). In a recent issue of *Nature*, Takashi Uehara and co-workers (4) elegantly demonstrate that NO-mediated S-nitrosylation of protein disulfide isomerase (PDI) inhibits PDI function, leads to dysregulated protein folding within the endoplasmic reticulum (ER), and consequently results in ER stress that promotes neuronal cell death. A causal role for this sequence of events in human NDD was supported by the demonstration that PDI is S-nitrosylated in the brains of patients suffering from PD or AD (but not in normal brains). Thus, the findings of Uehara *et al.* (4) provide additional evidence of a role for dysregulated protein S-nitrosylation (nitrosative stress) in NDD and indicate that ER dysfunction may serve as a critical common factor that couples NO-induced cellular stress to neurodegeneration.

A well-established model for understanding both NO-mediated neuromodulation and neurodegeneration entails a central role

for the *N*-methyl-D-aspartate type of neuronal glutamate receptor (NMDAR). Activation of NMDARs drives Ca²⁺ influx, which in turn activates the predominant NO-synthesizing enzyme in neurons, neuronal NO synthase (nNOS) (5). Many proteins are S-nitrosylated by physiological nNOS activity (6). However, excessive activity of the NMDAR leads to S-nitrosylation of additional proteins, including GAPDH (7) and MMP-9 (8), that facilitate neuronal cell death.

New studies on the NO biology of blood vessels point to the potential importance of mitochondrial sources of NO in signaling (9), and the idea has been raised that increased production of NO within mitochondria may lead to mitochondrial dysfunction that is characteristic of metabolic syndrome disorder (10), a condition that has been previously connected with oxidative mitochondrial injury. An analogous picture emerges in neurobiology: NO is now being implicated in the pathogenesis of NDD in which mitochondrial dysfunction is thought to play an essential role. The connection between mitochondrial dysfunction and NDD has been linked previously with oxidative stress. Indeed, the best-established laboratory model of PD relies on inducing mitochondrial dysfunction with rotenone or 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), both of which inhibit mitochondrial complex I (11); the generation of reactive oxygen species (ROS) and the depletion of ATP result. Neurons are particularly vulnerable to decreased ATP concentrations because of

ABSTRACT S-Nitrosylation, the covalent addition of a nitrogen monoxide group to a cysteine thiol, has been shown to modify the function of a broad spectrum of mammalian, plant, and microbial proteins and thereby to convey the ubiquitous influence of nitric oxide on cellular signal transduction and host defense. Accumulating evidence indicates that dysregulated, diminished, or excessive S-nitrosylation may be implicated in a wide range of pathophysiological conditions. A recent study establishes a functional relationship between inhibitory S-nitrosylation of the redox enzyme protein disulfide isomerase (PDI), defects in regulation of protein folding within the endoplasmic reticulum (ER), and neurodegeneration. Further, an examination of human brains afflicted with Parkinson's or Alzheimer's disease supports a causal role for the S-nitrosylation of PDI and consequent ER stress in these prevalent neurodegenerative disorders.

*Corresponding author,
staml001@mc.duke.edu.

Published online July 21, 2006
10.1021/cb600244c CCC: \$33.50
© 2006 by American Chemical Society

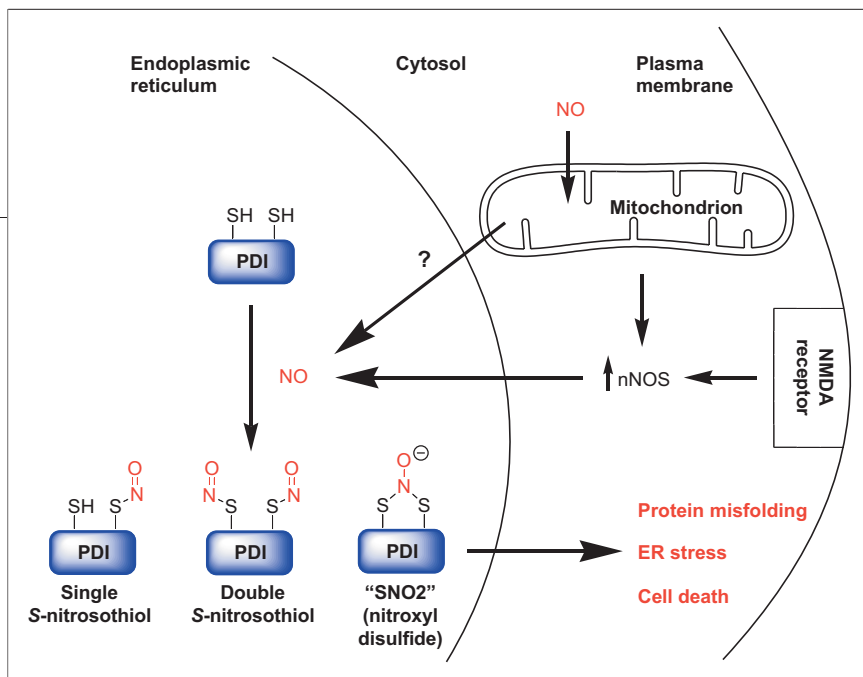


Figure 1. Mitochondrial dysfunction or NMDAR overexcitation promotes S-nitrosylation of PDI, which leads to protein misfolding, ER stress, and cell death. Mitochondria may serve as a source of NO or may activate cytosolic nNOS. NO derived from nNOS may aggravate the mitochondrial insult. S-Nitrosylated PDI may contain one or two S-nitrosothiols on each Trx domain. Alternatively, S-nitrosylation may produce an SNO₂, in which the NO group is shared by the active site thiols. For simplicity, one of the two Trx domains for each PDI is depicted.

their limited capacity to upregulate glycolysis (12), and the *substantia nigra* is particularly susceptible to oxidative injury, at least partly because of the presence of dopamine, which is readily oxidized to a reactive quinone (13). Remarkably, inhibition of complex 1 by rotenone or MPTP leads to not only production of ROS but also a marked increase in RNS (4, 14, 15). This increase in RNA (NOS activity) may partly result from neuronal depolarization and voltage-dependent Ca²⁺ influx mediated by ATP depletion (16), but mitochondria themselves may also serve as sources of both calcium and NO (derived from NOS and nitrite) (9, 17). Note that NO itself, operating through S-nitrosylation, is a well-characterized inhibitor of complex 1 (18). These events would create a vicious cycle of mitochondrial dysfunction, RNS generation, and ultimately cell death. Thus, accumulating evidence suggests that dysregulated or excessive protein S-nitrosylation driven by mitochondrial regulation of NO production (in addition to protein oxidation induced by mitochondrial ROS production) contributes to neuronal cell death. A prime example of such dysregulation includes mitochondrial-dependent S-nitrosylation of parkin, an E3

ubiquitin ligase, which leads to inhibition of protein ubiquitylation, defective protein degradation, and apoptosis, all consistent with a neurodegenerative phenotype (14) (Figure 1).

Independent of the emerging role for NO in neurodegeneration, a large body of work indicates an important role for the ER in many types of apoptosis (19). The ER serves an essential role in the oxidative folding, glycosylation, and targeting of nascent proteins to proper subcellular and extracellular destinations. In contrast to the highly reducing environments of the cytosol and mitochondria, the ER is characterized by a relatively positive redox potential, reflected in part by a low ratio of reduced/oxidized glutathione (GSH/GSSG ~ 3/1), which promotes protein oxidation. Strict quality-control systems ensure that only correctly folded proteins are exported from the ER, while misfolded proteins are retrotranslocated to the cytosol for degradation (20). Accumulation of misfolded proteins in the ER, however, leads to ER stress that triggers the protective unfolded protein response (UPR). The UPR entails the induction of chaperone proteins, increased degradation of misfolded proteins, and inhibition of

protein translation (21). Prolonged ER stress can nonetheless lead to activation of apoptosis and cell death. Several members of the Bcl-2 family, as well as caspase-12, reside in the ER, and these apoptotic regulators may communicate stress signals between the mitochondria and ER (22, 23).

That aberrant protein folding is associated with various NDDs, including PD and AD, is well established; interest is growing in elucidating the role of ER stress in these pathologies (24). Experiments in pancreatic β -cells, macrophages (25), and cerebellar granule cells (26) have shown that NO can induce ER stress. But what is the molecular basis of NO-induced ER stress? And how could it promote neurodegeneration? The new findings of Uehara *et al.* (4) indicate that PDI may hold some of the answers to these questions.

PDI is a diverse family of thiol/disulfide oxidoreductases, each of which possesses two thioredoxin (Trx) domains (27). The Trx domains of PDI use a Cys-Gly-His-Cys (CGHC) motif to catalyze protein folding through two major pathways: oxidizing substrate cysteine thiols to intramolecular disulfides and isomerizing "incorrect" disulfides. As nascent proteins are translocated into the ER, their reduced cysteine thiols attack the electrophilic CGHC disulfide on PDI to generate a transient PDI-substrate intermolecular disulfide. Subsequently, a second reduced thiol of the substrate attacks the intermolecular disulfide to generate an oxidized substrate and reduced PDI (Figure 2, top). In the reduced form, PDIs also isomerize substrate disulfides through a single nucleophilic cysteine residue within the CGHC motif (Figure 2, bottom). To regenerate oxidized PDI, the ER employs a unique flavin-dependent oxidase (Ero1) that reduces O₂ to H₂O₂ with concomitant oxidation of PDI.

Uehara *et al.* (4) establish that PDI is S-nitrosylated (to yield SNO-PDI) following either mitochondrial insult (rotenone) or NMDAR activation and that SNO-PDI is

inhibited from performing its described functions. *In vitro*, SNO-PDI is inefficient at oxidizing reduced rhodanese or isomerizing scrambled RNase A. Within neurons, S-nitrosylation of PDI results in defective folding of synphilin (a feature of PD). Furthermore, SNO-PDI induces ER stress and subsequent activation of the UPR pathway. Consistent with these findings, overexpression of PDI guards against NMDAR-dependent ER stress, polyubiquitinylation, and apoptosis. This protection is dependent upon the active site cysteines of PDI. Thus, Uehara *et al.* (4) link two paradigmatic neurotoxic stimuli, NMDAR activation and mitochondrial dysfunction, to S-nitrosylation of PDI (Figure 1). The immediate relevance of these findings for the etiology of human NDD was established by the demonstration that SNO-PDI can be detected in the brains of patients with PD or AD, but not in normal brains.

Uehara and co-workers (4) used PDI mutants to show that each Trx domain in PDI can be S-nitrosylated, but the stoichiometry and target residues were not determined. In SNO-PDI, therefore, a single SNO or two SNOs could be associated with each Trx domain (Figure 1). In addition, it has been suggested that modification by NO of proteins possessing a pair of vicinal thiols (as in the case of PDI and several other redox enzymes) may take the form of a nitroxyl disulfide (28) (SNO₂; Figure 1); this modification would also be consistent with the analysis of Uehara *et al.* (4).

Because cellular reductants such as glutathione and ascorbate are thought to participate in SNO catabolism, the presence of SNO-PDI *in vivo* raises the possibility that the relatively oxidizing environment of the ER may contribute to the stability of protein SNOs. In this regard, it is of interest that a recent report observed SNO-proteins in association with the ER (29), suggesting that PDI may be one of numerous substrates for S-nitrosylation in this organelle. Preferential stability of SNO-proteins within the ER would be consistent with a model in which

dysregulated production of NO in NDDs induces protein S-nitrosylation in many subcellular regions; however, targets such as SNO-PDI would be sustained in the oxidizing environment of the ER. Further, since Ero1-dependent oxidation of PDI is a rate-limiting step in the redox cycle of PDI (30, 31), SNO-PDI may be trapped in a state that can neither participate in protein folding nor be regenerated by Ero1. This analysis suggests that ROS (oxidative stress) and RNS (nitrosative stress) may have substantially different effects in the ER: ROS-dependent oxidation of PDI would presumably generate an active site disulfide, which remains a substrate for Ero1 and thus functional. Nitrosative stress may thus be a more likely cause of ER stress.

Earlier studies have indicated that PDI can break down S-nitrosoglutathione (32) and, when localized to the plasma membrane, may participate in cellular import of extracellular S-nitrosothiols (33, 34). In both cases, SNO-PDI was a suggested (short-lived) intermediate. Similarly, *in vitro* S-nitrosylation of the human Trx active site leads rapidly to an intramolecular disulfide, with release of HNO (nitroxyl) (35–37). In contrast, Uehara *et al.* (4) demonstrate that the Trx domains of PDI are stable to S-nitrosylation *in vitro* and *in vivo*. The question is why. It is important to note that the geometry and dihedral strain in protein disulfides can vary substantially (the reduction potential of PDI is in fact 100 mV less negative than thioredoxin (38)) and that the effects of S-nitrosylation

on either the geometry of vicinal thiols that serve as catalytic disulfides or on the global structure of PDI have not been studied. Thus, while a kinetic barrier to disulfide formation most likely contributes to the stability of SNO-PDI (39, 40), a thermodynamic explanation is not excluded. In addition, the SNO in PDI might reflect the presence of a double SNO or a SNO₂ (Figure 1), with lower reactivity than the single SNO. The biochemical and cellular properties that lead to such differences in SNO protein stability, as well as the roles of S-nitrosylation in the ER, are interesting topics for future investigation.

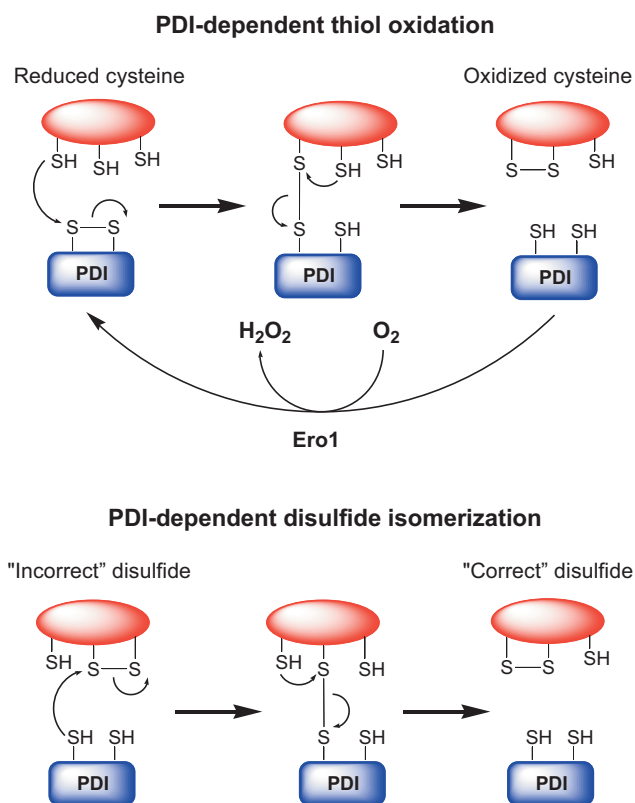


Figure 2. PDI catalyzes protein folding in the ER by oxidizing nascent protein thiols to disulfides (top) and rearranging “incorrect” disulfides (bottom). The flavoprotein Ero1 regenerates oxidized PDI by transferring electrons to molecular oxygen.

REFERENCES

- Garthwaite, J., and Boulton, C. L. (1995) Nitric oxide signaling in the central nervous system, *Annu. Rev. Physiol.* **57**, 683–706.
- Calabrese, V., Bates, T. E., and Stella, A. M. (2000) NO synthase and NO-dependent signal pathways in brain aging and neurodegenerative disorders: the role of oxidant/antioxidant balance. *Neurochem. Res.* **25**, 1315–1341.
- Mattson, M. P. (2000) Apoptosis in neurodegenerative disorders, *Nat. Rev. Mol. Cell Biol.* **1**, 120–129.
- Uehara, T., Nakamura, T., Yao, D., Shi, Z. Q., Gu, Z., Ma, Y., Masliah, E., Nomura, Y., and Lipton, S. A. (2006) S-nitrosylated protein-disulfide isomerase links protein misfolding to neurodegeneration, *Nature* **441**, 513–517.
- Dawson, V. L., Dawson, T. M., London, E. D., Bredt, D. S., and Snyder, S. H. (1991) Nitric oxide mediates glutamate neurotoxicity in primary cortical cultures, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6368–6371.
- Jaffrey, S. R., Erdjument-Bromage, H., Ferris, C. D., Tempst, P., and Snyder, S. H. (2001) Protein S-nitrosylation: a physiological signal for neuronal nitric oxide, *Nat. Cell Biol.* **3**, 193–197.
- Hara, M. R., Agrawal, N., Kim, S. F., Cascio, M. B., Fujimuro, M., Ozeki, Y., Takahashi, M., Cheah, J. H., Tankou, S. K., Hester, L. D., Ferris, C. D., Hayward, S. D., Snyder, S. H., and Sawa, A. (2005) S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding, *Nat. Cell Biol.* **7**, 665–674.
- Gu, Z., Kaul, M., Yan, B., Kridel, S. J., Cui, J., Strongin, A., Smith, J. W., Liddington, R. C., and Lipton, S. A. (2002) S-nitrosylation of matrix metalloproteinases: signaling pathway to neuronal cell death, *Science* **297**, 1186–1190.
- Chen, Z., Foster, M. W., Zhang, J., Mao, L., Rockman, H. A., Kawamoto, T., Kitagawa, K., Nakayama, K. I., Hess, D. T., and Stamler, J. S. (2005) An essential role for mitochondrial aldehyde dehydrogenase in nitroglycerin bioactivation, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12159–12164.
- Sydow, K., Daiber, A., Oelze, M., Chen, Z., August, M., Wendt, M., Ullrich, V., Mulsch, A., Schulz, E., Keane, J. F., Jr., Stamler, J. S., and Munzel, T. (2004) Central role of mitochondrial aldehyde dehydrogenase and reactive oxygen species in nitroglycerin tolerance and cross-tolerance, *J. Clin. Invest.* **113**, 482–489.
- Zhang, L., Dawson, V. L., and Dawson, T. M. (2006) Role of nitric oxide in Parkinson's disease. *Pharmacol. Ther.* **109**, 33–41.
- Almeida, A., Moncada, S., and Bolanos, J. P. (2004) Nitric oxide switches on glycolysis through the AMP protein kinase and 6-phosphofructo-2-kinase pathway, *Nat. Cell Biol.* **6**, 45–51.
- LaVoie, M. J., Ostaszewski, B. L., Weihofen, A., Schlossmacher, M. G., and Selkoe, D. J. (2005) Dopamine covalently modifies and functionally inactivates parkin, *Nat. Med.* **11**, 1214–1221.
- Chung, K. K., Thomas, B., Li, X., Pletnikova, O., Troncoso, J. C., Marsh, L., Dawson, V. L., and Dawson, T. M. (2004) S-nitrosylation of parkin regulates ubiquitination and compromises parkin's protective function, *Science* **304**, 1328–1331.
- He, Y., Imam, S. Z., Dong, Z., Jankovic, J., Ali, S. F., Appel, S. H., and Le, W. (2003) Role of nitric oxide in rotenone-induced nigro-striatal injury. *J. Neurochem.* **86**, 1338–1345.
- Abou-Sleiman, P. M., Muqit, M. M., and Wood, N. W. (2006) Expanding insights of mitochondrial dysfunction in Parkinson's disease, *Nat. Rev. Neurosci.* **7**, 207–219.
- Kozlov, A. V., Staniek, K., and Nohl, H. (1999) Nitrite reductase activity is a novel function of mammalian mitochondria, *FEBS Lett.* **454**, 127–130.
- Clementi, E., Brown, G. C., Feelisch, M., and Moncada, S. (1998) Persistent inhibition of cell respiration by nitric oxide: crucial role of S-nitrosylation of mitochondrial complex I and protective action of glutathione, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7631–7636.
- Xu, C., Bailly-Maitre, B., and Reed, J. C. (2005) Endoplasmic reticulum stress: cell life and death decisions, *J. Clin. Invest.* **115**, 2656–2664.
- Ellgaard, L., and Helenius, A. (2003) Quality control in the endoplasmic reticulum, *Nat. Rev. Mol. Cell Biol.* **4**, 181–191.
- Zhang, K., and Kaufman, R. J. (2006) The unfolded protein response: a stress signaling pathway critical for health and disease, *Neurology* **66**, S102–S109.
- Breckenridge, D. G., Germain, M., Mathai, J. P., Nguyen, M., and Shore, G. C. (2003) Regulation of apoptosis by endoplasmic reticulum pathways, *Oncogene* **22**, 8608–8618.
- Xu, W., Liu, L., Charles, I. G., and Moncada, S. (2004) Nitric oxide induces coupling of mitochondrial signalling with the endoplasmic reticulum stress response, *Nat. Cell Biol.* **6**, 1129–1134.
- Lindholm, D., Wootz, H., and Korhonen, L. (2006) ER stress and neurodegenerative diseases, *Cell Death Differ.* **13**, 385–392.
- Gotoh, T., and Mori, M. (2006) Nitric oxide and endoplasmic reticulum stress, *Arterioscler. Thromb., Vasc. Biol.* **26**, 1439–1446.
- He, J., Kang, H., Yan, F., and Chen, C. (2004) The endoplasmic reticulum-related events in S-nitrosoglutathione-induced neurotoxicity in cerebellar granule cells, *Brain Res.* **1015**, 25–33.
- Ellgaard, L., and Ruddock, L. W. (2005) The human protein disulfide isomerase family: substrate interactions and functional properties, *EMBO Rep.* **6**, 28–32.
- Houk, K. N., Hietbrink, B. N., Bartberger, M. D., McCarran, P. R., Choi, B. Y., Voyksner, R. D., Stamler, J. S., and Toone, E. J. (2003) Nitroxyl disulfides, novel intermediates in transnitrosation reactions, *J. Am. Chem. Soc.* **125**, 6972–6976.
- Greco, T. M., Hodara, R., Parastatidis, I., Heijnen, H. F., Dennehy, M. K., Liebler, D. C., and Ischiropoulos, H. (2006) Identification of S-nitrosylation motifs by site-specific mapping of the S-nitrosocysteine proteome in human vascular smooth muscle cells, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7420–7425.
- Mezghrani, A., Fassio, A., Benham, A., Simmen, T., Braakman, I., and Sitia, R. (2001) Manipulation of oxidative protein folding and PDI redox state in mammalian cells, *EMBO J.* **20**, 6288–6296.
- Molteni, S. N., Fassio, A., Ciriolo, M. R., Filomeni, G., Pasqualeto, E., Fagioli, C., and Sitia, R. (2004) Glutathione limits Ero1-dependent oxidation in the endoplasmic reticulum, *J. Biol. Chem.* **279**, 32667–32673.
- Sliskovic, I., Raturi, A., and Mutus, B. (2005) Characterization of the S-denitrosation activity of protein disulfide isomerase, *J. Biol. Chem.* **280**, 8733–8741.
- Ramachandran, N., Root, P., Jiang, X. M., Hogg, P. J., and Mutus, B. (2001) Mechanism of transfer of NO from extracellular S-nitrosothiols into the cytosol by cell-surface protein disulfide isomerase, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9539–9544.
- Zai, A., Rudd, M. A., Scribner, A. W., and Loscalzo, J. (1999) Cell-surface protein disulfide isomerase catalyzes transnitrosation and regulates intracellular transfer of nitric oxide, *J. Clin. Invest.* **103**, 393–399.
- Mitchell, D. A., and Marletta, M. A. (2005) Thioredoxin catalyzes the S-nitrosation of the caspase-3 active site cysteine, *Nat. Chem. Biol.* **1**, 154–158.
- Stoyanovsky, D. A., Tyurina, Y. Y., Tyurin, V. A., Anand, D., Mandavia, D. N., Gius, D., Ivanova, J., Pitt, B., Billiar, T. R., and Kagan, V. E. (2005) Thioredoxin and lipoic acid catalyze the denitrosation of low molecular weight and protein S-nitrosothiols, *J. Am. Chem. Soc.* **127**, 15815–15823.
- Amelle, D. R., and Stamler, J. S. (1995) NO⁺, NO, and NO⁻ donation by S-nitrosothiols: implications for regulation of physiological functions by S-nitrosylation and acceleration of disulfide formation, *Arch. Biochem. Biophys.* **318**, 279–285.
- Lundstrom, J., and Holmgren, A. (1993) Determination of the reduction-oxidation potential of the thioredoxin-like domains of protein disulfide isomerase from the equilibrium with glutathione and thioredoxin, *Biochemistry* **32**, 6649–6655.
- Schmidt, B., Ho, L., and Hogg, P. J. (2006) Allosteric disulfide bonds, *Biochemistry* **45**, 7429–7433.
- Stamler, J. S., and Toone, E. J. (2002) The decomposition of thionitrites, *Curr. Opin. Chem. Biol.* **6**, 779–785.

The Use of Small Molecules to Investigate Molecular Mechanisms and Therapeutic Targets for Treatment of Botulinum Neurotoxin A Intoxication

Tobin J. Dickerson* and Kim D. Janda*

Departments of Chemistry and Immunology, The Skaggs Institute for Chemical Biology, and Worm Institute for Research and Medicine (WIRM), The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

C*lostridium botulinum* and the disease it causes, botulism, have been known for centuries. This bacterium has been studied for >100 years, with a number of publications dating back to the early 19th century. Case studies were reported as far back as 1815; however, it was not until van Ermengem isolated an anaerobic bacillus from contaminated meat in 1897 that the causative agent of botulism was discovered (1). Cultivation of the bacillus and subsequent introduction into animals led to the development of the symptoms of botulism. The discovery that the organism produced a toxin occurred shortly after the identification of the bacillus; however, it was only in the final decade of the 20th century that a full picture of the toxin's structure, mechanism of action, and target substrates emerged (2).

Structurally, *C. botulinum* is a rod-shaped, Gram-positive, sporulating anaerobic bacillus that is widely distributed in the environment (3). Neurotoxins produced by *C. botulinum* are some of the most potent naturally occurring compounds known; the lethal dose for humans is $\sim 1 \text{ ng kg}^{-1}$ of body weight (4). Their exquisite toxicity, coupled with their highly specific mechanism of action, renders the botulinum neurotoxins (BoNTs) both highly dangerous and yet quite useful to medical science (5). BoNTs are typically associated with food poisoning, although they also are seen as a result of wound infections, inhalation, or as a colonizing infection in the intestinal tract of infants (6). BoNTs became a common public-health threat only after the advent of food preservation in the 19th century. Modern food-preparation practices have rendered botulism a rare occurrence from commercially prepared foods,

ABSTRACT Botulinum neurotoxins (BoNTs) are agents responsible for botulism, a disease characterized by peripheral neuromuscular blockade and subsequent flaccid paralysis. The potent paralytic ability of these toxins has resulted in their use as a therapeutic; however, BoNTs are also classified by the Centers for Disease Control and Prevention as one of the six highest-risk threat agents of bioterrorism. Consequently, a thorough understanding of the molecular mechanism of BoNT toxicity is crucial before effective inhibitors and, ultimately, an approved drug can be developed. In this article, we systematically detail BoNT intoxication by examining each of the discrete steps in this process. Additionally, rationally designed strategies for combating the toxicity of the most potent BoNT serotype are evaluated.

*Corresponding authors,
tobin@scripps.edu,
kdjanda@scripps.edu.

Received for review April 25, 2006
and accepted May 26, 2006.

Published online July 21, 2006
10.1021/cb600179d CCC: \$33.50

© 2006 by American Chemical Society

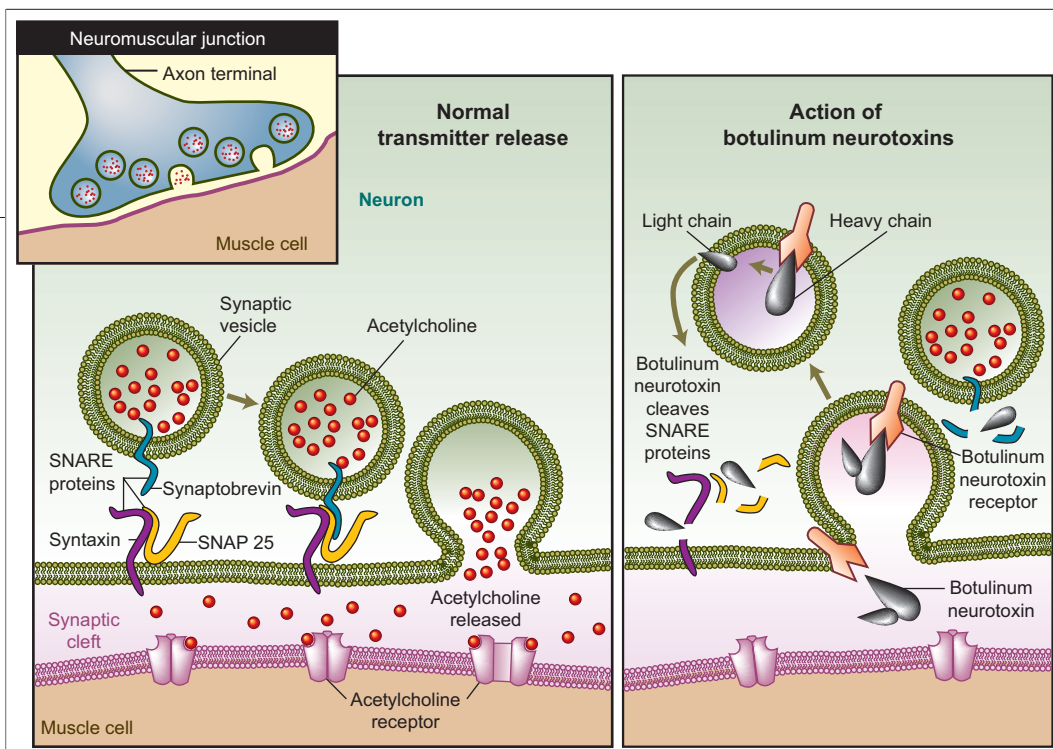


Figure 1. Mechanism of action of BoNTs. (Adapted with permission from Rowland, L. P. (2002) *New Engl. J. Med.* 347, 382. Copyright 2002 Massachusetts Medical Society). Shown are the individual stages of BoNT intoxication, including cell surface recognition, vesicle internalization, translocation of the LC protease into the cytosol, and proteolytic cleavage of one of the proteins of the SNARE complex. These steps lead to inhibition of neurotransmitter-containing vesicle release. BoNT/B, -D, -F, and -G cleave proteins of the VAMP family (blue), and BoNT/A, -C, and -E cleave SNAP-25 (yellow). BoNT/C can also cleave syntaxin (purple).

ease of production and transport, and need for prolonged intensive care, BoNTs are classified by the Centers for Disease Control and Prevention (CDC) as one of the six highest-risk threat agents for bioterrorism (Category A agents).

Yet, food poisoning and weapons of mass destruction are only two scenarios in which BoNTs play a role. Despite being extremely poisonous, BoNT is a highly effective therapeutic agent and valuable research tool (9).

although a small but significant number of cases occur annually from eating canned foods.

Botulism is characterized by generalized muscle weakness, and in more severe cases, it results in impaired respiratory function and autonomic dysfunction. In most severe cases, this leads to respiratory failure and death. One of the most fascinating aspects of BoNT intoxication is that host death does not result from target cell death and subsequent accumulated tissue destruction. Rather, death of the host results from a secondary event (*e.g.*, respiratory failure) that depends on toxin-induced inactivation of neurotransmitter release from otherwise viable nerve cells.

As previously stated, BoNT poisoning can occur accidentally via food consumption; however, incident rates are low. The major concern is its malicious use, especially in bioterrorism and biowarfare. During World War II, the extremely high potency of BoNTs induced both the Allied and Axis powers to evaluate these proteins as potential biological warfare agents (7). This work formally ended with the signing of the 1972 Biological Weapons Convention. However, recent events in the Middle East and Asia have confirmed the weaponization of this toxin by the former Soviet Union and by the Iraqi military before and during the 1991 Gulf War; the Japanese cult Aum Shinrikyo tried to use BoNT for bioterrorism (8). Because of their extreme potency and lethality,

At the beginning of the 20th century, it was observed that injecting BoNT could paralyze individual muscle groups without giving the recipient botulism (10). Indeed, preparations of BoNT serotype A (BoNT/A) have been approved by the U.S. Food and Drug Administration for use in treating strabismus, blepharospasm, and hemifacial spasm (11). The use of BoNT has also been extended to cover a wide variety of disorders, including those that do not have a neuromuscular basis (12), such as axillary hyperhidrosis (excessive sweating), myofascial pain and tension, migraine headaches, and multiple sclerosis. In addition, polypeptide fragments derived from the toxin are being evaluated as potential carrier molecules in the construction of oral and inhalation vaccines (13).

Molecular Mechanism of BoNT and Neurotransmitter Release. BoNT has seven serologically distinct serotypes (A–G); these proteins have a molecular weight of ~150,000 kDa (14). BoNTs are synthesized as single-polypeptide chains, and cleavage by intra- or extracellular proteases converts them into dimers consisting of a 100-kDa heavy chain (HC) coupled to a 50-kDa light chain (LC) by one or more disulfide bonds. The two-chain molecule is the active form of the toxin that poisons cholinergic transmission. Each serotype is produced as the primary toxin by a specific strain of bacterium and, although they share a high degree of homol-

ogy, they differ in their toxicity and molecular site of action. BoNT intoxication (as summarized in Figure 1) occurs through a multistep process involving each of the toxin functional domains and can be described as the outcome of discrete stages (15, 16).

Binding to the Target Cell and Internalization. BoNTs bind to cholinergic nerve terminals by their HC domains and are subsequently internalized by receptor-mediated endocytosis. The identity of the putative receptor used by BoNT/A has recently been reported to be the synaptic vesicle protein SV2 (17, 18). Additionally, a second component of cellular recognition is thought to be due to low-affinity interactions between the toxin and the gangliosides (19). A double-receptor model has been proposed whereby BoNTs also bind to both of these receptors before internalization can occur (19). In such a scenario, within the first step, the toxin associates with the cell membrane and the gangliosides in a low-affinity complex. Next, this complex migrates laterally until it interacts with a high-affinity binding site; the latter interaction then allows subsequent events, such as receptor-mediated endocytosis. Toxin binding to receptors also appears to be serotype-specific (20, 21). The precise identities of the receptors for each of the BoNTs have remained elusive. However, this area of research remains highly studied because the elucidation of these receptors may lead to novel therapeutics for the treatment of botulism.

Receptor-Mediated Endocytosis. Most researchers presume that the process of BoNT receptor-mediated endocytosis is basically identical to that of most ligands that are internalized into cells (16). In fact, this may be the case; however, it has been noted that a retrieval phase of the vesicle recycling mechanism may also be plausible (22). Nerves that exocytose have a vigorous and well-developed mechanism for membrane retrieval (23); such a recycling mechanism could be a viable route by which the toxin can enter nerve cells. Synaptotagmin II, a protein that serves as a receptor for BoNT/B, shows that vesicles are involved in this process. Synaptotagmin has an exposed domain in the lumen of vesicles. Thus, in the cycle, exocytosis would place synaptotagmin on the exterior of the nerve cell for a brief time, wherein toxin binding to its receptor(s) would occur and both synaptotagmin and BoNT would be internalized during membrane retrieval. Labeled derivatives of synaptotagmin antibodies have been used to monitor

membrane retrieval and reformation of intraneuronal vesicles in order to support this theory (24).

Translocation. It has been proposed that a pH-dependent structural rearrangement of the toxin inside an acidic compartment within the cell allows for toxin entry into the cytosol, a process common for several other bacterial toxins (25). Thus, the substrates for BoNTs are in the cytosol, and the LC protease must escape the endosome. Translocation is believed to take place wherein buried endosomal domains are exposed as the pH decreases (16). These domains then facilitate penetration of the lipid bilayer in a way that promotes translocation of the active region to the cytosol. This mechanism has been investigated via the pretreatment of neuromuscular junctions with chloroquine, a small molecule that can effectively and specifically raise the endosomal pH (26). This approach represents the first nonpeptidic approach for BoNT antagonism by preventing toxic escape from the endosome.

The notion that BoNT is internalized by pH-induced translocation is now widely accepted, but the exact nature of the membrane penetration remains unclear. Studies have been conducted that measured the change in resistance of artificial membranes as a function of the location of the toxin in an effort to clarify this mechanism (27). More recently, a new perspective has taken shape in which it has been proposed that the HC of BoNT can act both as a channel and as a chaperone (28). Substantiation of this hypothesis is given by the fact that BoNT/A and -E form ion channels in phospholipid bilayers and PC12 cell membranes under conditions similar to those believed to exist *in vivo*.

Inhibition of Neurotransmitter Release. SNARE proteins are involved in the fusion of synaptic vesicles with the plasma membrane; thus, the action of BoNT is to prevent exocytosis (29). At a more specific level, cleavage of SNARE proteins by BoNT inhibits the release of acetylcholine at the neuromuscular junction; this leads to inhibition of neurotransmission (15, 16). Cleavage of individual SNARE proteins does not prevent SNARE complex formation but results in a nonfunctional complex in which the coupling between Ca^{2+} influx and fusion is disrupted (30). The role of Ca^{2+} is fundamental to the process of BoNT-dependent inhibition of neurotransmitter release, because increasing the Ca^{2+} concentration in the synaptic terminal partially reverses the effect of BoNT/A (31).

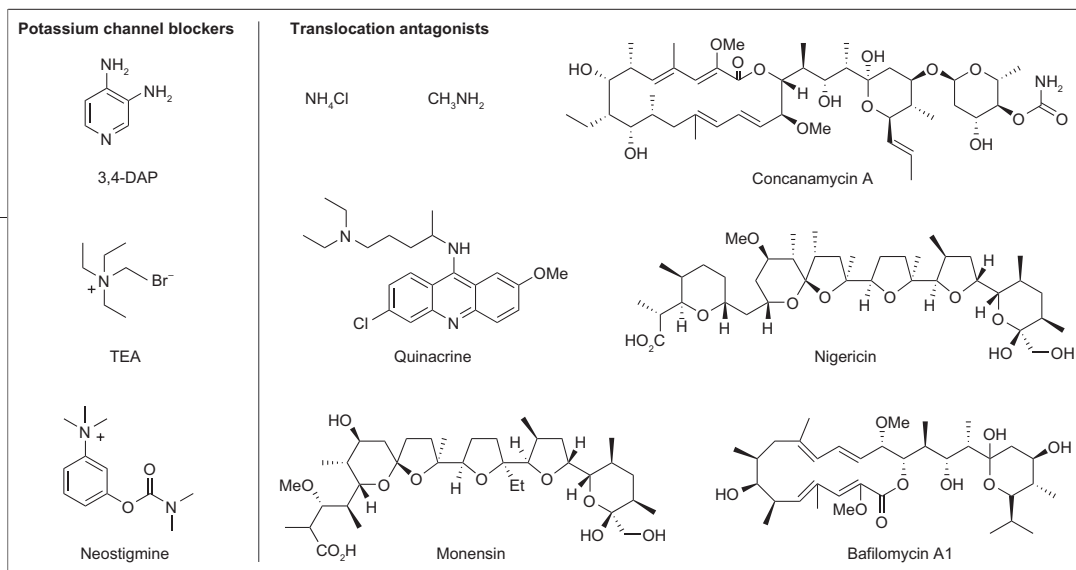


Figure 2. Chemical structures of various compounds used to block BoNT intoxication by inhibiting either potassium channels or toxin translocation out of endosomes.

Details defining the cleavage of the SNARE proteins on a molecular level (Figure 2) can be traced back to the BoNT LC, which functions as a zinc-dependent metalloprotease (32). A highly conserved segment of 20 amino acids located within these LCs displays the common zinc endopeptidase motif His-Glu-Xaa-Xaa-His. Each of the seven serotypes of BoNT cleaves one of the three SNARE proteins [synaptobrevin (vesicle-associated membrane protein, VAMP), SNAP-25 (synaptosomal-associated protein of 25 kDa), and syntaxin], which are necessary for vesicle fusion and acetylcholine release. It has been firmly established that VAMP is the target for BoNT/B, -D, -F, and -G. The target for BoNT/A and -E is SNAP-25, whereas BoNT/C can cleave both SNAP-25 and syntaxin.

KEYWORDS

Botulinum neurotoxin (BoNT): One of the seven related proteins (A–G) secreted by *Clostridium botulinum* that lead to botulism. These toxins are among the most toxic species known, and certain preparations are commercially called Botox. The molecule is composed of two chains, termed heavy and light chains. The heavy chain is critical in toxin recognition of the target cellular surface and translocation out of endosomal vesicles into the cytoplasm, whereas the light chain is a zinc-dependent metalloprotease that cleaves specific SNARE proteins within the cell.

Botulism: The bacterium *C. botulinum* causes this disease, which results in peripheral nerve paralysis and can ultimately lead to death. The causative agent is BoNT, a proteinaceous toxin.

Cholinergic transmission: A mechanism of neurotransmission that employs acetylcholine as a neurotransmitter. Botulism interferes with this process by preventing the release of neurotransmitter at synapses, thus leading to flaccid paralysis.

Therapeutic Strategies for Treating BoNT Intoxication. Because of their exceptional potency and ease of production, BoNTs are formidable biothreat agents. In particular, BoNT/A is considered the deadliest serotype, and as such, we have focused our discussion of therapeutic approaches on this specific toxin. Currently, no approved pharmacological treatments exist for BoNT intoxication. Although an effective vaccine is available for immunoprophylaxis (33), the development of protection is slow, and multiple inoculations and annual boosters are required to

produce adequate antibody titers. Additionally, the CDC also distributes two equine antitoxins for treatment of adult botulism. Although equine antibodies are broadly effective, they can cause adverse reactions, such as serum sickness and anaphylaxis (34). Clearly, a pharmacological intervention, especially one that would be effective after BoNT internalization, is highly desirable. More recently, the orphan drug Human Botulism Immune Globulin (BIG) has been developed as a therapeutic antitoxin that neutralizes BoNT. Notably, this is the first licensed product for the treatment of patients suffering from botulism. Although this antibody approach has proven efficacy in the treatment of infant botulism (35), antibodies are not a small-molecule therapeutic approach and thus are beyond the scope of this article (36, 37).

Potassium Channel Blockers. Potassium channels play crucial physiological roles in almost all types of cells in all organisms (38). In brief, potassium channels form a remarkably diverse group of ion channel structures; the first type of potassium channel to be described was the classic voltage-activated channel found in the squid. Other types of potassium channels include hyperpolarizing voltages that are modulated by intracellular metabolites and second messengers, transient outward currents or A-currents, and large-conductance Ca²⁺-sensitive, K⁺ channels, or BKCa channels, in which neurons are characterized by calcium-activated potassium currents that contribute to re-polarization and firing. BKCa channels exist as a complex of two different subunits, the pore-forming α subunit and a β regulatory subunit. BKCa channels are generally believed to be tetraethylammonium (TEA)-sensitive, and they are therapeutic targets for BoNT poisoning.

The best-known K⁺ channel blocker for botulism poisoning is 3,4-diaminopyridine (3,4-DAP) (Figure 2) (39). 3,4-DAP is highly effective in antagonizing muscle paralysis after BoNT/A exposure *in vitro* and is the least toxic of the currently available K⁺ blockers. After a rat diaphragm muscle was paralyzed by BoNT/A exposure, 3,4-DAP induced a rapid and pronounced increase of

twitch tensions (39). 3,4-DAP continued to work, and compared with the control, little or no decrease was seen for ≥ 8 h after its addition. Combining *in vivo* and *in vitro* recording techniques, Adler and co-workers (40) studied the actions of 3,4-DAP in the rat extensor digitorum longus muscle after local inhibition of neuromuscular transmission by BoNT. The results showed that 3,4-DAP markedly potentiated twitch tensions in BoNT/A-intoxicated muscle. The sensitivity of the extensor digitorum longus muscle to 3,4-DAP did not diminish with time or with repeated application. Because a major obstacle to the clinical use of 3,4-DAP is its brief duration, other investigations have attempted to deliver the compound via osmotic minipump infusion (41). It was also found that, for muscle function to be maintained, the drug must be delivered continuously for the entire period of BoNT intoxication.

3,4-DAP is an impressive compound for BoNT treatment; however, several problems have been associated with its use. First, its efficacy is primarily limited to BoNT/A. Second, while it seems to be very beneficial for increasing muscle strength, it only provides limited improvement for respiratory muscles, and no spontaneous ventilation has been seen. Third, it has toxicity issues, and seizures have been noted with its use, mainly from its penetration across the blood–brain barrier. To surmount some of these difficulties, researchers combined 3,4-DAP with neostigmine and TEA bromide (39). Researchers had hoped that combining 3,4-DAP and TEA, both of which are K^+ channel blockers, would reduce the dose of 3,4-DAP needed and its toxic effects. Furthermore, this combination would have produced a greater Ca^{2+} influx than was attainable with either inhibitor alone. Accordingly, the increased Ca^{2+} entry should have resulted in an enhanced acetylcholine release. However, inhibition rather than enhancement was observed; this suggests that the postsynaptic inhibitory action of TEA on the nicotinic ion channel was sufficient to counteract any beneficial action of TEA on transmitter release. Although TEA was unsuccessful, this concept of “combinatorial” channel blockers is still believed to be valid for the treatment of BoNT poisoning. The second combination therapy of neostigmine and 3,4-DAP was considered viable, because neostigmine would increase the persistence of acetylcholine by inhibiting the activity of acetylcholinesterase in peripheral tissues. However, this combination was insufficient to restore neuromuscular transmission.

The efficacy of potassium channel blockers in antagonizing the action of BoNT/A is generally attributed to their ability to enhance the influx of Ca^{2+} as a result of inhibiting voltage-dependent K^+ currents. Substances that can increase intracellular Ca^{2+} can partially overcome the paralysis due to BoNT poisoning. Continued efforts to optimize agents such as 3,4-DAP to increase their efficacy and reduce their toxicity are thus considered of significant interest. However, debate exists about the viability of this approach in the treatment of BoNT intoxication; for a potassium channel blocker to be effective, it must be administered over a period of weeks or months.

Antagonists of Toxin Binding to Target Cells. The initial step in the mechanism of BoNT poisoning is the binding of the toxin to the cellular membranes of target neurons. Thus, blocking the interaction between BoNT and the cognate cellular receptor can inhibit nerve paralysis. Monoclonal antibodies have been investigated in this role and have been reviewed elsewhere (36, 37). Two approaches can be envisioned to accomplish the goal of antagonizing toxin–cell interactions: molecules that can coat the toxin with a small molecule and interfere with its ability to interact with a cell, and molecules that bind to the cellular receptor, thus blocking toxin binding. In the former case, polysialylated gangliosides such as GT1b were observed >30 years ago to be potential receptors for BoNT/A (42, 43) and more recently were shown to inhibit BoNT/A binding to synaptosomes (44) and to quench BoNT/A fluorescence (45). In the instance of molecules that compete with BoNT for the cellular binding site, lectins from two sources, one derived from animals and the other from plants (*Limax flavus* and *Triticum vulgare*, respectively, both of which possess affinity for sialic acid), could serve as competitive antagonists of all BoNT serotypes as well as tetanus toxins (46). In the context of BoNT/A, the amount of time necessary to cause neuromuscular paralysis in mouse phrenic-nerve hemidiaphragm preparations was increased from 78 to 128 min. Importantly, this study was the first to report on a small molecule that could potentiate the activity of all BoNT serotypes.

Antagonists of pH-Dependent BoNT Translocation. The lethal effects of BoNT involve the inhibition of synaptic transmission at the skeletal neuromuscular junction. Despite variations in the SNARE protein targets of different BoNT serotypes, a common point in the action of each of these toxins is the encapsulation of the holo-

Another possible therapy for the inhibition of BoNT neurotoxicity is to limit the metallopeptidase activity of the BoNT light chains.

toxin in an endocytotic vesicle that then undergoes acidification. This acidification is needed for BoNT to induce muscle failure; agents that inhibit acidification delay the onset of the paralysis *in vitro* and lengthen the time the toxin is susceptible to neutralization with antisera (47). Thus, endocytotic vesicle acidification is a logical point at which to inhibit the activity of BoNT serotypes.

Simpson (48) has reported that ammonium chloride and methylamine produce concentration- and time-dependent antagonism of the onset of neuromuscular blockade caused by BoNT/A–C. These amines exerted their effects only when they were added before the toxin was introduced or ~10–20 min later. At concentrations that produce antagonism of BoNT-induced paralysis onset, these amines did not inactivate toxin molecules nor did they produce irreversible changes in tissue function. Presumably, the amines did not inhibit BoNT from binding to its receptor(s) and did not reverse neuromuscular blockade; rather, they acted solely to antagonize internalization of the toxins.

Acidification of endocytotic vesicles depends on a vesicular H⁺-ATPase that acts as a proton pump to accumulate protons from the cytoplasm into the lumen of a vesicle. In 1994, Simpson (49) reported that bafilomycin A1, an inhibitor of this

ATPase, is a universal antagonist of BoNTs (Figure 2). This compound produced a concentration-dependent blockage of neuromuscular transmission without affecting nerve- or muscle-action potentials. Application of proton ionophores can also deplete this pH gradient without affecting ATP hydrolysis. Sheridan (50) presented evidence that two monocarboxylate polyether ionophores, nigericin and monensin (Figure 2), increase membrane permeability to H⁺ and K⁺, or H⁺, Na⁺, and K⁺, respectively, and block endosomal acidification by acting as proton shunts to neutralize pH gradients.

Nanomolar concentrations of both compounds were able to block BoNT effects at neuromuscular junctions up to 3-fold times more than unprotected muscles. Unfortunately, higher concentrations of these antibiotics also blocked synapses. More recently, concanamycin A (Con A) has also been examined as an inhibitor of endosomal acidification (Figure 2) (51). Interestingly, Con A prevented SNAP-25 cleavage in pretreated cultures or those treated up to 15 min after toxin exposure, whereas the addition of the compound 40 min later was not protective.

Groups have also examined several clinically used antimalarial drugs (aminoquinolines) for their effectiveness in antagonizing BoNT/A-induced neuromuscular blockade (52). These studies concluded that these compounds may produce their protective efficacy through the blockade of endosomal acidification. Alternatively, it has been suggested that, because these drugs block channel formation, they may be a key element in LC release from the endosome. Lastly, Adler and co-workers (53) have investigated drug combinations for additive or synergistic effects. Here, quinacrine and the metal chelator *N,N,N',N'*-tetrakis(2-pyridylmethyl)-ethylenediamine (TPEN), when added together, protected up to 100% more than either drug alone.

Inhibition of the BoNT Metalloprotease LC. As detailed earlier, BoNT impairs neuronal exocytosis through specific proteolysis of the SNARE proteins. SNARE assembly into a low-energy ternary complex is thought to be intimately involved with membrane fusion and thus neurotransmitter release (54). The site-specific SNARE hydrolysis is catalyzed by the BoNT LC, which comprises a unique group of endopeptidases. Thus, another possible therapy for the inhibition of BoNT neurotoxicity is to limit the metallopeptidase activity of the BoNT LC.

Current research efforts to identify BoNT protease inhibitors have concentrated only on BoNT/A and -B, primarily because of their potent toxicity and extended duration. Two approaches can be envisioned for inhibitor development, peptide sequences/peptidomimetics based upon the native SNARE protein substrate or small organic molecules that specifically bind to the toxin and inactivate it. Several effective inhibitors based upon peptide scaffolds have been reported (55–57); however, these molecules are unlikely to become leads for new pharmaceuticals because of their short *in vivo* lifetimes. Initial efforts in this area focused on the use of

KEYWORDS

***Clostridium botulinum*:** A rod-shaped, spore-forming bacterium that is widely distributed in the environment. The neurotoxin secreted by this bacterium, BoNT, is among the most toxic species known.

Ganglioside: A molecule that is composed of a lipid domain and a carbohydrate domain that contains one or more sialic acid moieties. These compounds are embedded within the plasma membrane of numerous cell types and are critically important in signal transduction events.

Lectin: Receptor proteins that specifically bind to carbohydrates. Lectins are found in both plants and animals and have diverse roles in each. They play a large role in the normal function of the immune system.

SNARE protein: Acronym for “soluble *N*-ethylmaleimide sensitive factor attachment receptor”. These proteins make up a large superfamily whose primary function is to mediate vesicle fusion in mammalian and yeast cells.

Synaptosome: The isolated synapse of a neuron purified by homogenization of nerve tissue commonly used in assays addressing BoNT paralysis.

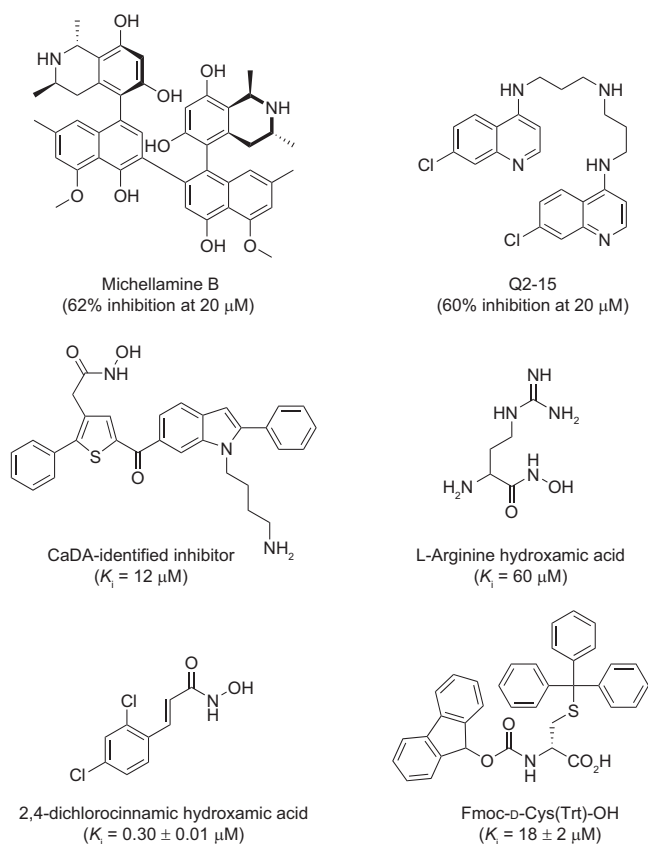


Figure 3. BoNT/A LC metalloprotease inhibitors.

heavy-metal chelators that possessed cell permeability as potential therapeutic targets. In particular, TPEN has been extensively studied for its ability to antagonize the catalytic action of BoNT in mouse phrenic-nerve hemidiaphragm preparations and in a mouse model (42, 58, 59). However, TPEN does not possess any inherent specificity for BoNT over other metalloproteins; this may preclude its use in a clinical setting because of various side effects that could result from chelation of other critical metals.

Other studies targeted toward the identification of low-molecular-weight, small-molecule inhibitors of BoNT/A include screening of the National Cancer Institute Diversity Set and a series of 4-aminoquinolines that were originally identified to prolong the time required for BoNT/A to block neuromuscular transmission (Figure 3) (60). Interestingly, several compounds possessing >50% inhibition (at 20- μM concentration) were identified. From these lead “hits”, modeling studies predicted

common pharmacophore scaffolds; however, these initial leads have yet to yield any new compounds of greater potency. An *in silico* screen of 2.5 million compounds has also been conducted recently in an effort to isolate BoNT/A-selective inhibitors. After extensive structure-guided modification of the selected scaffolds, an inhibitor with a K_i of 12 μM was found (61). This inhibitor was designed to display competitive kinetics by chelating the active site zinc atom through a hydroxamic acid moiety.

The success of this study validates the use of the cationic dummy atom approach (CaDA) to develop BoNT inhibitors by molecular dynamics simulations. However, a key conclusion in that article, that effective BoNT inhibition requires a length of 10 atoms, is not supported by our own studies on BoNT metalloprotease inhibitors.

Recently, we have initiated a multifaceted research program aimed at identifying novel small-molecule inhibitors of the BoNT/A LC metalloprotease. Given the presence of a critical zinc ion in the LC active site, we speculated that the hydroxamate zinc-binding functionality, when coupled to a suitable scaffold to impart

specificity, would provide potent BoNT/A inhibitors. Our initial studies reported that arginine hydroxamic acid, a single modified amino acid, could modestly inhibit BoNT/A in a high-throughput FRET-based assay developed for the screening of compound libraries ($K_i = 60 \mu\text{M}$) (62). The native SNAP-25 cleavage site is between Gln197 and Arg198; thus, inhibition by arginine derivatives was not entirely unexpected. However, it was surprising that little difference in inhibition efficiency was observed between the D- and L-isomers of arginine hydroxamic acid.

In a further study, the *in situ* preparation of a library of hydroxamic acids has been described (63). Here, 150 commercially available carboxylic acids were converted to hydroxamic acids in a facile two-step procedure to allow for rapid lead identification. On the basis of the initial screen, 4-chlorocinnamic hydroxamate displayed an IC_{50} of 15 μM and was considered a promising lead for further development. A small series of compounds

were synthesized to explore the structure–activity relationships of this lead. Amazingly, substitutions of the chloro substituent were not tolerated, whereas a 2,4-dichloro-substituted compound was found to be the most potent nonpeptidic inhibitor of BoNT to date ($K_i = 0.30 \pm 0.01 \mu\text{M}$) (Figure 3). Not surprisingly, in light of the hydroxamate zinc-binding functionality, this compound was found to be a competitive inhibitor of BoNT/A. It is important to note that the inhibitors in this study contradict the previously articulated hypothesis that in order to be an effective inhibitor of BoNT, a small molecule must have a length of ≥ 10 carbon atoms (61).

Before any BoNT/A LC inhibitor can advance to animal models of BoNT intoxication, efficacy in cellular models must be shown. We recently reported the first demonstration of a BoNT/A LC inhibitor that shows protection from SNAP-25 cleavage in a cellular model (64). Quite surprisingly, this compound is simply a protected amino acid used in peptide synthesis, Fmoc-D-Cys(Trt)-OH (Figure 3). Kinetic analysis of this compound revealed that it also competitively inhibits BoNT/A ($K_i = 18 \mu\text{M}$). Computational docking studies found that the predicted binding constant for the inhibitor ($10 \mu\text{M}$) was in close agreement with the experimentally determined value. Structurally, this model also revealed that a significant amount of binding energy can be attributed to burying the Fmoc group in a hydrophobic pocket, while the carboxylic acid moiety was positioned in close proximity to several positively charged residues. Interestingly, while competitive inhibition was observed, the docking model predicted no interaction between the active site zinc-binding residues and the inhibitor. In light of the potency of this compound combined with its ready availability, it was next tested in a cellular model of BoNT intoxication. At a concentration of $30 \mu\text{M}$, almost total protection of SNAP-25 cleavage was observed in Neuro-2a cells, whereas complete protection was seen when the compound was added to cells at a concentration of $60 \mu\text{M}$. The discovery of potent LC protease inhibitors of the BoNTs could be a crucial step in rescuing nerve activity after toxin internalization. However, to date, research has yet to advance any nonpeptide molecules from enzymatic assays into the corresponding cellular and animal models.

Therapeutics with an Undefined Mechanism of

Action. Theaflavins are unique active ingredients produced when green tea ferments into black or oolong tea. Nishimura and co-workers (65, 66) have reported how

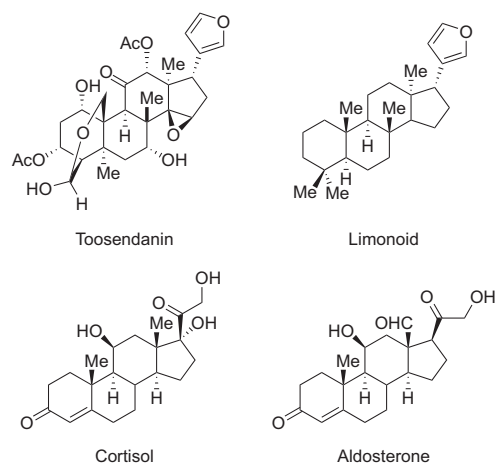


Figure 4. Structure of the limonoid toosendanin. For comparison, the structures of the general limonoid skeleton, as well as representative common steroids, are also shown.

the thearubigin fraction, complex phenols formed during fermentation by polymerization of theaflavins, of *Camellia sinensis* extract protected against the effects of BoNT/A, -B, and -E in mouse phrenic-nerve diaphragm assays. In a series of studies aimed at elucidating the mechanism of action of thearubigin, the binding of ^{125}I -labeled BoNT/A, -B, or -E to rat cerebrocortical synaptosomes was inhibited by these polyphenols. The authors suggest that the thearubigin fraction attenuated the effects of BoNT via simple binding to the toxin. Although which molecule or molecules in the thearubigin fraction are responsible for the loss of toxin activity is unknown, such fractions are made up of a plethora of polyphenols, and multiple compounds are likely involved.

Limonoids are tetranortriterpenoids with a 4,4,8-trimethyl-17-furanylsteroid skeleton derived from euphane or tirucallane triterpenoids, and they generally have an intense, bitter taste (67). Limonoids from the *Melia toosendan* tree have been found to be effective anthelmintics (68); a major limonoid constituent in *M. toosendan* is the compound toosendanin (Figure 4), which appears to have multiple modes of action in insects (69).

Interest in the application of limonoid natural products in pest management remains high. Although this area of research remains fertile, a series of reports detailing the activity of toosendanin on neurotransmitter release in motor-nerve terminals have emerged over the

past 2 decades. In total, these studies have indicated that toosendanin is a selective presynaptic blocker that inhibits quantal release of acetylcholine (70). In contrast with BoNT, the blocking effect of toosendanin has always been preceded by a facilitator phase that depends on the presence of Ca^{2+} in the medium. It is noteworthy that, during the facilitatory period, the tolerance of the neuromuscular junction toward BoNT was enhanced significantly (71).

Further studies investigating the mechanism of toosendanin provided evidence that it causes a decrease in, and ultimately the disappearance of, Ca^{2+} sensitivity, thus, inhibiting the voltage-dependent potassium current in neuroblastoma glioma cells (72). Toosendanin was also shown to induce submicroscopic changes in the neuromuscular junction, namely, a reduction in the number of synaptic vesicles and an increase in the width of synaptic cleft (73). Furthermore, toosendanin inhibits the delayed rectifier potassium channel by intracellular and/or extracellular addition (74).

Toosendanin is membrane-permeable; thus, its modulation of ion channels may be due to its ability to span membrane bilayers and alter channel integrity. The relevance of this finding is that BoNT is capable of forming ion channels in artificial bilayers and in PC12 cell membranes (75). In addition, treatment of synaptosomes with toosendanin imparted resistance to BoNT/A-mediated proteolytic cleavage of SNAP-25. This protective effect did not result from inhibition of the endopeptidase activity; thus, interference must result from disruption of the BoNT LC before it becomes a functional protease (76). Along this line of investigation, BoNT

translocation and channel formation are known to correlate (16). Taken together, the effects of toosendanin on BoNT might be achieved by interfering with the LC translocation. However, the precise mechanism of BoNT antagonism by toosendanin remains unclear.

Conclusion and Outlook. BoNTs are one of the deadliest agents of bioterrorism. Consequently, specific pharmaceutical agents are urgently needed to treat BoNT intoxication. Although an investigational pentavalent toxoid and a horse polyclonal serum are currently available from the CDC and a recombinant vaccine is under development, post-exposure vaccination is virtually useless because of the rapid onset of the toxin. A drug for the prevention or treatment of botulism would be exciting, yet no small molecule has advanced to even phase I clinical trials. Future work in this area will likely be driven by several factors, including the development of robust high-throughput assays for the rapid identification of compound efficacy in cellular models and the complete elucidation of the molecular mechanism of BoNT intoxication. Furthermore, by analyzing natural products such as toosendanin that protect against BoNT-induced paralysis, researchers may also uncover novel molecular scaffolds for BoNT inhibitors. In total, this area of research epitomizes the importance of conducting studies at the interface of chemistry and biology; only by having a firm understanding of the chemical and biological processes at play can an efficient route for the development of new therapeutics be charted.

Acknowledgment: This work has been supported by the Skaggs Institute of Chemical Biology and the National Institutes of Health (AI066507 and BT010-04).

REFERENCES

- van Ermengem, E. (1897) Über einen neuen anaeroben Bacillus und seine Beziehungen zum Botulismus, *Z. Hyg. Infekt.* 26, 1–56.
- Franciosa, G., Aureli, P., and Schechter, R. (2003) *Clostridium botulinum*, *Food Sci. Technol.* 125, 61–90.
- Johnson, E. A., and Bradshaw, M. (2001) *Clostridium botulinum* and its neurotoxins: a metabolic and cellular perspective, *Toxicon* 39, 1703–1722.
- Schantz, E. J., and Johnson, E. A. (1992) Properties and use of botulinum toxin and other microbial neurotoxins in medicine, *Microbiol. Rev.* 56, 80–99.
- Turton, K., Chaddock, J. A., and Acharya, K. R. (2002) Botulinum and tetanus neurotoxins: structure, function and therapeutic utility, *Trends Biochem. Sci.* 27, 552–558.
- Hatheway, C. L., and Johnson, E. A. (1998) Clostridium. the spore-bearing anaerobes, in *Topley and Wilson's Microbiology and Microbial Infections, Vol. 2. Systematic Bacteriology*, (Balows, A., and Duerden, B., Eds.) pp 731–782, Arnold Publishers: London.
- Christopher, G. W., Cieslak, T. J., Pavlin, J. A., and Eitzen, E. M., Jr. (1997) Biological warfare. A historical perspective, *JAMA, J. Am. Med. Assoc.* 278, 412–417.
- Zilinskas, R. A. (1997) Iraq's biological weapons. The past as future? *JAMA, J. Am. Med. Assoc.* 278, 418–424.
- Montecucco, C., and Molgo, J. (2005) Botulinum neurotoxins: revival of an old killer, *Curr. Opin. Pharmacol.* 5, 274–279.
- Johnson, E. A. (1999) Clostridial toxins as therapeutic agents: benefits of nature's most toxic proteins, *Annu. Rev. Microbiol.* 53, 551–575.
- Shukla, H. D., and Sharma, S. K. (2005) *Clostridium botulinum*: a bug with beauty and weapon, *Crit. Rev. Microbiol.* 31, 11–18.
- Munchau, A., and Bhatia, K. P. (2000) Uses of botulinum toxin injection in medicine today, *Br. Med. J.* 320, 161–165.
- Kobayashi, R., Kohda, T., Kataoka, K., Ihara, H., Kozaki, S., Pascual, D. W., Staats, H. F., Kiyono, H., McGhee, J. R., and Fujihashi, K. (2005) A novel neurotoxoid vaccine prevents mucosal botulism, *J. Immunol.* 174, 2190–2195.

14. Oguma, K., Fujinaga, Y., and Inoue, K. (1995) Structure and function of *Clostridium botulinum* toxins, *Microbiol. Immunol.* **39**, 161–168.
15. Schiavo, G., Matteoli, M., and Montecucco, C. (2000) Neurotoxins affecting neuroexocytosis, *Physiol. Rev.* **80**, 717–766.
16. Simpson, L. L. (2004) Identification of the major steps in Botulinum toxin action, *Annu. Rev. Pharmacol. Toxicol.* **44**, 167–193.
17. Dong, M., Yeh, F., Tepp, W. H., Dean, C., Johnson, E. A., Janz, R., and Chapman, E. R. (2006) SV2 is the protein receptor for botulinum neurotoxin A, *Science* **312**, 592–596.
18. Mahrhold, S., Rummel, A., Bigalke, H., Davletov, B., and Binz, T. (2006) The synaptic vesicle protein 2C mediates the uptake of botulinum neurotoxin A into phrenic nerves, *FEBS Lett.* **580**, 2011–2014.
19. Montecucco, C. (1986) How do tetanus and botulinum toxins bind to neuronal membranes? *Trends Biochem. Sci.* **11**, 314–317.
20. Black, J. D., and Dolly, J. O. (1986) Interaction of ¹²⁵I-labeled botulinum neurotoxins with nerve terminals. I. Ultrastructural autoradiographic localization and quantitation of distinct membrane acceptors for types A and B on motor nerves, *J. Cell Biol.* **103**, 521–534.
21. Black, J. D., and Dolly, J. O. (1986) Interaction of ¹²⁵I-labeled botulinum neurotoxins with nerve terminals. II. Autoradiographic evidence for its uptake into motor nerves by acceptor-mediated endocytosis, *J. Cell Biol.* **103**, 535–544.
22. Matthews, G. (1996) Neurotransmitter release, *Annu. Rev. Neurosci.* **19**, 219–233.
23. Betz, W. J., and Angleson, J. K. (1998) The synaptic vesicle cycle, *Annu. Rev. Physiol.* **60**, 347–363.
24. Matteoli, M., Takei, K., Perin, M. S., Sudhof, T. C., and De Camilli, P. (1992) Exo-endocytotic recycling of synaptic vesicles in developing processes of cultured hippocampal neurons, *J. Cell Biol.* **117**, 849–861.
25. Sandvig, K. (2003) Transport of toxins across intracellular membranes, in *Bacterial Protein Toxins*, (Bums, D. L., Barbieri, J. T., and Iglewski, B. H., Eds.) pp 157–172, American Society for Microbiology, Washington, DC.
26. Simpson, L. L. (1982) The interactions between aminoquinoline and presynaptically acting neurotoxins, *J. Pharmacol. Exp. Ther.* **222**, 43–48.
27. Hoch, D. H., Romero-Mira, M., Ehrlich, B. E., Finkelstein, A., DasGupta, B. R., and Simpson, L. L. (1985) Channels formed by botulinum, tetanus, and diphtheria toxins in planar lipid bilayers: relevance to translocation of proteins across membranes, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1692–1696.
28. Koriazova, L. K., and Montal, M. (2003) Translocation of botulinum neurotoxin light chain protease through the heavy chain channel, *Nat. Struct. Biol.* **10**, 13–18.
29. Bajjalieh, S. M. (1999) Synaptic vesicle docking and fusion, *Curr. Opin. Neurobiol.* **9**, 321–328.
30. Humeau, Y., Doussau, F., Grant, N. J., and Poulain, B. (2000) How botulinum and tetanus neurotoxins block neurotransmitter release, *Biochemie* **82**, 427–446.
31. Meunier, F. A., Schiavo, G., and Molgo, J. (2002) Botulinum neurotoxins: from paralysis to recovery of functional neuromuscular transmission, *J. Physiol. (Paris)* **96**, 105–113.
32. Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R., and Montecucco, C. (1992) Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin, *Nature* **359**, 832–835.
33. Amon, S. S., Schechter, R., Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Fine, A. D., Hauer, J., Layton, M., Lillibridge, S., Osterholm, M. T., O'Toole, T., Parker, G., Perl, T. M., Russell, P. K., Swerdlow, D. L., and Tonat, K. (2001) Botulinum toxin as a biological weapon: medical and public health management, *JAMA, J. Am. Med. Assoc.* **285**, 1059–1070.
34. Black, R. E., and Gunn, R. A. (1980) Hypersensitivity reactions associated with botulinum antitoxin, *Am. J. Med.* **69**, 567–570.
35. Amon, S. S., Schechter, R., Maslanka, S. E., Jewell, N. P., and Hatheway, C. L. (2006) Human botulinum immune globulin for the treatment of infant botulism, *N. Engl. J. Med.* **354**, 462–471.
36. Marks, J. D. (2004) Deciphering antibody properties that lead to potent botulinum neurotoxin neutralization, *Mov. Disord.* **19**, S101–S108.
37. Byrne, M. P., and Smith, L. A. (2000) Development of vaccines for the prevention of botulism, *Biochemie* **82**, 955–966.
38. Rudy, B. (1988) Diversity and ubiquity of K channels, *Neuroscience* **25**, 729–749.
39. Adler, M., Scovill, J., Parker, G., Lebeda, F. J., Piotrowski, J., and Deshpande, S. S. (1995) Antagonism of botulinum toxin-induced muscle weakness by 3,4-diaminopyridine in rat phrenic nerve hemidiaphragm preparations, *Toxicon* **33**, 527–537.
40. Adler, M., Macdonald, D. A., Sellin, L. C., and Parker, G. W. (1996) Effect of 3,4-diaminopyridine on rat extensor digitorum longus muscle paralyzed by local injection of botulinum neurotoxin, *Toxicon* **34**, 237–249.
41. Adler, M., Capacio, B., and Deshpande, S. S. (2000) Antagonism of botulinum toxin A-mediated muscle paralysis by 3,4-diaminopyridine delivered via osmotic minipumps, *Toxicon* **38**, 1381–1388.
42. Simpson, L. L., and Rapport, M. M. (1971) The binding of botulinum toxin to membrane lipids: sphingolipids, steroids and fatty acids, *J. Neurochem.* **18**, 1751–1759.
43. Simpson, L. L., and Rapport, M. M. (1971) Ganglioside inactivation of botulinum toxin, *J. Neurochem.* **18**, 1341–1343.
44. Kitamura, M., Iwamori, M., and Nagai, Y. (1980) Interaction between *Clostridium botulinum* neurotoxin and gangliosides, *Biochim. Biophys. Acta* **628**, 328–335.
45. Kamata, Y., Yoshimoto, M., and Kozaki, S. (1997) Interaction between botulinum neurotoxin type A and ganglioside: ganglioside inactivates the neurotoxin and quenches its tryptophan fluorescence, *Toxicon* **35**, 1337–1340.
46. Bakry, N., Kamata, Y., and Simpson, L. L. (1991) Lectins from *Triticum vulgare* and *Andara* universal antagonists of botulinum neurotoxin and tetanus toxin, *J. Pharmacol. Exp. Ther.* **258**, 830–836.
47. Simpson, L. L. (1982) The interaction between aminoquinolines and presynaptically acting neurotoxins, *J. Pharmacol. Exp. Ther.* **222**, 43–48.
48. Simpson, L. L. (1983) Ammonium chloride and methylamine hydrochloride antagonize clostridial neurotoxins, *J. Pharmacol. Exp. Ther.* **225**, 546–552.
49. Simpson, L. L., Coffield, J. A., and Bakry, N. (1994) Inhibition of vacuolar adenosine triphosphatase antagonizes the effects of clostridial neurotoxins but not phospholipase A2 neurotoxins, *J. Pharmacol. Exp. Ther.* **269**, 256–262.
50. Sheridan, R. E. (1996) Protonophore antagonism of botulinum toxin in mouse muscle, *Toxicon* **34**, 849–855.
51. Keller, J. E., Cai, F., and Neale, E. A. (2004) Uptake of botulinum neurotoxin into cultured neurons, *Biochemistry* **43**, 526–532.
52. Deshpande, S. S., Sheridan, R. E., and Adler, M. (1997) Efficacy of certain quinolines as pharmacological antagonists in botulinum neurotoxin poisoning, *Toxicon* **35**, 433–445.
53. Adler, M., Dineterman, R. E., and Wannemacher, R. W. (1997) Protection by the heavy metal chelator *N,N,N',N'*-tetrakis(2-pyridylmethyl)ethylenediamine (TPEN) against the lethal action of botulinum neurotoxin A and B, *Toxicon* **35**, 1089–1100.
54. Sollner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993) SNAP receptors implicated in vesicle targeting and fusion, *Nature* **362**, 318–324.
55. Schmidt, J. J., and Stafford, R. G. (2002) A high-affinity competitive inhibitor of type A botulinum neurotoxin protease activity, *FEBS Lett.* **532**, 423–426.

56. Sukonpan, C., Oost, T., Goodnough, M., Tepp, W., Johnson, E. A., and Rich, D. H. (2004) Synthesis of substrates and inhibitors of botulinum neurotoxin type A metalloprotease, *J. Pept. Res.* **63**, 181–193.
57. Yiadom, K. P. A. B., Muhie, S., and Yang, D. C. H. (2005) Peptide inhibitors of botulinum neurotoxin by mRNA display, *Biochem. Biophys. Res. Commun.* **335**, 1247–1253.
58. Simpson, L. L., Coffield, J. A., and Bakry, N. (1993) Chelation of zinc antagonizes the neuromuscular blocking properties of the seven serotypes of botulinum neurotoxin as well as tetanus toxin, *J. Pharmacol. Exp. Ther.* **267**, 720–727.
59. Sheridan, R. E., and Deshpande, S. S. (1995) Interactions between heavy metal chelators and botulinum neurotoxins at the mouse neuromuscular junction, *Toxicon* **33**, 539–549.
60. Burnett, J. C., Schmidt, J. J., Stafford, R. G., Panchal, R. G., Nguyen, T. L., Hermone, A. R., Vennerstrom, J. L., McGrath, C. F., Lane, D. J., Sausville, E. A., Zaharevitz, D. W., Gussio, R., and Bavarí, S. (2003) Novel small molecule inhibitors of botulinum neurotoxin A metalloprotease activity, *Biochem. Biophys. Res. Commun.* **310**, 84–93.
61. Park, J. G., Sill, P. C., Makiy, E. F., Garcia-Sosa, A. T., Millard, C. B., Schmidt, J. J., and Pang, Y.-P. (2006) Serotype-selective, small-molecule inhibitors of the zinc endopeptidase of botulinum neurotoxin serotype A, *Bioorg. Med. Chem.* **14**, 395–408.
62. Boldt, G. E., Kennedy, J. P., Hixon, M. S., McAllister, L. A., Barbieri, J. T., Tzipori, S., and Janda, K. D. (2006) Synthesis, characterization and development of a high-throughput methodology for the discovery of botulinum neurotoxin A inhibitors, *J. Comb. Chem.*, published online May 19, <http://dx.doi.org/10.1021/cc060010h>.
63. Boldt, G. E., Kennedy, J. P., and Janda, K. D. (2006) Identification of a potent botulinum neurotoxin A protease inhibitor using in situ lead identification chemistry, *Org. Lett.* **8**, 1729–1732.
64. Boldt, G. E., Eubanks, L. M., and Janda, K. D. (2006) Identification of a botulinum neurotoxin A protease inhibitor displaying efficacy in a cellular model, *Chem. Commun.*, published online May 3, <http://dx.doi.org/10.1039/b603099h>.
65. Satoh, E., Ishii, T., Shimizu, Y., Sawamura, S., and Nishimura, M. (2001) Black tea extract, thearubigin fraction, counteract the effects of botulinum neurotoxins in mice, *Br. J. Pharmacol.* **132**, 797–798.
66. Satoh, E., Ishii, T., Shimizu, Y., Sawamura, S.-L., and Nishimura, M. (2002) The mechanism underlying the protective effect of the thearubigin fraction of black tea (*Camellia sinensis*) extract against the neuromuscular blocking action of botulinum neurotoxins, *Pharmacol. Toxicol.* **90**, 199–202.
67. Nakatani, M. (1999) Limonoids from *Melia toosendan* (Meliaceae) and their antifeedant activity, *Heterocycles* **50**, 595–609.
68. Xie, Y. S., Fields, P. G., Isman, M. B., Chen, W. K., and Zhang, X. (1995) Insecticidal activity of *Melia toosendan* extracts and toosendanin against three stored-product insects, *J. Stored Prod. Res.* **31**, 259–265.
69. Zhou, J.-B., Minami, Y., Yagi, F., Tadera, K., and Nakatani, M. (1997) Antifeeding limonoids from *Melia toosendan*, *Heterocycles* **45**, 1781–1786.
70. Shih, Y.-L., Wei, N., Yang, Y., and Wang, Z. (1980) Toosendanina presynaptic neuromuscular blocking agent, *Acta Physiologica Sin.* **32**, 293–297.
71. Shih, Y.-L., and Hsu, K. (1983) Anti-botulism effect of toosendanin and its facilitatory action on miniature end-plate potentials, *Jpn. J. Physiol.* **33**, 677–680.
72. Li, M.-F., Wu, Y., Wang, Z.-F., and Shi, Y.-L. (2004) Toosendanin, a triterpenoid derivative, increases Ca^{2+} current in NG108-15 cells via L-type channels, *Neurosci. Res.* **49**, 197–203.
73. Wang, Z.-F., and Shi, Y.-L. (2001) Inhibition of large-conductance Ca^{2+} -activated K^+ channels in hippocampal neurons by toosendanin, *Neuroscience* **104**, 41–47.
74. Wang, Z.-F., and Shi, Y.-L. (2001) Modulation of inward rectifier potassium channel by toosendanin, a presynaptic blocker, *Neurosci. Res.* **40**, 211–215.
75. Tang, M.-Z., Wang, Z.-F., and Shi, Y.-L. (2003) Toosendanin induces outgrowth of neuronal processes and apoptosis in PC12 cells, *Neurosci. Res.* **45**, 225–231.
76. Wang, Z.-F., Zhou, J.-Y., Ren, X.-M., Tang, M.-Z., and Shi, Y.-L. (2003) Antagonism of botulinum toxin type A-induced cleavage of SNAP-25 in rat cerebral synaptosome by toosendanin, *FEBS Lett.* **555**, 375–379.

Reading Dynamic Kinase Activity in Living Cells for High-Throughput Screening

Michael D. Allen[†], Lisa M. DiPilato[†], Meghdad Rahdar[†], Yunzhao R. Ren[†], Curtis Chong[†], Jun O. Liu^{†,*}, and Jin Zhang^{†,*}

[†]Department of Pharmacology and Molecular Sciences, and ^{*}The Solomon H. Snyder Department of Neuroscience and Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Kinases are crucial nodes in the signaling network and emerging drug targets of the 21st century (1). High-throughput screening with various kinase targets (2, 3) has been an effective approach for the identification of new leads for drug development and pharmacological probes for mechanistic studies. We sought to develop a new method for live-cell, high-throughput screening that is generally applicable to most, if not all, kinase targets within the kinome (Figure 1, panel a). This live-cell-based screening method should be ideally suited for examining dynamic responses of endogenous kinase targets, for evaluating drug candidates that ultimately perform within cellular environments, and for identifying compounds with unique mechanisms of action (3). Furthermore, such a method may be extended to follow multiple components of kinase-mediated signaling pathways to screen for pathway modulators. In this context, we chose to focus on protein kinase A (PKA) (4, 5) and the cyclic adenosine monophosphate (cAMP)/PKA pathway downstream of many G-protein coupled receptors (GPCRs), which represent one of the largest classes of current drug targets (6).

Dynamic tracking of specific kinase activity in living cells can be provided by a class of genetically encoded reporters based on FRET (7, 8), which offers sensitive ratiometric fluorescence readout, precise molecular targeting, and high spatiotemporal resolution

compared to other fluorescent kinase sensors (9). A general modular design for such kinase activity reporters based on a phosphorylation-dependent conformational switch has been successfully applied to many serine/threonine and tyrosine kinases, with many more under development (8). Although these FRET-based reporters are powerful tools for monitoring kinase activities in single-cell imaging, their tremendous value cannot be fully realized without a high-throughput compatible assay platform. To extend the application of these live-cell imaging tools to high-throughput screening, we tested the ratiometric readout of A-kinase activity reporter (AKAR) (10), in a high-throughput plate reader format. This assay format, when combined with the use of kinase activity reporters, is expected to allow simple, fast, and convenient high-throughput reading of dynamic kinase activities with high spatiotemporal resolution. Promising results were obtained using a recently reported version of the AKAR reporter, AKAR2 (11) (not shown), although the Z' factor (12), which compares the assay dynamic range to data variation, is <0.5, making this particular reporter unsuitable for high-throughput screening.

We rationalized that further improvement of sensor dynamic range would be necessary for application of AKAR in high-throughput format, as such assays are often affected by cell–cell heterogeneity and loss of sensitivity for averaged signals. By opti-

ABSTRACT Protein kinases, as crucial signaling molecules, represent an emerging class of drug targets, and the ability to assay their activities in living cells with high-throughput screening should provide exciting opportunities for drug discovery and chemical and functional genomics. Here, we describe a general method for high-throughput reading of dynamic kinase activities using ratiometric fluorescent sensors, and showcase an example of reading intracellular activities of protein kinase A (PKA) and the cyclic adenosine monophosphate (cAMP)/PKA pathway downstream of many G-protein coupled receptors (GPCRs). We further demonstrate the first compound screen based on the ability of compounds to modulate dynamic kinase activities in living cells and show that such screening of a collection of clinical compounds has successfully identified modulators of the GPCR/cAMP/PKA pathway.

*Corresponding author,
jzhang32@jhmi.edu.

Received for review May 15, 2006
and accepted June 28, 2006.

Published online July 21, 2006
10.1021/cb600202f CCC: \$33.50

© 2006 by American Chemical Society

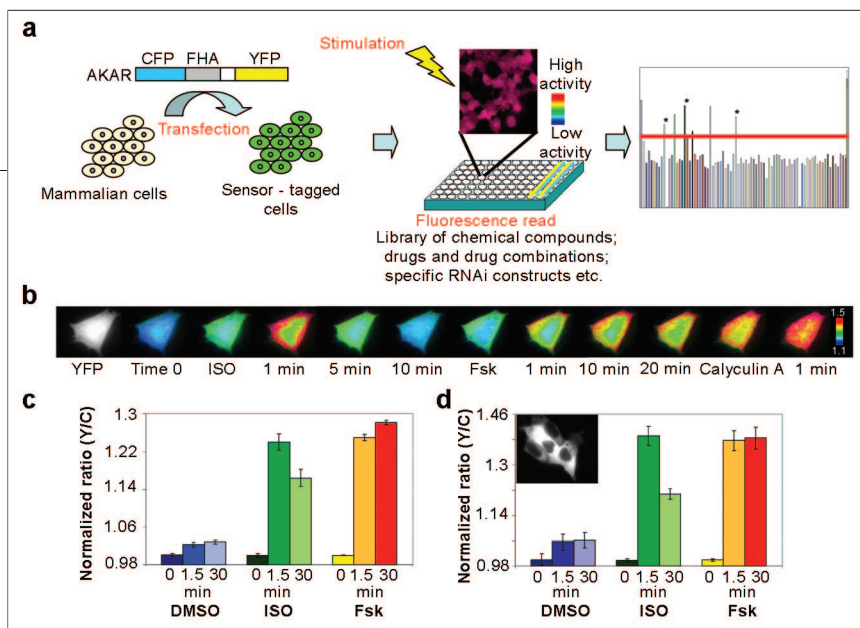


Figure 1. A high-throughput activity assay based on the improved AKAR. **a)** Scheme of using live-cell, high-throughput assays based on FRET reporters (e.g., AKAR) for drug screens, pharmacological profiling, and functional genomics studies. **b)** A representative time course of emission ratio (yellow/cyan) change of AKAR3, indicated by pseudocolor images. A HEK-293 cell expressing AKAR3 was stimulated with 1 μ M ISO, followed by 50 μ M Fsk, and Calyculin A, a phosphatase inhibitor. **c)** Representative data from five independent runs in 96-well format showing emission ratio (yellow/cyan) changes of AKAR3 in HEK-293 cells treated with ISO, Fsk, and 2% DMSO. **d)** Representative data from three independent runs showing emission ratio (yellow/cyan) changes of AKAR3-NES in HEK-293 cells treated with indicated drugs. The inset shows a representative fluorescence image of cells expressing AKAR-NES. Error bars represent standard deviation ($n = 3$).

mizing the change in distance (11) and orientation (13) between donor-acceptor fluorophores (M. Allen, J. Zhang, Johns Hopkins, unpublished data), we generated an improved version that doubled the response amplitude of AKAR2 (Supplementary Figure 1). This reporter, AKAR3, was tested in a 96-well plate format. As in single-cell imaging (Figure 1, panel b), addition of β -adrenergic agonist isoproterenol (ISO) to HEK-293 cells expressing AKAR3 in 96-well plates induced an increase in yellow-to-cyan emission ratio of $22.1\% \pm 0.7\%$ ($n = 3$), followed by a slight decrease (Figure 1, panel c). When cells were treated with adenylyl cyclase activator forskolin (Fsk), sustained responses were observed with an average emission ratio change of $25.7\% \pm 0.7\%$ ($n = 3$). As negative controls, addition of buffer or 2% dimethyl sulfoxide (DMSO) generated minimal changes in emission ratios. This 96-well plate assay was sensitive and reproducible, with an S/N ratio of 30.3, a Z' factor of 0.84, and a coefficient of variation (CV) of 1.8% for the Fsk-stimulated response.

CV measures the variability around the mean in relation to the size of the mean and, along with the Z' factor, provides statistical evaluation of the assay.

Kinase activities are often spatially compartmentalized (14), and compounds targeting spatiotemporal regulation of kinases could be new classes of modulators. To provide specific readout of compartmentalized PKA activities, we took advantage of the genetic targetability of such activity reporters and introduced a C-terminal nuclear export signal (NES) to AKAR3. In the plate reader assay using AKAR3-NES (Figure 1, panel d), cytosolic PKA activity was recorded without contamination of nuclear activity, which has slower kinetics (10) due to diffusional translocation of the catalytic subunit from cytosol to nucleus. As a result, ISO stimulated a larger increase in emission ratio followed by a more rapid decrease than that in the AKAR3 assay (Figure 1, panel d). The maximum signal for the assay also improved, showing a ratio change of $37.8\% \pm 3.2\%$ ($n = 3$). Thus, sensor targeting has

led to increased temporal and spatial resolution. This technique could be used to reveal how individual signaling microdomains, such as kinase-containing signaling complexes (14), are affected by drugs or other perturbations. This adds another level of capacity to these nonimage-based, high-throughput assays.

To follow the activity of the signaling pathway and to assay related drug targets in parallel, we further sought to develop a live-cell-based, high-throughput assay for cAMP, a second messenger that is generated *via* activation of Gs-coupled GPCRs and exerts its effects by activating PKA and other effectors (4). Similar approaches for improvement were applied to create a reporter that more than doubled the response of previous Indicator for cAMP Using Epac (ICUE) (15), which generates a decrease in FRET upon binding of cAMP (Supplementary Figure 2). When tested in plate reader format, this improved version ICUE3 showed consistent responses to ISO, which were inhibited by cotreatment with β -adrenergic antagonist propranolol (Pro) (Figure 2, panel a). A maximum response of $24.1\% \pm 2.7\%$ was obtained with Fsk stimulation in the presence of 100 μ M 3-isobutyl-1-methylxanthine (IBMX), a phosphodiesterase (PDE) inhibitor (Figure 2, panel b). This ICUE assay leaves room for improvement with a Z' factor of 0.51, an S/N ratio of 17.8, and a CV of 12%. Possible sources of the variation and reduced signal are transfection efficiency and expression variability. Indeed, experiments using a stable cell line showed significant improvement of the assay, indicated by a maximum response of $43.3\% \pm 2.3\%$ ($n = 3$) and a Z' factor of 0.78. Thus, a live-cell, high-throughput cAMP assay was developed as a functional assay for Gs-coupled GPCRs (6), demonstrating both the assay platform and sensor improvement strategy could be generally applicable.

The development of activity assays for the cAMP/PKA pathway allowed for parallel kinetic profiling of a panel of agonists and

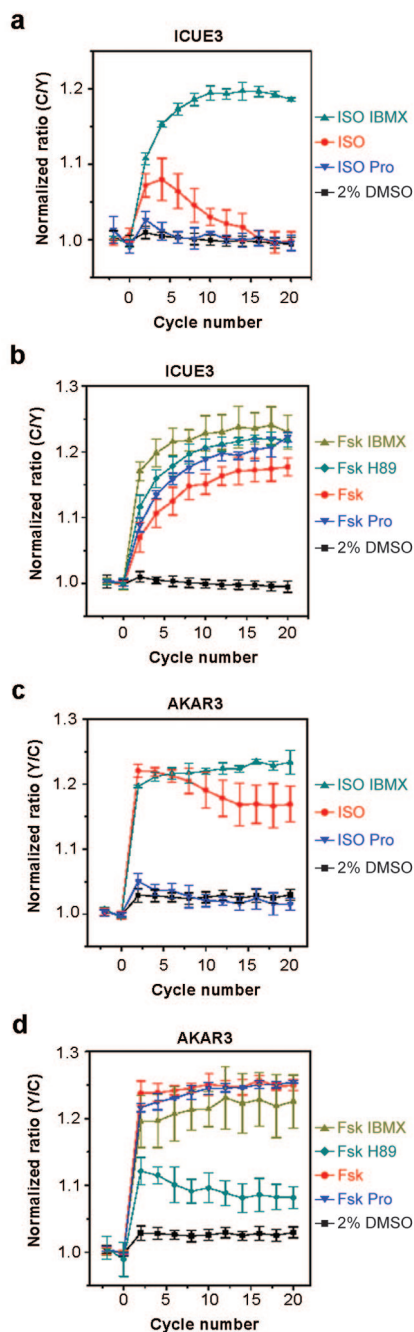


Figure 2. High-throughput activity assays for cAMP and PKA in pharmacological profiling. a) and b) Representative data from five independent runs showing emission ratio (cyan/yellow) changes of ICUE3 in HEK-293 cells treated with indicated drugs. 1 cycle = 92 s. c and d) Representative data from five independent runs showing emission ratio (yellow/cyan) changes of AKAR3 in HEK-293 cells treated with indicated drugs, ISO, Fsk, IBMX, Pro. 1 cycle = 64 s. Error bars represent standard deviation ($n = 3$).

IBMX sustained responses and increased their amplitude, indicating PDEs play an important role in switching off cAMP/PKA signaling in this cell system. Addition of H89, a PKA inhibitor, diminished the Fsk-stimulated response of AKAR3 but not that of ICUE3 (Figure 2, panels b and d). Interestingly, H89 increased the ICUE3 response, similar to the effect caused by combination of Fsk and IBMX (Figure 2, panel b). Confirmed by single-cell imaging (not shown), this finding suggests that disruption of a feedback loop possibly involving PKA-dependent activation of PDE and/or inhibition of adenylyl cyclase (16) could directly contribute to enhanced cAMP accumulation. Thus, use of these assays allows parallel evaluation of key targets in the same signaling pathway and should facilitate understanding of complex drug effects.

Upon development of these live-cell, high-throughput activity assays, a new collection of Food and Drug Administration-approved drugs and other clinically relevant compounds, the Johns Hopkins Clinical Compound Library (17), became available, which lent itself to the first compound screen based on the ability of compounds to modulate dynamic kinase activities in living cells. Such screening could identify clinical compounds that activate or inhibit PKA by affecting GPCR/cAMP/PKA signaling pathway, which may provide new insight into their clinical effects and point to new opportunities of developing more effective drugs that capitalize on their abilities to modify the activity of this pathway.

antagonists, revealing their differential effects on PKA activity and cAMP dynamics. ISO-induced responses can be inhibited by 10 μ M Pro (Figure 2), which had no effect on Fsk-stimulated responses. Supplementing

In the primary screening, 160 compounds, a subset of the library (see Supporting Information), were added to AKAR3-expressing HEK-293 cells in individual wells of 96-well plates to a final concentration of 10 μ M after a base line reading. Time courses were recorded to monitor any fluorescence changes upon drug addition and to identify hits that activate PKA as potential agonists. A standard agonist, for example ISO, was then added to all wells, and changes in emission ratios were calculated to identify compounds that inhibit such changes as potential antagonists. Screening using ICUE stable cells was performed in 96-well plates in a similar fashion (see Supporting Information). Of note, both agonists and antagonists could be identified in one primary screening, and some kinetic information can be obtained that facilitates early and direct characterization of hits.

Most drugs caused no stimulation or inhibition of AKAR responses (Figure 3). In some cases, abnormal fluorescence changes were observed upon addition of compounds to the cells, which correlated with their fluorescent or colorimetric properties or toxicity. Potential agonists (Figure 3, panel a) were identified and individually evaluated. Three compounds that stimulated AKAR responses on their own without generating abnormal fluorescence changes in individual channels were identified as ISO, ritaline, and epinephrine. Unsurprisingly, ISO, a general β -adrenergic agonist, was identified as an agonist from the library. Also a well-known adrenergic agonist, epinephrine activates both β -adrenergic receptors (β -AR) and α -adrenergic receptors. In β 2AR-expressing HEK-293 cells, 10 μ M epinephrine stimulates cAMP production and PKA activation, as confirmed by single-cell imaging experiments (Figure 3, panel c). On the other hand, β 2AR-specific agonist ritaline stimulated a moderate and gradual response in HEK-293 cells expressing AKAR3, when tested in single-cell imaging (Figure 3, panel d). This small response

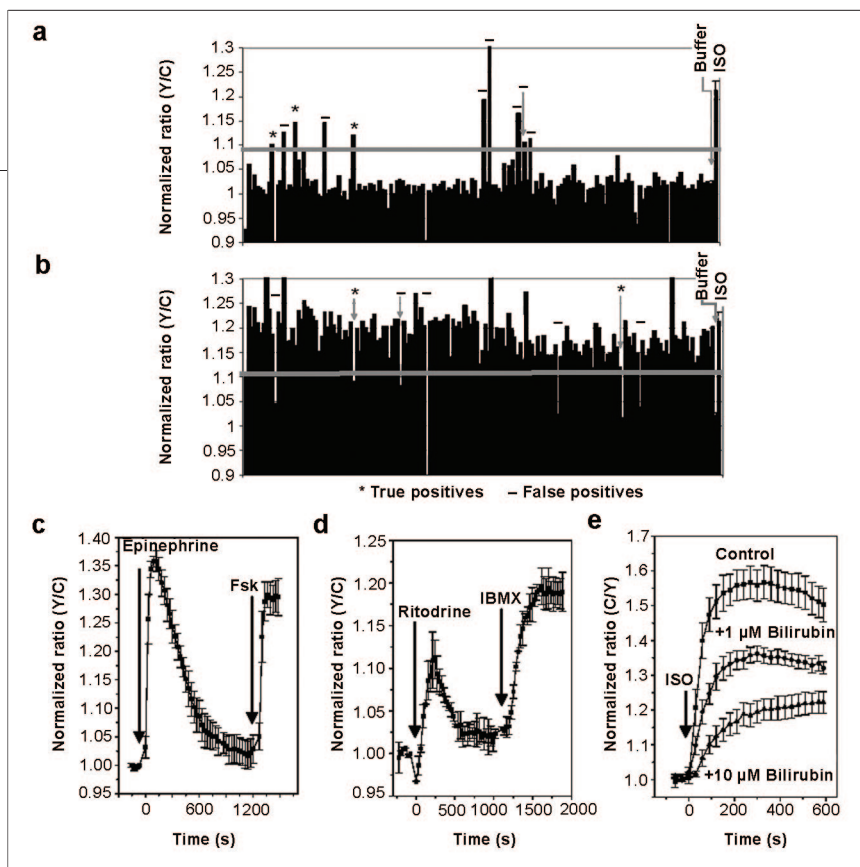


Figure 3. Clinical compound screen. **a**) Normalized emission ratio (yellow/cyan) from cells expressing AKAR3 treated with individual library compounds, compared to the negative control in which only buffer was added and the positive control in which ISO (250 nM) was added (agonist screen). **b**) Normalized emission ratio (yellow/cyan) from cells expressing AKAR3 first treated with individual library compounds for ~15 min then stimulated by ISO (250 nM) (antagonist screen). The positive control in which only ISO was added and the negative control in which only buffer was added are shown as labeled. The asterisks indicate responses of positive hits, minus signs label false positives, including fluorescent or colored compounds, cytotoxic, or unconfirmed hits by the secondary assay. **c**) Representative time courses of emission ratio (yellow/cyan) change of cells expressing AKAR3 treated with 10 μM epinephrine followed by 50 μM Fsk in single-cell imaging experiments. Error bars represent standard deviation ($n = 3$). **d**) Representative time courses of emission ratio (yellow/cyan) change of cells expressing AKAR3 treated with 50 μM ritodrine followed by 100 μM IBMX in single-cell imaging experiments. Error bars represent standard deviation ($n = 2$). **e**) Representative time courses of emission ratio (cyan/yellow) change of cells expressing ICUE pretreated with bilirubin (1 μM and 10 μM) for 15 min, followed by ISO stimulation in presence of IBMX. Error bars represent standard deviation ($n = 22, 9, \text{ and } 10$).

could be enhanced by blocking PDE activity with IBMX (Figure 3, panel d), while IBMX alone did not generate any such responses in HEK-293 cells (not shown). Importantly all three known agonists among the 160 drugs were identified as positive hits in the screen.

One compound potentially inhibited the ISO-stimulated responses in both AKAR and ICUE assays and was identified to be Pro, the general β -adrenergic antagonist previously used in our profiling experiment. Interestingly, epinephrine was also identified as a hit in the antagonist category, indicating

that, after inducing a transient response (Figure 3, panel c), continuous presence of epinephrine decreased the cell response to subsequent stimulation by ISO. An additional potential antagonist was identified from the ICUE assay, and the inhibitory effects by this compound, bilirubin, were verified in single cell imaging experiments. The ISO-stimulated ICUE response was inhibited by 40% by 1 μM bilirubin (Figure 3, panel e), and the inhibition was more than 60% by 10 μM bilirubin. Although bilirubin has been shown to be a protective antioxi-

dant, very high levels can lead to its accumulation in the brain, causing kernicterus (18). The ability of bilirubin to inhibit PKA *in vitro* at tens of micromolar concentration was suggested to play a role in the neurotoxic effects of bilirubin in patients of kernicterus (19). Our results suggest an alternative mechanism of inhibition of PKA by bilirubin, that is, through inhibition of cellular cAMP production. It is of interest to note the ICUE assay may be more sensitive for identifying potential antagonists, while the AKAR assay is more sensitive for agonists, as an amplification step is incorporated in the AKAR assay when a single active PKA molecule phosphorylates multiple AKAR substrates.

In summary, we have developed a new method for live-cell, high-throughput screening based on measurement of dynamic activities of signaling molecules, particularly protein kinases. Furthermore, a compound screen based on the ability of compounds to modulate dynamic kinase activities in live cells was performed using a collection of clinical compounds, from which known and new modulators of the GPCR/cAMP/PKA signaling pathway were identified. This method captures many of the crucial elements of high-throughput assays. However, further optimization, miniaturization, and automation would be necessary to achieve higher throughput.

This new screening method should complement, yet offer unique advantages over, existing methods including purified target-based biochemical screens and endpoint focused phenotypic screens. Such activity-based screens, which can be combined with phenotypic screens (20), should allow direct measurement of dynamic cellular activities of defined targets, or the activity of a signaling pathway when used concurrently. Compared to *in vitro* assays (2), living cells are used as reaction vessels with targets of interest, cofactors, and regulators present at endogenous levels in their natural cellular environment, where spatio-

temporal control of signaling activities can be specifically followed. With the complexity of live systems maintained, the quality of the screening process should be increased, enabling discovery of compounds with unique mechanisms of action (3).

It is our expectation that these simple yet powerful activity assays based on FRET reporters should find immediate application in high-throughput screens for pharmacological reagents and drug candidates, as well as in parallel tracking of multiple physiological and pharmacological events at subcellular locations in living cells in chemical and functional genomics studies. Furthermore, this assay platform should be generally applicable to most kinases in the kinome, as various kinase activity sensors could be engineered using the general design, improved in a similar fashion and adapted to this assay format. The successful bridging of high-throughput technology with dynamic live-cell activity measurement has thus laid a foundation for high-capacity mechanistic studies and multifaceted drug discovery processes targeting protein kinases.

METHODS

Gene Construction. Different variants of fluorescent protein were PCR-amplified and incorporated into AKAR2 and AKAR2 T391A to replace the original enhanced cyan fluorescent protein or Citrine (11). Cytoplasmic targeting of AKAR3 was achieved by genetically adding a nuclear export signal (NES) LPPLERLTL at the C-terminal end (21). ICUE constructs containing new fluorescent protein variants were created using the same method as above. All constructs were initially generated in pRSET B (Invitrogen), then subcloned into pcDNA3 (Invitrogen) after a Kozak sequence for mammalian expression, except for AKAR3 NES, which utilized 3' restriction sites in pcDNA3 to introduce the NES.

Cell Culture and Imaging. HEK-293 cells were plated onto sterilized glass coverslips in 35-mm dishes and grown to ~50% confluency in DMEM (10% fetal bovine serum (v/v) at 37 °C with 5% CO₂). Cells were transfected with calcium phosphate and allowed to grow for 12–24 h before imaging. After a wash with Hanks' balanced salt solution (HBSS), cells were maintained in buffer in the dark at 20–25 °C and treated with various reagents as indicated. Cells were imaged on a Zeiss Axiovert 200M microscope with a 40×/1.3NA oil-immersion objective lens and

cooled CCD camera as previously described (22). Briefly, dual emission ratio imaging used a 420DF20 excitation filter, a 450DRLP dichroic mirror, and two emission filters (475DF40 for cyan and 535DF25 for yellow). The ratios of yellow-to-cyan (AKAR) or cyan-to-yellow (ICUE) emissions were then calculated at different time points and normalized by dividing all ratios by the emission ratio before stimulation, setting the basal emission ratio as 1.

Live Cell Plate Reading. HEK-293 cells were transfected with AKAR3 or ICUE3 using calcium phosphate at 40% confluency and allowed to grow for 40 h. Cells were then trypsinized and plated in a Costar 3603 96-well plate (Corning) at a density of 150,000 cells per well. After incubation for another 24 h, cells were washed once with HBSS and left in 150 μL of HBSS at 20–25 °C. Fluorescence reading was taken on FLUOstar OPTIMA fluorescence microplate reader (BMG Labtechologies, Inc.) using a 420DF20 excitation filter and two emission filters (470DF40 for cyan and 535DF25 for yellow). A baseline was established in three cycles, each consisting of a full plate reading of yellow intensity, followed by a reading of cyan intensity. Each cycle lasted between 64 and 92 s. Cells were then treated with ISO, Fsk, H89, Pro (Sigma), and IBMX (Sigma) as indicated. Readings were taken in additional cycles. FRET change was calculated as the percent increase of emission ratios (yellow-to-cyan for AKAR and cyan-to-yellow for ICUE) over baseline for each well during a given cycle. To assess the efficacy of our assay, several statistical parameters were calculated including Z' factor, CV, and S/N ratio (12). The Z' factor reflects both dynamic range and variation with the following equation in which σ^+ and σ^- are the standard deviations of the positive and negative control samples, respectively, and μ^+ and μ^- are the average responses for the positive and negative controls, respectively.

$$Z' = 1 - [(3\sigma^+ + 3\sigma^-)/(\mu^+ - \mu^-)]$$

CV is defined as the standard deviation of a set of responses divided by the mean of the set. Finally, signal-to-noise ratio was calculated for the assay as the difference between the mean positive control signal and mean negative control signal divided by the standard deviation of the negative control responses.

Live Cell Clinical Compound Screen. HEK-293 cells transiently expressing AKAR or stably expressing ICUE were trypsinized and plated in a Costar 3603 96-well plate (Corning) at a density of 150,000 cells per well. After incubation for 24 h, cells were washed once with HBSS and left in 190 μL of HBSS at 20–25 °C. Fluorescence readings were taken as described above with each cycle lasting 90 s. Following baseline acquisition, cells in each experimental well were treated with a compound from the Johns Hopkins Clinical Compound Library to a final concentration of 10 μM. Control cells were treated with 10 μL of 10% FBS in a solution of phosphate-buffered saline (PBS) at pH 7.4, the solution used to dissolve library compounds. Readings were taken for 10 cycles spanning a time

of ~15 min, after which cells in experimental wells and positive controls were treated with 0.25 μM ISO (AKAR) or 0.25 μM ISO plus 100 μM IBMX (ICUE), while negative control cells received 0.5% (v/v) DMSO in HBSS. Ten final cycles were then performed. FRET responses were calculated as described above. Negative control (10% FBS, 0.5% DMSO) and positive control (10% FBS, 0.25 μM ISO) curves were generated. Agonist hits were defined as compounds eliciting responses that are larger than 6 times of standard deviation above the baseline or 50% of the positive control. Antagonist hits were defined as compounds that decreased the ISO-stimulated response by 50% or by 6 times of the standard deviation. Hits identified based on only one extreme outlier data point in the entire time course were considered false positives. From the plate-reading data, individual channels (yellow and cyan) of all hits were also examined for abnormal changes upon addition of library compounds. An example of detection of a compound with inherent yellow fluorescence is an increase in yellow emission upon excitation with little or no change in cyan emission. Fluorescence or colorimetric properties of compounds were further verified by comparing to literature data or through direct fluorescence measurement. Such hits were labeled as false positives.

Fluorescence Microscopy-Based Secondary Screen. Fluorescence microscopy-based secondary screen was used to examine the remaining hits. HEK-293 cells expressing AKAR or ICUE were followed via fluorescence microscopy before and after treatment with these compounds at a final concentration of 10 μM. Compounds that caused cell rounding, blebbing, lift-up, or severe shrinkage were considered cytotoxic and labeled as false positives. Compounds that did not affect emission ratio dynamics either before or after treatment with ISO (agonist and antagonist secondary screen, respectively) were also labeled as false positives. Those that do antagonize or agonize in both primary and secondary screens were labeled as true positives.

Acknowledgment: We thank A. Miyawaki, R. Tsien, and A. Hires for providing various Venus constructs; B. Curtin and J. Violin for technical help; and M. Li and Q. Ni for helpful discussion. This work was supported by an award from the American Heart Association and the National Institutes of Health (R01 DK073368).

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

1. Cohen, P. (2002) Protein kinases—the major drug targets of the twenty-first century? *Nat. Rev. Drug Discovery* 1, 309–315.
2. von Ahnen, O., and Bomer, U. (2005) High-throughput screening for kinase inhibitors, *Chem-BioChem* 6, 481–490.
3. Zaman, G. J., Garritsen, A., de Boer, T., and van Boeckel, C. A. (2003) Fluorescence assays for high-throughput screening of protein kinases, *Comb. Chem. High Throughput Screening* 6, 313–320.

4. Tasken, K., and Aandahl, E. M. (2004) Localized effects of cAMP mediated by distinct routes of protein kinase A, *Physiol. Rev.* **84**, 137–167.
5. Taylor, S. S., Yang, J., Wu, J., Haste, N. M., Radzio-Andzelm, E., and Anand, G. (2004) PKA: a portrait of protein kinase dynamics, *Biochim. Biophys. Acta* **1697**, 259–269.
6. Williams, C. (2004) cAMP detection methods in HTS: selecting the best from the rest, *Nat. Rev. Drug Discovery* **3**, 125–135.
7. Miyawaki, A. (2003) Visualization of the spatial and temporal dynamics of intracellular signaling, *Dev. Cell* **4**, 295–305.
8. Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002) Creating new fluorescent probes for cell biology, *Nat. Rev. Mol. Cell Biol.* **3**, 906–918.
9. Rothman, D. M., Shultz, M. D., and Imperiali, B. (2005) Chemical approaches for investigating phosphorylation in signal transduction networks, *Trends Cell Biol.* **15**, 502–510.
10. Zhang, J., Ma, Y., Taylor, S. S., and Tsien, R. Y. (2001) Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14997–15002.
11. Zhang, J., Hupfeld, C. J., Taylor, S. S., Olefsky, J. M., and Tsien, R. Y. (2005) Insulin disrupts beta-adrenergic signalling to protein kinase A in adipocytes, *Nature* **437**, 569–573.
12. Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays, *J. Biomol. Screening* **4**, 67–73.
13. Nagai, T., Yamada, S., Tominaga, T., Ichikawa, M., and Miyawaki, A. (2004) Expanded dynamic range of fluorescent indicators for Ca²⁺ by circularly permuted yellow fluorescent proteins, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10554–10559.
14. Wong, W., and Scott, J. D. (2004) AKAP signalling complexes: focal points in space and time, *Nat. Rev. Mol. Cell Biol.* **5**, 959–970.
15. DiPilato, L. M., Cheng, X., and Zhang, J. (2004) Fluorescent indicators of cAMP and Epac activation reveal differential dynamics of cAMP signaling within discrete subcellular compartments, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16513–16518.
16. Murthy, K. S., Zhou, H., and Makhlof, G. M. (2002) PKA-dependent activation of PDE3A and PDE4 and inhibition of adenylyl cyclase V/VI in smooth muscle, *Am. J. Physiol. Cell Physiol.* **282**, C508–C517.
17. Chong, C. R., Qian, D. Z., Pan, F., Wei, Y., Pili, R., Sulivan, D. J., Jr., and Liu, J. O. (2006) Identification of type 1 inosine monophosphate dehydrogenase as an antiangiogenic drug target, *J. Med. Chem.* **49**, 2677–2680.
18. Sedlak, T. W., and Snyder, S. H. (2004) Bilirubin benefits: cellular protection by a biliverdin reductase antioxidant cycle, *Pediatrics* **113**, 1776–1782.
19. Hansen, T. W., Mathiesen, S. B., and Walaas, S. I. (1996) Bilirubin has widespread inhibitory effects on protein phosphorylation, *Pediatr. Res.* **39**, 1072–1077.
20. Clemons, P. A. (2004) Complex phenotypic assays in high-throughput screening, *Curr. Opin. Chem. Biol.* **8**, 334–338.
21. Ullman, K. S., Powers, M. A., and Forbes D. J. (1997) Nuclear export receptors: from importin to exportin, *Cell* **90**, 967–970.
22. Ananthanarayanan, B., Ni, Q., and Zhang, J. (2005) Signal propagation from membrane messengers to nuclear effectors revealed by reporters of phosphoinositide dynamics and Akt activity, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15081–15086.

Photoactive Analogues of the Haloether Anesthetics Provide High-Resolution Features from Low-Affinity Interactions

Jin Xi[†], Renyu Liu[†], Matthew J. Rossi[‡], Jay Yang[§], Patrick J. Loll[‡], William P. Dailey[†], and Roderic G. Eckenhoff^{†,*}

[†]Department of Anesthesiology & Critical Care, and [‡]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, [‡]Department of Biochemistry & Molecular Biology, Drexel University College of Medicine, Philadelphia, Pennsylvania 19102 and [§]Department of Anesthesiology, Columbia University, New York, New York

The molecular targets for the volatile inhaled anesthetic targets remain uncertain, both in identity and number. Site-directed mutagenesis and electrophysiology has allowed insight into anesthetic binding sites in some putative targets (1, 2), while photoaffinity labeling with halothane (1-bromo-1-chloro-2,2,2-trifluoroethane), an extensively used inhalational anesthetic, has allowed selection of candidate targets from complex mixtures (3, 4). Photolabeling has also provided for residue-level localization in the nicotinic acetylcholine receptor (5, 6) and the GPCR rhodopsin (7), together with estimates of stoichiometry. Halothane decomposes to bromine atom and chlorotrifluoroethyl radical on exposure to short UV light, which then adduct into nearby regions of proteins. Halothane photolabeling has been a powerful approach to discover both sites and targets, but is limited by the requirement for 250 nm illumination, photoselectivity for adduction sites, the creation of two reactive moieties, and the extreme expense of the radiolabeled product.

We have previously reported on a diazine-based analogue of halothane, finding that it had anesthetic activity and photoincorporated into specific proteins (8). Attempts to radiolabel this molecule through a base-catalyzed exchange were not successful, and using radiolabeled precursors was not considered cost-effective due to the low efficiency of the reactions. Further, like halothane, this diazine compound is a haloalkane anesthetic, and haloalkanes are no longer used in the United States because of their toxicity profiles. Haloethers represent a much more relevant class of compounds, since all current clinical anesthetics fall into this class, including the three most commonly used

ABSTRACT The difficulty in obtaining binding target and site information for low-affinity drugs, like the inhaled anesthetics, has limited identification of their molecular effectors. Because such information can be provided by photoactive analogues, we designed, synthesized, and characterized a novel diazirnyl haloether that closely mimics isoflurane, the most widely used clinical general anesthetic. This compound, H-diaziflurane, is a nontoxic, potent anesthetic that potentiates GABA-gated ion channels in primary cultures of hippocampal neurons. Calorimetric and structural characterizations show that H-diaziflurane binds a model anesthetic host protein with similar energetics as isoflurane and forms photoadducts with residues lining the isoflurane binding site. H-diaziflurane will be immediately useful for identifying targets and sites important for the molecular pharmacology of the inhaled haloether anesthetics.

*Corresponding author,
roderic.eckenhoff@uphs.upenn.edu.

Received for review May 17, 2006
and accepted June 23, 2006.

Published online July 14, 2006
10.1021/cb00207d CCC: \$33.50

© 2006 by American Chemical Society

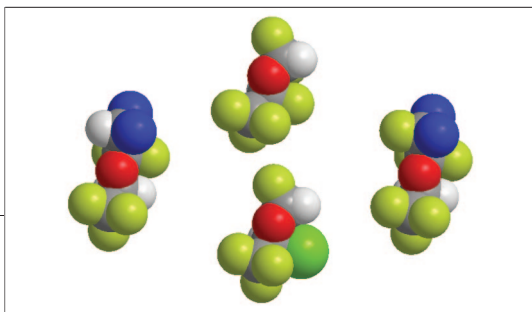


Figure 1. Space-filling representation of H-diaziflurane (left), F-diaziflurane (right), desflurane (top middle) and isoflurane (bottom middle).

isoflurane (1-chloro-2,2,2-trifluoroethyl difluoroethyl ether), and desflurane (1-fluoro-2,2,2-trifluoroethyl difluoroethyl ether). Growing evidence argues against the unitary nature of anesthetic mechanisms, so it appears likely that different targets and binding sites transduce anesthesia from the haloalkanes as compared to the ethers (9, 10). Thus, photoactive reagents for the ethers are essential to complement those presently available for the haloalkanes (8), alcohols (11), or injectable drugs (12, 13). Accordingly, we sought to develop and test a photoactivatable compound to mimic haloether anesthetic activity. We have produced such a compound and show that it binds in a crystallographically identified isoflurane binding site in a naturally occurring protein.

RESULTS AND DISCUSSION

Two candidate diazifluranes were designed and synthesized (Figure 1). Both are colorless volatile liquids, with similar physicochemical parameters as isoflurane (Table 1). Spectroscopic data are given in the experimental section. Both F- and H-diaziflurane were found to produce reversible immobilization of tadpoles (Figure 2). Even when exposed to the highest concentration achievable in pond water for both compounds, all animals were found to recover completely with no detectable untoward effects up to 24 h later.

The F-diaziflurane compound was a more potent immobilizing compound (Figure 2), and exhibits a higher affinity for apoferritin (Table 2; Figure 3) than H-diaziflurane, but mass spectroscopic analysis of UV-exposed apoferritin/F-diaziflurane mixtures revealed no evidence for photoadduction, so we did not pursue characterization of this molecule further. The lack of

agents: sevoflurane (fluoromethyl 2,2,2-trifluoro-1-(trifluoromethyl)ethyl ether),

photoadduction is probably due to excessive stabilization of the carbene by the attached fluorine, with resultant internal rearrangement being favored over adduction. Substitution of hydrogen at this position allowed protein adducts to be detected by MS. Two tryptic fragments of apoferritin L-chain were found to have a +162 Da mass only in the UV plus H-diaziflurane sample (Figure 4). These fragments were MS-MS-sequenced to detect adducts at the residue level. The 1052 Da fragment (890 + 162) starts at L19 and ends at R25. Sequencing in both directions places the adducted residue at L24 (see Supplementary Figure 1). The 1864 Da fragment (1702 + 162) starts at A26 and ends at R39. Sequencing in both directions places the adducted residue at S27 (see Supplementary Figure 2). Both of these residues line the previously identified (14) isoflurane binding cavity in the ferritin L-chain.

Interestingly, the single change of atom on the diaziryl carbon lowered the immobilizing potency by almost 10-fold (Figure 2) and the apoferritin binding affinity by almost 20-fold (Figure 3). The potential basis for this effect is twofold. First, the fluorine makes the F-diaziflurane molecule more hydrophobic (Table 1), and there is a well-known correlation between hydrophobicity and anesthetic potency (Overton/Meyer rule). At the same time, the fluorine increases the (calculated) molecular

TABLE 1. Photolabel properties

Compound	MW	Adduct MW	$\Sigma 320^a$	Water sol.	Octanol/Water	Dipole ^b
F-diaziflurane	208	180	167	0.5 mM	>800	1.84 D
H-diaziflurane	190	162	110	2.0 mM	400	0.90 D
Isoflurane	185	N/A ^c	0	20 mM	120	2.0 D

^a $\Sigma 320$ = extinction coefficient at 320 nm. ^bDipole moments were calculated using *ab initio* molecular orbital theory at the geometry-optimized B3LYP/6-31G* level of theory. ^cN/A, not applicable.

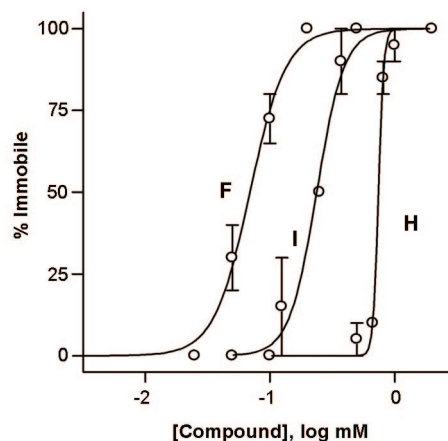


Figure 2. Immobilizing potency of F- (F), H-diaziflurane (H), and isoflurane (I) in *Xenopus* tadpoles. Each point is the mean of three groups of 10 animals. Curves are least squares regressions to variable slope Hill relationships. EC_{50} values were F-diaziflurane, 0.07 mM, 95% CI = 0.06–0.08 mM; isoflurane, 0.24 mM, 95% CI = 0.20–0.28 mM; H-diaziflurane, 0.75 mM, 95% CI = 0.73–0.76 mM.

TABLE 2. Photolabel/apoferritin binding parameters from ITC^a

Compound	<i>n</i>	<i>K_i</i> , M ⁻¹	ΔH cal mol ⁻¹	ΔS cal mol ⁻¹ °K ⁻¹
F-diaziflurane	3.6 ± 0.1	1.4E5 ± 8.63	-1.3E4 ± 300	-20
H-diaziflurane	4.3 ± 0.3	7.7E4 ± 5.6E3	-1.2E4 ± 1.1E3	-20
Isoflurane	6.4 ± 0.2	1.50E4 ± 1.4E3	-6200 ± 280	-2.0

^aValues refer to interaction with the apoferritin 24-mer, single class binding model.

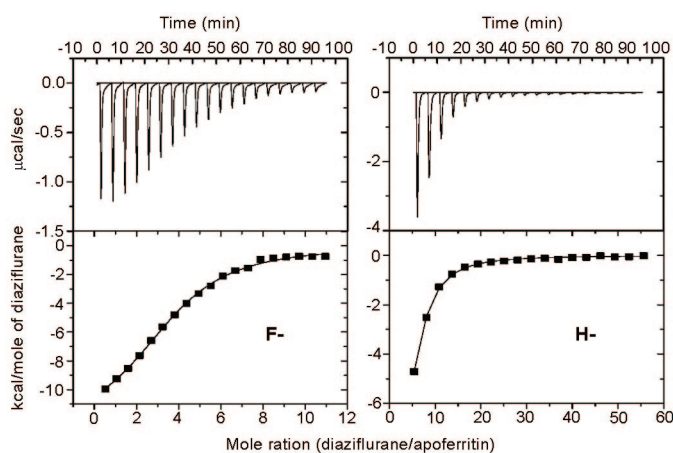


Figure 3. Isothermal titration calorimetry (ITC) of F- (left) or H-diaziflurane (right) into apoferritin. Both raw (top) and derived (bottom) data are shown, and the parameters of a single-class binding site fit are given in Table 2.

dipole from 0.9 (H-diaziflurane) to 1.8 D (F-diaziflurane). A degree of polarity is thought to be important for improving interactions within protein cavities (15).

The advantage of using diazirine chemistry for photolabeling is readily apparent in our ability to obtain high-resolution structural information on the adduction site; the long UV illumination does not impair crystallization or significantly alter resolution. We used two different crystals to collect two complete diffraction data sets and carry out two independent structure determinations for apoferritin labeled by H-diaziflurane. The crystals of photolabeled apoferritin proved to be isomorphous with those of the previously reported apoferritin–isoflurane complex (14), and the structure of the protein molecule does not appear to have been perturbed by the labeling process. Specifically, the rms differences in α carbon positions between our two structures and the apoferritin–isoflurane complex are 0.1 Å or lower, which is on the same order as the expected errors in the refined

structures. The variation of *B*-values with residue number is essentially identical for all three structures.

Electron density was noted in the same cavity where the binding of isoflurane and halothane has previously been observed (15). This cavity is centered on a crystallographic twofold symmetry axis that relates two ferritin monomers, and so ligands bound in this site must exhibit twofold disorder. Thus, even if every ligand binds in the site in exactly the same orientation, the electron density map will show a twofold averaged view, reflecting the presence of both that orientation and its symmetry-related conformer. The maximum occupancy for any ligand is therefore 50%, and if multiple orientations are present, the occupancy for each orientation will drop accordingly. Clear electron density was observed in the

cavity and had an elongated shape that matched well the backbone of H-diaziflurane (Figures 5 and 6). No clear electron density was seen to bridge the protein and

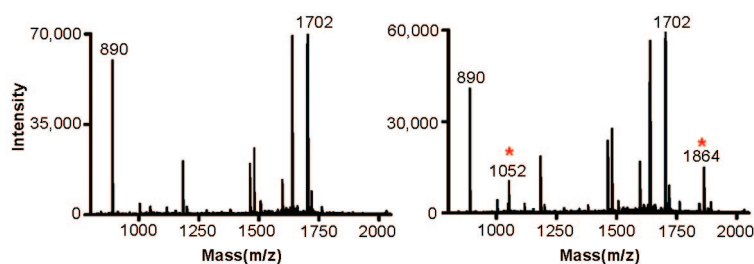


Figure 4. Mass spectra of trypsin-digested apoferritin L after either UV (left) or UV + H-diaziflurane (right) exposure. Note the appearance of two new peaks (red asterisks) in the lower spectra at 1052 and at 1864. These represent single adductions of the parent peptides 890 and 1702. No apoferritin adducts were detected with UV + F-diaziflurane. Tryptic peptide coverage for the two photolabels was 40–45%. Data were obtained with a 4700 Proteomics Analyzer (Applied Biosystems) via MALDI-TOF/TOF with 4000 series Explorer software (Applied Biosystems) for data collection, and Data Explorer software (Applied Biosystems) for data analysis.

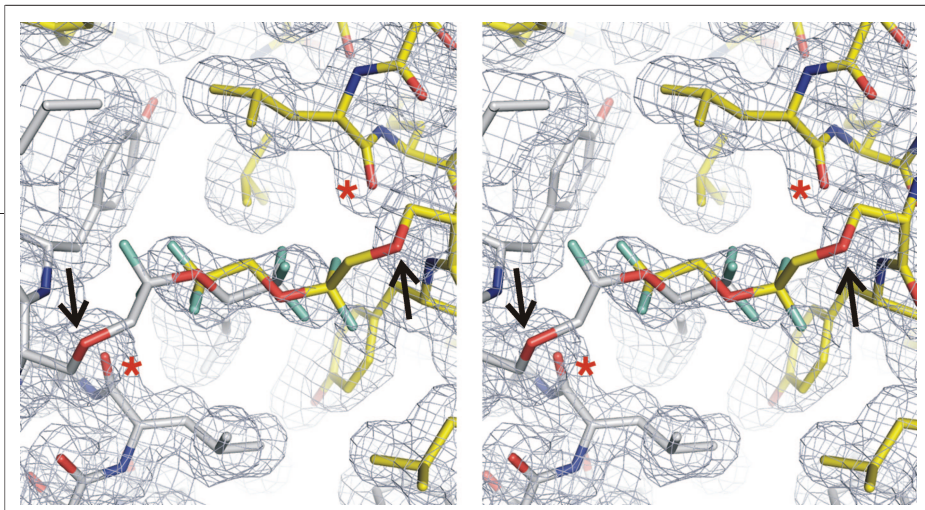


Figure 5. Stereo image of $2F_o - F_c$ electron density in the vicinity of the H-diaziflurane binding site. Elemental color code: oxygen, red; nitrogen, blue; fluorine, teal. One ferritin chain is colored yellow, and the symmetry-related chain is colored gray. The presumptive principal site of attachment of the photolabel is the γ oxygen of Ser-27 and is indicated by a black arrow. The secondary attachment site, the carbonyl oxygen of Leu-24, is indicated by a red asterisk. An $F_o - F_c$ omit map can be found in the Supplementary Figure 3. Figures 5 and 6 and Supplementary Figure 3 were prepared using PyMOL (<http://pymol.sourceforge.net/>).

the ligand electron density found in the cavity, and thus, the precise position of the covalent attachment site for the adduct cannot be conclusively obtained from the electron density map. The attachment site was modeled as the Ser-27 side chain, which allowed for the generation of a model that fits well to the electron density, displays good geometry, and agrees with the mass spectrometric data. However, the MS data also suggested adduction to the Leu-24, so we constructed a geometrically reasonable model in which the adduct is attached to the carbonyl oxygen of Leu-24. This latter model fits the electron density reasonably well, but not as well as when Ser-27 is used as the attachment point, explaining our choice of Ser-27 in the final refinement. It is not certain why no density is seen to bridge the body of the adduct with the protein; however, it is reasonable to surmise that the lack of density is due to the existence of more than one adduction site in the cavity, which leads to nonoverlapping orientations of the “business end” of the photolabel. When combined with less than fully efficient photoincorporation (perhaps <50%) and the inherent twofold disorder of ligands in this cavity, this heterogeneity could easily cause the electron density at the point of attachment to fall well below detectable levels. H-diaziflurane’s ability to photoadduct to these two residues may indicate a relative lack of photoselectivity

(although both bonds may be to oxygen). This is a desirable feature in a photolabel, in that it gives confidence that binding sites are being reliably reported. It is of interest, however, that we could not detect adduction to Tyr-28. Tyrosines and other aromatics are well-known to be preferred photochemical targets. Halothane, for example, generates a carbon-centered free radical on 250 nm irradiation, which appears to prefer either tyrosine or tryptophan in the nicotinic acetylcholine receptor

(6), serum albumin (16), and in a model four-helix bundle protein (17). The lack of tyrosine labeling in the apoferritin L-chain may relate to a steric hindrance to ligand movement within the cavity. It is not yet clear if there are residues that the UV-generated carbene will not react with, but apoferritin may be a good candidate for the mutagenesis necessary to find out.

Despite being of lower potency and of lower affinity for a model anesthetic-binding protein, the H-diaziflurane exhibits an important molecular activity associated with general anesthetics. H-diaziflurane increased the peak magnitude of the GABA-gated chloride current in a concentration-dependent manner (Figure 7), with an EC_{50} of $\sim 350 \mu\text{M}$ (isoflurane EC_{50} is 200–300 μM). At greater concentrations, this photolabel appeared to evoke a current in the absence of GABA, as has also

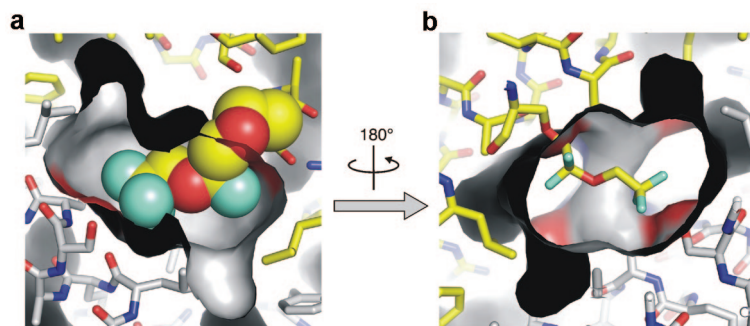


Figure 6. The binding cavity for H-diaziflurane. a) The binding site viewed from the same perspective as Figure 5. The molecular surface of the cavity is shown, colored according to the atom types forming the surface. The adduct is shown in a space-filling representation. b) View from the bottom of the cavity, looking outward toward the surface of the protein. Here the adduct is shown in a stick representation. In both panels a and b, only one of the two symmetry-related adduct positions is shown for the sake of clarity.

been reported for other volatile anesthetics (18), resulting in an apparent decrease in the potentiation of GABA-evoked current, probably due to cross desensitization.

It should be possible to further optimize H-diaziflurane's structure in a way that recovers the dipole moment, increases hydrophobicity, and retains the hydrogen on the diazirine carbon. One possibility is monochlorination of the methylene group. This will make it more like isoflurane and, at the same time, acidify the hydrogen on the 2-carbon, improving its ability to undergo exchange reactions with either deuterium or tritium for labeling purposes. Radiolabeling is especially useful in combination with Edman degradation to identify drug binding sites in complex membrane proteins not well-suited to MS approaches, such as the nicotinic acetylcholine receptor (6). Of relevance here is our finding that a backbone atom (leucine carbonyl oxygen) is adducted by H-diaziflurane, and the Protein Data Bank (PDB) clearly shows that backbone atoms contribute to the surface of many cavities. It is probable that adduction of such atoms will impede the Edman chemistry and result in an under-reporting of labeled residues in these proteins. It is less clear whether such an adduction would influence MS sequencing.

In summary, the combined measurements show that H-diaziflurane (a) is a nontoxic, reversible, immobilizing compound in *Xenopus* tadpoles; (b) produces effects on

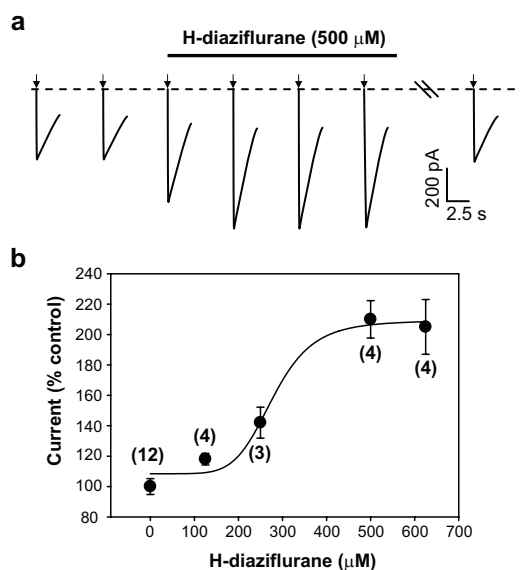


Figure 7. H-diaziflurane potentiates GABA-evoked current in voltage-clamped hippocampal neurons in primary culture. **a)** GABA (5 μM) was applied by pressure pulse (20 psi \times 50 ms) every 45 s (arrows), and the diaziflurane was applied by bath perfusion. The current magnitude returned to the control level 10 min after initiation of drug washout. **b)** Summary of H-diaziflurane effects on GABA-evoked current (mean \pm SEM). The points are fit to a Hill equation of variable slope. ($\text{EC}_{50} = 280 \pm 80 \mu\text{M}$, slope = 6.1 ± 13). Numbers in parentheses are the number of cells recorded.

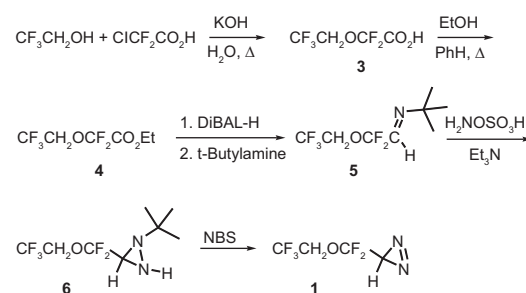
mammalian GABA-gated channels that are not different from inhaled anesthetics; (c) binds in the same specific apoferritin site as isoflurane, interacting with the same residues; and (d) photoadducts to these same residues. H-diaziflurane is therefore a satisfactory photolabel mimic of the haloether anesthetics and should be useful to identify their sites of action within molecular targets and, with minor modifications, to identify molecular targets in complex mixtures, such as synaptic membranes.

METHODS

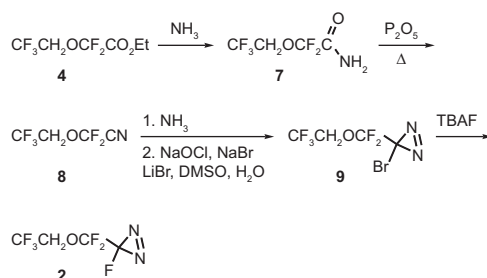
Materials. Equine apoferritin was obtained from Sigma. All other chemicals were of reagent grade or better and were obtained from Sigma or Aldrich. CAUTION: All diazirines and diazo compounds are potentially explosive and should be treated with care. All new compounds exhibited IR and high-resolution mass spectra consistent with the assigned structures. Final purified products were >98% pure by gas chromatographic (GC) analysis using a 30 m dimethylsilicone capillary column and flame ionization detection.

Synthesis of Compound 1 (H-Diaziflurane) and Compound 2 (F-Diaziflurane). The preparation of diazirine **1** is shown (Scheme 1). (2,2,2-Trifluoroethoxy)-2,2-difluoroacetic acid (**3**) was prepared according to the literature procedure (19). Esterification of **3** using ethanol under Dean-Stark conditions yielded ethyl ester **4** in 69% yield. Reduction of **4** with DIBAL-H in toluene followed by condensation of the crude aldehyde with *tert*-butyl amine produced imine **5** in 40% yield. Treatment of **5** with hydroxylamine-*O*-sulfonic acid in triethylamine/ethanol produced diaziridine **6** in modest yield. Conversion of **6** to diazirine **1** was accomplished using *N*-bromosuccinimide according to our previously developed conditions (8).

The preparation of diazirine **2** started by conversion of ethyl ester **4** to amide **7** (Scheme 2). Dehydration of **7** using P_2O_5 produced nitrile **8** in good yield. Treatment of **8** with ammonia gave the intermediate amidine that was directly converted to bromodiazirine **9** using Graham's procedure (20). Replacement of bromide in **9** by fluoride in **2** was accomplished using tetrabutylammonium fluoride (TBAF).



Scheme 1. Synthesis of H-diaziflurane (**1**)



Scheme 2. Synthesis of F-diazirurane

Preparation of (2,2,2-Trifluoroethoxy)-2,2-difluoroacetic acid (3). This compound was prepared from chlorodifluoroacetic acid, trifluoroethanol, and aqueous potassium hydroxide in 60% yield according to the literature procedure (19).

Preparation of Ethyl (2,2,2-Trifluoroethoxy)-2,2-difluoroacetate (4). A mixture of 172 g (0.89 mol) of acid **3**, 1 L of benzene, and 82 mL (65 g, 1.4 mol) of anhydrous ethanol was heated under reflux for 24 h, while water was collected in a Dean-Stark trap. A total of 19 mL of water was collected. The benzene was removed by careful distillation, and the product was fractionated through a Vigreux column to yield 137 g (69%) of clear colorless oil, bp 138–140 °C. ¹H NMR (CDCl₃): δ 4.36 (q, *J*_{H-H} = 7.1 Hz, 2H), 4.27 (q, *J*_{H-F} = 8.0 Hz, 2H), 1.36 (t, *J*_{H-H} = 7.1 Hz, 3H). ¹³C NMR: δ 158.9 (t, *J*_{C-F} = 41 Hz), 122.3 (q, *J*_{C-F} = 277 Hz), 114.0 (t, *J*_{C-F} = 274 Hz), 63.9, 61.2 (qt, *J*_{C-F} = 38 Hz, 6 Hz), 13.6. ¹⁹F NMR: δ -74.9 (t, *J*_{H-F} = 9 Hz, 3F), -81.4 (s, 2F). HRMS (CI+): *m/z* calcd for C₆H₈F₅O₃ (M + H), 223.0393; found, 223.0394.

Preparation of (2,2,2-Trifluoroethoxy)-2,2-difluoroacetamide (7). An ice-cooled 500 mL 3-neck flask was filled with 68.8 g (0.31 mol) of ester **4**, 230 mL of anhydrous ether, and a stir bar. One joint was fitted with a reflux condenser attached to an oil bubbler, one neck was sealed, and one neck had a glass dispersion tube. Gaseous anhydrous ammonia was slowly bubbled into the stirred solution over the course of 2 h. The mixture was allowed to warm to RT, and the solvents and excess ammonia were removed *in vacuo* leaving a mass of white crystals, 58.6 g (98%), mp 54 °C. ¹H NMR (DMSO-*d*₆): δ 8.31 (bs, 1H), 8.13 (bs, 1H), 4.66 (q, *J*_{H-F} = 8.8 Hz, 2H). ¹³C NMR: δ 160.1 (t, *J*_{C-F} = 36 Hz), 123.0 (q, *J*_{C-F} = 277 Hz), 115.1 (t, *J*_{C-F} = 275 Hz), 60.8 (qt, *J*_{C-F} = 36 Hz, 6 Hz). ¹⁹F NMR: δ -72.6 (t, *J*_{H-F} = 9 Hz, 3F), -80.0 (s, 2F). HRMS (CI+): *m/z* calcd for C₄H₅F₅NO₂ (M + H), 194.0240; found, 194.0240.

Preparation of (2,2,2-Trifluoroethoxy)-2,2-difluoroacetonitrile (8). Crude acetamide **7** (58 g, mol) was added to a 1 L round-bottom flask along with a stir bar and 57.6 g (0.203 mol) of P₂O₅. A short-path distillation apparatus was attached. The flask was slowly heated in an oil bath to 180 °C, and product was collected in the range of 55–65 °C. Redistillation gave 36.7 g (0.21 mol) of clear colorless liquid in 69% yield, bp 60 °C. The nitrile **8** is hygroscopic. ¹H NMR (CDCl₃): 4.31 (q, *J*_{H-F} = 7.7 Hz, 2H). ¹³C NMR: δ 121.5 (q, *J*_{C-F} = 277 Hz), 108.9 (t, *J*_{C-F} = 259 Hz), 108.7 (t, *J*_{C-F} = 58 Hz), 62.4 (qt, *J*_{C-F} = 39, 4 Hz). ¹⁹F NMR: δ -61.0 (q, *J*_{F-F} = 2 Hz, 2F), -74.7 (tt, *J*_{H-F} = 8 Hz, *J*_{F-F} = 2 Hz, 3F). HRMS (CI+): *m/z* calcd for C₄H₅F₅NO (M + H), 174.00; found, 174.01.

Preparation of 2,2-Difluoro-2-(2,2,2-trifluoroethoxy)bromodiazirine (9). A 250 mL 3-necked, round-bottom flask was filled with 4.0 g (23 mmol) of (2,2,2-trifluoroethoxy)-2,2-difluoroacetonitrile (**8**) and a stir bar. A dry ice/acetone-cooled gas condenser was attached to the center neck, one joint was closed, and the third was attached via safety traps to a cylinder of ammonia. The flask was cooled to -78 °C, and ammonia was slowly condensed into it. After the volume in the flask had approximately tripled, the addition was stopped, and the

cooling bath was removed. The solution was allowed to reflux for 1.5 h. The gas condenser was removed and attached to another flask, and the ammonia was allowed to distill from the reaction flask and collect in the second flask. Once all the ammonia had evaporated, the solid yellow mass of crystals was dried *in vacuo* to leave 4.2 g (21.9 mmol, 95%) of yellow crystalline solid, mp 44 °C. The amidine was used without further purification for the preparation of diazirine **9**.

A 5-L 3-necked, round-bottom flask containing a stir bar was fitted with 100 mL of DMSO, 10.2 g (11.7 mmol) of dry LiBr, and 4.2 g (21.9 mmol) of crude amidine. One neck of the flask was fitted with a 500 mL pressure-equalizing dropping funnel containing a solution of 40 g (0.39 mol) of NaBr dissolved in 180 mL of 12% sodium hypochlorite. A glass stopper was added to the third neck of the flask. A vacuum adapter was connected to the third neck and led through a series of three traps maintained at -40, -78, and -196 °C to a mechanical vacuum pump. The system was evacuated, and once the pressure of the system reached 0.1 Torr, the stir bar was started and the sodium hypobromite solution was added to the flask over the course of ~1 min. The products were continuously pumped through the train of U-traps for the next 20 min, then the reaction was stopped. Volatile material from the -78 °C trap was transferred under vacuum onto a small amount of P₂O₅ to remove any water. Transfer of volatile material from this flask into a gas storage bulb gave 0.98 g (16%) of clear colorless liquid. Final purification of **9** used preparative GLC. ¹H NMR: δ 4.23 ppm (q, *J*_{H-F} = 6.3 Hz); ¹⁹F NMR: δ -75.1 (m, 3F), -75.9 ppm (m, 2F).

Preparation of 2,2-Difluoro-2-(2,2,2-trifluoroethoxy)fluorodiazirine (2). Solvent was evaporated from 10 mL of 1 M TBAF in THF to leave a powdery white solid. Five milliliters of 1,2-dichlorobenzene was added to the flask along with a Teflon-coated magnetic stir bar, and the mixture was stirred to dissolve the solid. A high vacuum was applied to the flask, and the remaining water and tetrahydrofuran (THF) were removed and trapped in a dry ice-cooled trap. A solution of 5 mL of dichlorobenzene and bromodiazirine **9** (0.9 g, 4 mmol) was added to the ice-cooled flask. The mixture was allowed to stir for several hours and then was refrigerated overnight. The next day, the mixture was fractionated through a train of three U-traps cooled to -35, -78, and -196 °C under high vacuum. Clear colorless liquid (0.4 g, 48%) was collected in the -196 °C U-trap. The product was purified by preparative GC using a 1/4 in. × 10 feet column of 10% SF-96 on Chromasorb W. ¹H NMR: δ 4.3 (dq, *J*_{F-F} = 2.8 Hz, *J*_{H-F} = 8 Hz); ¹³C NMR: δ 122.22 (q, *J*_{C-F} = 277 Hz), 117.9 (qd, *J*_{C-F} = 265, 40 Hz), 68.7 (d, *J*_{C-F} = 277 Hz), 61.6 (qt, *J*_{C-F} = 38, 5 Hz); ¹⁹F NMR: δ -74.9 (tt, *J*_{H-F} = 8 Hz, *J*_{F-F} = 2 Hz, 3F), -80.4 (dq, *J*_{F-F} = 7 Hz, *J*_{F-F} = 2 Hz, 2F), -170.4 (t, *J*_{F-F} = 7 Hz).

Preparation of *tert*-Butyl-(2,2-Difluoro-2-(2,2,2-trifluoroethoxy)ethylidene) Amine (5). A solution of 21.0 g (95 mmol) of ester **4** and 100 mL of anhydrous ether was cooled with stirring to -78 °C under a nitrogen atmosphere. A solution of 75 mL (0.11 mol) of 1.5 M DiBAL-H in toluene was added dropwise *via* a syringe over the course of 20 min. The solution was stirred for an additional 2 h at -78 °C and was allowed to warm and stir at 0 °C for 30 min. After it was recooled to -78 °C, the mixture was quenched with 4 mL of methanol. After warming to RT, the mixture was slowly added with good stirring to 300 mL of a 5% HCl solution. The resulting mixture was extracted with three 100 mL portions of ether. The ether was dried over sodium sulfate and carefully evaporated under reduced pressure below RT until most of the ether had been removed. Freshly distilled *tert*-butyl amine (15 mL, 290 mmol) was added, and the mixture heated to reflux for 2 h. Careful distillation through a short-path apparatus yielded 9.0 g (40%) of clear colorless liquid imine **5**, bp 138–140 °C. ¹H NMR (CDCl₃): 7.49 (t, *J*_{H-F} = 4.4 Hz, 1H), 4.30 (q, *J*_{H-F} = 8 Hz, 2H), 1.26 (s, 9H). ¹³C NMR: δ 146.8 (t, *J*_{C-F} = 35 Hz),

123.0 (q, $J_{C-F} = 277$ Hz), 118.5 (t, $J_{C-F} = 265$ Hz), 61.0 (qt, $J_{C-F} = 37$, 6 Hz), 59.0, 29.1. ^{19}F NMR: δ -74.9 (t, $J_{F-F} = 8$ Hz, 3F), -79.3 (bs, 2F). HRMS (Cl⁺): m/z calcd for $C_8H_{12}F_5NO$ (M + H), 234.0917; found, 234.0923.

Preparation of 1-*tert*-Butyl-3-(Difluoro-(2,2,2-trifluoroethoxy)methyl)diaziridine (6). A 100 mL round-bottom flask was filled with 2.4 g (10 mmol) of imine **5**, 5 mL of anhydrous triethylamine, and 9 mL of absolute ethanol. The solution was cooled in an ice bath, and 2.4 g (21 mmol) of hydroxylamine-*O*-sulfonic acid was added in one portion with good stirring. The mixture was stirred in an ice bath for 2 h, then placed in a freezer overnight. The next day, solvent was removed under aspirator pressure below RT. The resulting white slurry was extracted with ether (3 \times 25 mL). The combined ether extracts were evaporated to leave 1.2 g (~50%) of clear colorless oil sufficiently pure for the next reaction. An analytical sample was purified by preparatory GC using a 1/4 in. \times 10 ft column of 10% SF-96 on Chromasorb W. 1H NMR (CDCl₃): 7.50 (t, $J_{H-F} = 4$ Hz, 1H), 4.30 (q, $J_{H-F} = 8$ Hz, 2H), 1.26 (s, 9H). ^{19}F NMR: δ -80.3 (m, 3F), -84.8 (m, 2F). HRMS (Cl⁺): m/z calcd for $C_8H_{14}F_5N_2O$ (M + H), 249.1026; found, 249.1016.

Preparation of 3-(Difluoro-(2,2,2-trifluoroethoxy)methyl)-3H-diazirine (1). To a solution of 1.0 g (4 mmol) of crude diaziridine **6** in 3 mL of 1,2-dichloroethane was added 0.7 g (4 mmol) of *N*-bromosuccinimide in one portion. The mixture was stirred at RT for 4 h. The volatile material was evaporated under high vacuum and collected in a U-trap cooled to -78 °C. Product **1** was purified by preparatory GLC using a 1/4 in. \times 10 ft column of 10% SF-96 on Chromasorb W to give a volatile, clear, colorless liquid (250 mg, ~30%). 1H NMR (CDCl₃): 4.18 (q, $J_{H-F} = 7.7$ Hz, 2H), 1.63 (t, $J_{H-F} = 7.7$ Hz, 1H). ^{13}C NMR: δ 122.0 (q, $J_{C-F} = 277$ Hz), 119.8 (t, $J_{C-F} = 262$ Hz), 60.8 (qt, $J_{C-F} = 38$, 4 Hz), 20.7 (t, $J_{C-F} = 44$ Hz). ^{19}F NMR: δ -75.0 (tt, $J_{H-F} = 8$ Hz, $J_{F-F} = 2$ Hz, 3F), -76.6 (tt, $J_{H-F} = 8$ Hz, $J_{F-F} = 2$ Hz, 2F).

Physical Properties: Water Solubility and Hydrophobicity.

Octanol water partition coefficients were used to estimate hydrophobicity. Excess F- and H-diaziflurane were solubilized in water with vigorous vortexing and brief sonication. After centrifugation at 1000g for 10 min, the solution was loaded into 10 mL gas-tight Hamilton syringes. Absorbance at 330 nm of an aliquot was recorded to allow calculation of maximal water solubility (after using methanolic solutions to calculate extinction coefficients), and then exactly 1 mL of octanol was drawn into the syringe and mixed by rotation for an hour. The octanol-water mixture was allowed to completely separate for another hour, and then absorbance of the water phase was measured again. Molar partition was calculated by multiplying the difference in water absorbance by the ratio of water-to-octanol volume, and dividing by the ending water absorbance.

Isothermal Titration Calorimetry. Briefly, titrations were performed at 20 °C using a Microcal, Inc. VP ITC. The thermodynamic parameters for the binding of F- and H-diaziflurane to equine apoferritin at 20 °C were determined by ITC using a Microcal, Inc. VP ITC (www.microcalorimetry.com). The ITC consists of a matched pair of sample and reference vessels (1.43 mL) enclosed in an adiabatic enclosure and a rotating stirrer-syringe for titrating aliquots of the ligand solution into the sample vessel. The sample cell contained 0.008 μ M sample of protein, and the reference cell contained water. Saturated photolabel (0.4 mM F-diaziflurane or 2.0 mM H-diaziflurane) was loaded in the syringe (volume = 0.28 mL) for injecting into the sample. Four separate titrations were performed, including three controls, ligand into buffer, buffer into protein, and buffer into buffer, which were then used to correct the experimental titration, ligand into protein. Origin 5.0 (Microcal Software, Inc.) was used to fit thermodynamic parameters to the heat profiles.

Photolabeling. Small aliquots of pH 7 phosphate buffer containing 4 mg mL⁻¹ apoferritin with and without saturated F- or H-diaziflurane were placed in a 1 mm path length quartz cuvette and exposed to 305 nm light (Rayonet RPR-3000 lamp: emission from 280 to 320 nm) at 2 mm distance for 5 min. Control samples received only UV irradiation. The UV-treated apoferritin was then passed through an HPLC C-18 analytical column to separate the labeled apoferritin H and L subunits, which were then resuspended in 0.1% TFA. Small aliquots of UV-treated samples L-chain were trypsinized and examined with MALDI-TOF for a shift in trypsin fragment molecule weight (MW) that would indicate covalent incorporation of the adduct (+180 Da for F-diaziflurane; +162 Da for H-diaziflurane). After identification of adducted fragments, MALDI-TOF/TOF approaches were used to sequence the fragment in an effort to identify the adducted residues.

Crystallization. Aliquots of H-diaziflurane-labeled apoferritin were washed by repeated centrifuge filtration, and then used to grow large, single crystals in hanging drop experiments, as previously described (14), except that these hanging drops were in standard multiwell plates. Labeled apoferritin crystals took several weeks to grow to full size and were stable thereafter for months. Approximately 3-month-old crystals were mounted in nylon loops, cryoprotected by passing them rapidly through 30% (w/w) glycerol in reservoir buffer, and flash-cooled in liquid nitrogen.

Diffraction Data Collection and Processing. Two complete diffraction data sets were collected from two crystals at beamline X6A of the National Synchrotron Light Source. Crystals were maintained at 100 K in a stream of cold nitrogen gas during data collection. Data processing was carried out using d*TREK (21) (details are given in Supplementary Table S1). Structures were refined independently for the two data sets. In each case, protein coordinates from PDB file 1XZ1 were used as a starting model for the refinements. Initial models containing protein atoms only were used for automated water placement and refinement using Arp-Warp and Refmac (22, 23). Cadmium ions in the lattice were identified as peaks in the anomalous difference density maps. Ligands were placed into difference maps using XtalView (24), and refinement of the complex structure was completed using Refmac. Stereochemical restraints for H-diaziflurane were derived from the Dundee PRODRG server (<http://davapc1.bioch.dundee.ac.uk/programs/prodrgr/>). Refinement details are shown in Supplementary Table S1. The two independent refinements converged to essentially identical structures, with an rms deviation in C α positions = 0.036 Å. Coordinates for the complex from the crystal 1 refinement have been deposited in the PDB, accession number 2GYD.

Tadpole Studies. *Xenopus* tadpoles were used to examine the anesthetic potency of both F- and H-diaziflurane. Ten tadpoles were placed in 20 mL sealed glass vials containing pond water and increasing concentrations of the compounds, presolubilized by vigorous shaking/sonication of pondwater with aliquots of neat compound. Volumes were adjusted to minimize gas volume in the vials. Tadpoles were observed for the loss-of-righting reflex for 5 min, and then immediately transferred into large containers of fresh pondwater and carefully observed for recovery. The tadpoles were then observed intermittently for 24 h. Control experiments verified an absence of effect of the manipulations on tadpole activity in the absence of added compound. Percent immobile at each concentration were fitted to variable slope Hill plots. Further control studies were able to reproduce the expected EC₅₀ for isoflurane of ~0.25 mM.

GABA_A Receptor Electrophysiology. E18 hippocampi preserved in Hibernate E solution were obtained from BrainBits and immediately plated onto poly-D-lysine-coated 10 mm glass coverslips following the vendor's suggested protocol. Experiments were

performed on neuronal cultures after 10–14 d *in vitro*, and only with H-diaziflurane. A coverslip with neurons was placed inside a recording chamber continuously perfused with an external medium (in mM: 140 NaCl, 2.8 KCl, 1 CaCl₂, 1 MgCl₂, 10 dextrose, 10 HEPES, pH 7.3 with NaOH and 1 μM TTX), and whole cell patch clamp recordings were obtained using a 1.5 mm glass capillary microelectrode with a resistance of 2–3 MΩ when filled with an internal solution (in mM: 140 CsCl, 4 NaCl, 2 MgCl₂, 10 K-EGTA, 10 HEPES, 2 Na-ATP, pH 7.2 with CsOH). The holding potential was fixed at –60 mV thus giving an inward flowing, GABA-evoked current. GABA (5 μM) diluted in the external medium was applied by pressure injection (20 psi × 50 ms) from a pipet placed ~1 cell diameter from the patched neuron. This concentration of GABA corresponds to approximately an EC₁₅ in this preparation; thus, the agonist concentration was well below saturation, a condition required to detect the increase in the current magnitude due to other allosteric modulators. A compound application interval of 45 s prevented cumulative desensitization and a decline in the response amplitude. Data acquisition and compound application were controlled by pClamp v9.0 (Molecular Devices).

Accession codes: Structure factors and final coordinates have been deposited in the RCSB PDB with ID code 2GYD.

Acknowledgments: X-ray diffraction data were collected at the X6A beam line, National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy under contract No. DE-AC02-98CH10886. X6A is funded by NIH/NIGMS under agreement Y1 GM-0080-03. This work was supported by GM-55876 (R.G.E.), and NS-45718 (J.Y.). The authors thank Susan Lin and Thuy Tran for dedicated technical assistance.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Forman, S. A., Miller, K. W., and Yellen, G. (1995) A discrete site for general anesthetics on a postsynaptic receptor. *Mol. Pharmacol.* **48**, 574–581.
- Mascia, M. P., Trudell, J. R., and Harris, R. A. (2000) Specific binding sites for alcohols and anesthetics on ligand-gated ion channels. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9305–9310.
- Eckenhoff, M. F., Chan, K., and Eckenhoff, R. G. (2002) Multiple specific binding targets for inhaled anesthetics in the mammalian brain. *J. Pharmacol. Exp. Ther.* **300**, 172–179.
- Xi, J., Liu, R., Asbury, G. R., Eckenhoff, M. F., and Eckenhoff, R. G. (2004) Inhalational anesthetic-binding proteins in rat neuronal membranes. *J. Biol. Chem.* **279**, 19628–19633.
- Eckenhoff, R. G. (1996) An inhalational anesthetic binding domain in the nicotinic acetylcholine receptor. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2807–2810.
- Chiara, D. C., Dangott, L. J., Eckenhoff, R. G., and Cohen, J. B. (2003) Identification of nicotinic acetylcholine receptor amino acids photolabeled by the volatile anesthetic halothane. *Biochemistry.* **42**, 13457–13467.
- Ishizawa, Y., Pidikiti, R., Liebman, P. A., and Eckenhoff, R. G. (2002) G protein-coupled receptors as direct targets of inhaled anesthetics. *Mol. Pharmacol.* **61**, 945–952.
- Eckenhoff, R. G., Knoll, F. J., Greenblatt, E. P., and Dailey, W. P. (2002) Halogenated diazirines as photolabel mimics of the inhaled haloalkane anesthetics. *J. Med. Chem.* **45**, 1879–1886.
- Eckenhoff, M. F., and Eckenhoff, R. G. (1998) Quantitative autoradiography of halothane binding in rat brain. *J. Pharmacol. Exp. Ther.* **285**, 371–376.
- Greenblatt, E. P., and Meng, X. (2001) Divergence of volatile anesthetic effects in inhibitory neurotransmitter receptors. *Anesthesiology* **94**, 1026–1033.
- Pratt, M. B., Husain, S. S., Miller, K. W., and Cohen, J. B. (2000) Identification of sites of incorporation in the nicotinic acetylcholine receptor of a photoactivatable general anesthetic. *J. Biol. Chem.* **275**, 29441–29451.
- Ziebell, M. R., Nirthanan, S., Husain, S. S., Miller, K. W., and Cohen, J. B. (2004) Identification of binding sites in the nicotinic acetylcholine receptor for [3H]azietomidate, a photoactivatable general anesthetic. *J. Biol. Chem.* **279**, 17640–17649.
- Darbandi-Tonkabon, R., Hastings, W. R., Zeng, C. M., Akk, G., Manion, B. D., Bracamontes, J. R., Steinbach, J. H., Mennerick, S. J., Covey, D. F., and Evers, A. S. (2003) Photoaffinity labeling with a neuroactive steroid analogue. 6-azi-pregnanolone labels voltage-dependent anion channel-1 in rat brain. *J. Biol. Chem.* **278**, 13196–13206.
- Liu, R., Loll, P. J., and Eckenhoff, R. G. (2005) Structural basis for high-affinity volatile anesthetic binding in a natural 4-helix bundle protein. *FASEB J.* **19**, 567–576.
- Liu, R., and Eckenhoff, R. G. (2005) Weak polar interactions confer albumin binding site selectivity for haloether anesthetics. *Anesthesiology* **102**, 799–805.
- Eckenhoff, R. G., Petersen, C. E., Ha, C. E., and Bhagavan, N. V. (2000) Inhaled anesthetic binding sites in human serum albumin. *J. Biol. Chem.* **275**, 30439–30444.
- Johansson, J. S., Scharf, D., Davies, L. A., Reddy, K. S., and Eckenhoff, R. G. (2000) A designed four-alpha-helix bundle that binds the volatile general anesthetic halothane with high affinity. *Biophys. J.* **78**, 982–993.
- Yang, J., Isenberg, K. E., and Zorumski, C. F. (1992) Volatile anesthetics gate a chloride current in postnatal rat hippocampal neurons. *FASEB J.* **6**, 914–918.
- Huang, D. G., Rozov, L. A., Halpern, D. F., and Vernice, G. G. (1993) Preparation of the isoflurane enantiomers. *J. Org. Chem.* **58**, 7382–7387.
- Graham, W. H. (1965) The halogenation of amidines. I. Synthesis of 3-halo- and other negatively substituted diazirines. *J. Amer. Chem. Soc.* **87**, 4396–4397.
- Pflugrath, J. W. (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1718–1725.
- Lamzin, V. S., Perrakis, A., and Wilson, K. S. (2001) in International Tables for Crystallography (Rossmann, M. G., and Arnold, E., Eds.) pp 720–722, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
- McRee, D. E. (1999) XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.

Probing the Barrier Function of the Outer Membrane with Chemical Conditionality

Natividad Ruiz[†], Tao Wu[‡], Daniel Kahne^{*§}, and Thomas J. Silhavy^{†*}

[†]Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, [‡]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 01238, and [§]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

The outer membrane (OM) is a defining feature of Gram-negative bacteria. It is located outside the peptidoglycan cell wall and it functions as a barrier to protect these organisms from toxic agents in their environment. This membrane is a lipid bilayer, but it is asymmetric, containing phospholipids in the inner leaflet and glycolipids such as lipopolysaccharide (LPS) in the outer leaflet. Integral OM β -barrel proteins function as porins to allow passage of small hydrophilic molecules across the barrier. Numerous lipoproteins are localized to the OM as well (1).

All of the molecular components of the OM are synthesized in the cytoplasm or the inner leaflet of the inner membrane (2). Somehow these components are transported across the aqueous periplasmic space that separates the inner and outer membranes and then assembled in the OM bilayer. Since the OM delimits the cell, its biogenesis occurs in direct contact with the external environment on one side and the bacterial periplasm on the other, both of which lack an obvious energy source such as ATP. Although it has long been of interest to understand how the OM is built and its integrity maintained during cell growth and division, identifying factors involved in OM biogenesis has proven difficult (2).

We recently described a genetic technique called chemical conditionality to probe OM biogenesis. In chemical conditionality, one starts with a mutant strain which has a defective OM with a compromised barrier function (*i.e.*, increased permeability) and which therefore is more sensitive to toxic small molecules such as antibiotics. This mutant strain is used in selections that demand resistance to toxic small molecules in order to isolate suppressors with improved OM barrier function; these suppressor mutations are in genes that specify factors involved in OM biogenesis (3). Using this

ABSTRACT A key function of biological membranes is to exclude toxic small molecules while allowing influx of nutrients. Cells achieve this by controlling the composition of different types of proteins and lipids within the membrane by a process called membrane biogenesis. We have recently proposed a strategy to identify genes involved in membrane biogenesis in Gram-negative bacteria such as *Escherichia coli* by selecting for suppressors of mutations that render the outer membrane (OM) leaky. We predicted that different small molecules could select different suppressors because mutations that answer a specific selection will correct the membrane permeability defect to different degrees depending on the structure of the small molecule. We have tested this hypothesis by selecting for resistance to bile acids in an *imp4213* strain, which contains a compromised OM owing to a defect in lipopolysaccharide biogenesis. We report here that a suppressor mutation in *yaeT*, which specifies an essential protein involved in the assembly of β -barrel proteins in the OM, confers resistance to a specific subset of bile acids in the *imp4213* strain. *YaeT* is conserved from bacteria to man because it is involved in OM biogenesis in mitochondria and chloroplasts. These results demonstrate that structurally different toxic small molecules select different, and highly specific, genetic solutions for correcting membrane-permeability defects. The remarkable chemical specificity of the *imp4213* suppressors provides insights into the molecular nature of the OM permeability barrier.

*Corresponding author,
tsilhavy@princeton.edu.

Received for review March 21, 2006
and accepted May 30, 2006.

Published online July 7, 2006
10.1021/cb600128v CCC: \$33.50

© 2006 by American Chemical Society

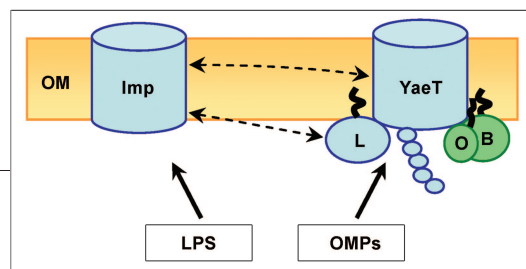


Figure 1. Imp and the YaeT–lipoprotein complex. The β -barrel protein Imp assembles LPS in the OM, while the YaeT complex assembles OMPs. The YaeT complex is composed of the β -barrel protein YaeT and the three lipoproteins YfgL (labeled “L”), YfiO (labeled “O”), and NlpB (labeled “B”). The five periplasmic POTRA domains of YaeT are shown as ovals. Recently, it has been shown that YfgL is not required for the association of YfiO and NlpB with YaeT but that YfiO and NlpB do require each other to form a stable complex with YaeT (17). The dotted lines denote genetic interactions between Imp and YfgL and YaeT.

approach, we identified YfgL, a component of the cellular machinery that assembles β -barrel OM proteins (OMPs) in Gram-negative bacteria such as *Escherichia coli* (Figure 1) (3, 4). Starting with a strain carrying the *imp4213* mutation (5), which alters Imp, the OMP that assembles LPS at the OM (6, 7), *yfgL* was identified by selecting for secondary mutations that partially suppress the OM permeability defects conferred by *imp4213*. In particular, *yfgL* was identified by selecting for resistance to the glycolipid derivative of vancomycin, chlorobiphenyl vancomycin (3, 8).

Results obtained with chemical conditionality were striking in two respects. First, there existed a continuum of suppressors that ameliorated the permeability defects conferred by *imp4213* to varying degrees (3). For example, intragenic suppressors carrying a second mutation in the *imp4213* gene confer resistance to all antibiotics tested, including vancomycin, chlorobiphenyl vancomycin, and erythromycin, and the mixture of bile salts present in the common laboratory media MacConkey agar. These intragenic suppressors appear to restore the OM barrier to nearly wild-type effectiveness. At the other end of the continuum, some of the suppressor mutations obtained by selection for resistance to bile salts confer resistance to bile salts only. These suppressor strains remain sensitive to vancomycin, chlorobiphenyl vancomycin, and erythromycin because the suppressor mutations restore the barrier function of the OM to a limited degree. The *yfgL* suppressors restore the barrier function of the OM to an intermediate degree, lying in the middle of the continuum; they confer resistance to chlorobiphenyl vancomycin and bile salts, but not vancomycin or erythromycin.

The second striking aspect of the results was the remarkable chemical specificity exhibited by the *yfgL* suppressors (3). The *imp4213* strains carrying the *yfgL*

suppressors were resistant to chlorobiphenyl vancomycin, but they remained sensitive to vancomycin (8). Heretofore such chemical specificity has been observed only with mutations that alter the drug target. However, *yfgL* mutations do not alter the drug target; they prevent access of the drug to its target by altering the OM permeability barrier (3).

We predicted that we might identify genes that specify additional components of the OM assembly machinery by looking for *imp4213* suppressors that expand the continuum. If so, we were interested in whether such suppressors would exhibit the same remarkable chemical specificity as the *yfgL* suppressors. Therefore, we examine here an uncharacterized *imp4213* suppressor isolated on MacConkey agar (3). This medium contains a crude mixture of bile salts, yet many pure bile salts of closely related chemical structures are readily available. Here, we show that the chemical conditionality approach revealed another component of the OM β -barrel assembly machinery, YaeT, and that the *yaeT* suppressor also exhibits remarkable chemical specificity. These results demonstrate that small molecules can be exquisitely sensitive probes of OM structure, and so different small molecules can serve as reagents to discover different genes involved in the assembly of the membrane.

RESULTS AND DISCUSSION

Chemical conditionality is a chemical genetics approach that uses toxic small molecules in selections employing strains with permeability defects to create particular chemical conditions that demand specific suppressor mutations which decrease membrane permeability (3). Thus, chemical conditionality exploits the fact that toxic molecules must first cross a membrane before reaching their intracellular target (3). When applying chemical conditionality, one searches for mutations that inhibit or decrease the efficiency of this entry step with the idea that such mutations will be in genes encoding factors involved in the membrane biogenesis. Here, we show that different suppressors obtained using chemical conditionality exhibit chemical specificity because the entry of each chemical into the cell is determined by its physicochemical properties.

We found that a previously isolated mutation that suppresses sensitivity of *imp4213* cells to the mixture of bile salts present in MacConkey agar (3) maps to *yaeT* (see Methods). This suppressor allele, which we named

yaeT6, carries a 6-bp in-frame insertion (agaaac) between codons 218 and 219 of *yaeT*. In the wild-type *yaeT* coding sequence, the agaaac sequence is repeated in tandem, while in the mutant allele, this sequence appears three times in tandem. Therefore, it is likely that this mutation arose as a replication error caused by DNA polymerase slippage (9). At the protein level, the suppressor mutation resulted in an in-frame insertion of Gln–Lys after amino acid 218 of YaeT, in a region predicted to be part of a large periplasmic domain that encompasses the first 482 amino acids (10). YaeT, a homologue of Omp85 (10) that is essential for the assembly of OMPs (4), contains five periplasmic POTRA (for polypeptide-transport-associated) domains that have a predicted chaperone-like function (Figure 1) (11). POTRA domains, which are conserved among the Omp85 family members, are composed of three β -strands and two α -helices (11). The two additional amino acids encoded by the *yaeT6* allele are located three residues upstream of the first predicted α -helix of the third POTRA domain of YaeT, suggesting that they might alter YaeT function in OMP assembly.

YaeT is part of an OM multiprotein complex that contains the three OM lipoproteins, YfgL, YfiO, and NlpB (Figure 1) (4). Therefore, chemical conditionality identified two factors, YaeT and YfgL, which are involved in OM biogenesis and are part of the same multiprotein complex. Despite this, there is an important difference between the suppressor quality of the *yfgL* and *yaeT6* alleles that we isolated using chemical conditionality (3). While both *yaeT6* and *yfgL* alleles suppress sensitivity of *imp4213* cells to MacConkey agar, only *yfgL* mutations suppress sensitivity to chlorobiphenyl vancomycin (8).

The Remarkable Chemical Specificity of the *yaeT6*.

Previously, we showed that *yfgL* mutations increase *imp4213* resistance to chlorobiphenyl vancomycin but not to vancomycin (8), because this suppressor mutation decreases the permeability of the OM to the former compound but not the latter (3). We tested whether the *yaeT6* allele would also exhibit such remarkable specificity for small molecules. Since *yaeT6* confers resistance to MacConkey agar and this medium is a mixture of bile salts, we compared the effects of *yaeT6* on the susceptibility of *imp4213* strains to a variety of purified bile salts.

MacConkey agar (12) contains two types of bile salts, sodium cholate ($3\alpha,7\alpha,12\alpha$ -trihydroxy-5 β -cholanolic acid) and sodium deoxycholate ($3\alpha,12\alpha$ -dihydroxy-5 β -cholanolic acid), which differ by the presence of a single hydroxyl group (Figure 2, panel a). Wild-type strains of *E. coli* are resistant to bile salts and, therefore, can grow on MacConkey agar. In contrast, cells carrying the *imp4213* allele cannot grow in this medium (5).

Strains carrying *imp4213* are sensitive to sodium deoxycholate (5), but we did not know how this mutation affects growth in the presence of sodium cholate. Therefore, we first examined whether *imp4213* cells are also sensitive to sodium cholate by determining the minimal inhibitory concentration (MIC) values of purified sodium cholate or sodium deoxycholate for *imp⁺* and *imp4213* strains in Luria-Bertani broth at 37 °C. As expected, the wild-type *imp⁺* strain was resistant to high levels of both sodium cholate and sodium deoxycholate, while the *imp4213* mutant was especially sensitive to sodium deoxycholate (Table 1). The MIC difference between *imp⁺* and *imp4213* strains for sodium deoxycholate was >200. The *imp4213* allele also increased sensitivity to sodium cholate, but only by a factor of 8 compared with the *imp⁺* strain (Table 1).

To determine whether the pronounced increase in sensitivity to sodium deoxycholate was specific to this bile salt or whether it reflected a general increase in permeability for bile salts that have similar physical properties, we tested the sensitivity of *imp4213* cells to a regioisomer of sodium deoxycholate, sodium chenodeoxycholate ($3\alpha,7\alpha$ -dihydroxy-5 β -cholanolic acid; Figure 2, panel a). Both sodium deoxycholate and sodium chenodeoxycholate differ from sodium cholate in that they only have two hydroxyl groups; however, the location of one of these hydroxyl groups differs between the two isomers (Figure 2, panel a). As a result, sodium

TABLE 1. Sensitivity to bile salts of wild-type and *imp4213* strains

Strain	MIC (mg mL ⁻¹)		
	Cholate	Deoxycholate	Chenodeoxycholate
<i>imp⁺</i>	67	>67	>67
<i>imp4213</i>	8.3	0.3	0.5
<i>imp4213 yfgL8</i>	8.3	4.2	33.35
<i>imp4213 yaeT6</i>	8.3	8.3	16.7

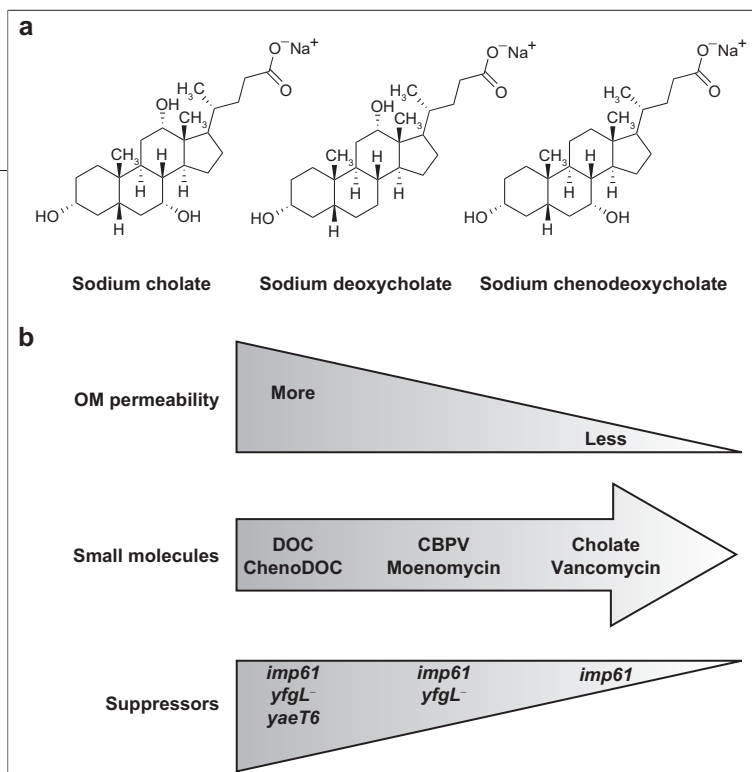


Figure 2. Toxic small molecules identify suppressors of *imp4213* with various degrees of OM permeability. **a**) Chemical structures of bile salts used in this study. **b**) Suppressors of *imp4213* can be arranged in a continuum (bottom triangle) according to the degree of permeability of their OM (top triangle). OM permeability can be probed with the toxic small molecules indicated (center arrow) as described in Results and Discussion. DOC and CBPV denote deoxycholate and chlorobiphenyl vancomycin, respectively.

cholate has a lower partition coefficient ($X \text{ Log } P$: 2.818) (13) than dihydroxy bile salts do ($X \text{ Log } P$: 3.992) (13). We found that in *imp4213* cells the MIC was marginally higher for sodium chenodeoxycholate than for sodium deoxycholate, but it was still >100-fold lower than in wild-type (Table 1). Thus, the *imp4213* allele increases sensitivity to all three bile salts tested, but it renders cells 15- and 20-fold more sensitive to sodium chenodeoxycholate and sodium deoxycholate, respectively, than to sodium cholate.

We next determined how the MIC values for *imp4213* cells changed in the presence of suppressor mutations that we obtained by selecting for growth on MacConkey plates. Both suppressor strains *imp4213 yfgL8* and *imp4213 yaeT6* exhibited the same sensitivity to sodium cholate as their parent *imp4213* strain, but they were much more resistant (>14-fold) to both sodium deoxycholate and sodium chenodeoxycholate (Table 1). Thus, both the *yaeT6* and *yfgL* suppressor mutations exhibit extraordinary specificity for dihydroxy bile salts compared with the trihydroxy bile salt sodium cholate (Figure 2, panel a).

Previously, we arranged toxic small molecules in a continuum that corresponds to the degree to which

membrane integrity had been restored by the suppressor mutation (3). The unrefined mixture of “bile salts” present in MacConkey agar was positioned to the left of the continuum, as resistance to this mixture appeared to demand the lowest level of restoration. Resistance to vancomycin, which requires the highest level of restoration, was positioned on the right. Here, we have refined and expanded this continuum (Figure 2, panel b). The use of purified bile salts has revealed that, while sodium deoxycholate and sodium chenodeoxycholate are indeed located on the leftmost end of the continuum, sodium cholate is actually located on the rightmost end together with vancomycin. The only mutations that increase resistance to compounds located at the rightmost end are intragenic suppressors in *imp* (data not shown) (3).

***yaeT6* Does Not Alter the Levels of YfgL in the YaeT Complex.** Since YaeT is an essential protein in *E. coli* (4, 14, 15), *yaeT6* cannot be a complete loss-of-function allele. However, it can be either a partial loss-of-function or a gain-of-function allele. These possibilities can be distinguished by diploid analysis. Complementation tests were performed using an *imp4213* strain that carried *yaeT6* at the native *yaeT* chromosomal locus and a wild-type *yaeT* allele expressed under the control of an inducible promoter (P_{BAD}) *in trans* from either a multicopy number plasmid (pBAD18-*yaeT*) or a single-copy locus near the *att* site ($\Delta\text{latt-lom}$::*blaP_{BAD} yaeT araC*) (4). In both cases, induction of wild-type *yaeT* with arabinose restored sensitivity to MacConkey agar of the *imp4213* strain, demonstrating that the *yaeT6* allele is recessive, as the phenotypes conferred by this mutant allele cannot be observed when the mutant gene and a wild-type *yaeT* gene are co-expressed. Thus, since *yaeT6* is recessive and *yaeT* is essential, the *yaeT6* allele cannot be a null (*i.e.* total loss-of-function) but rather a partial loss-of-function allele.

A partial loss-of-function allele could result from decreased levels of YaeT, if the two-amino acid insertion decreased the stability, targeting, or folding efficiency of YaeT6. However, immunoblot analysis revealed that the steady levels of YaeT6 are similar to those of wild-type YaeT (data not shown). Since *yaeT6* does not change levels of YaeT, it is likely to affect its activity and/or interaction with other proteins. Because YaeT exists in a complex with YfgL and loss-of-function mutations in *yfgL* suppress certain *imp4213* defects (3, 4), it is possible that *yaeT6* could suppress by altering the YaeT–YfgL

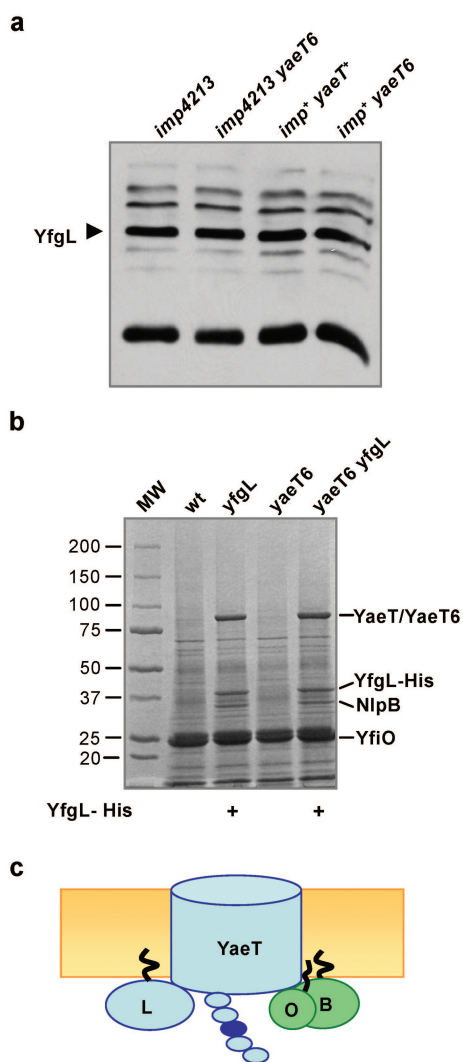


Figure 3. The *yaeT6* allele does not alter YfgL levels nor the composition of the YaeT–lipoprotein complex. **a)** Steady-state levels of YfgL in cells were analyzed by immunoblot using YfgL antiserum. The levels of YfgL remain unchanged in the presence of the *yaeT6* allele in both *imp4213* and wild-type *imp* backgrounds. The strains used from left to right are NR698, NR810, NR842, and NR850, and their respective relevant genotypes are shown above the lanes. **b)** Ni-affinity purification of His-tagged YfgL shows that YaeT6 does not change the composition of the YaeT–lipoprotein complex. The strains used from left to right are NR842, NR847 (pTW006), NR850, and NR853 (pTW006); their relevant genotypes are shown above the lanes, and the presence of the pTW006 plasmid is indicated with a “+” below the lanes. The bands corresponding to the members of the YaeT–lipoprotein complex are labeled. The size of the molecular markers (MW lane) is shown in kDa. **c)** The composition of the YaeT complex is not changed in the presence of YaeT6 and the His-tagged YfgL lipoprotein. The third POTRA domain of YaeT is shown in dark blue to represent the change caused by the *yaeT6* mutation.

ciation of YfgL from the YaeT–lipoprotein complex or by alterations in the composition of this complex.

***yaeT6* Does Not Render YfgL Nonfunctional.** Although the levels of YfgL in the YaeT complex are not changed by the *yaeT6* suppressor, it is possible that this mutation renders YfgL nonfunctional. Alternatively, *yaeT6* could be suppressing in an YfgL-independent manner. Since the loss of YfgL function leads to decreased OMP levels (3, 4, 16), if *yaeT6* rendered YfgL nonfunctional, we would expect that the *yaeT6* allele would cause the same phenotypes as loss-of-function alleles of *yfgL*. To assess the effect of *yaeT6* on OMP levels, we assayed the levels of two OmpAs, OmpA and LamB, by immunoblot using the levels of the periplasmic maltosebinding protein (MBP) as a loading standard. We found that the *yaeT6* allele reduced the levels of LamB and OmpA reproducibly by ~10% in *imp4213* cells (Figure 4, panels a and c). In contrast, the loss of YfgL function reduced LamB levels 33% and OmpA 14% in *imp4213* cells (3). Likewise, in an *imp+* background, the *yfgL* allele causes a greater reduction in OMP levels than *yaeT6*. While the presence of the *yaeT6* allele results in a 10% reduction in LamB levels, a *yfgL* null allele reduces them by 30% (Figure 4, panels b and c). These data show that the *yaeT6* allele does not render YfgL nonfunctional; therefore, suppression of *imp4213* by *yaeT6* does not involve the loss of YfgL function.

Synthetic Phenotypes of *yaeT6*. We previously reported two synthetic interactions between *imp* and

interaction so as to reduce the amount of this lipoprotein in the complex, which in turn might destabilize YfgL. We therefore monitored the levels of YfgL present at steady state and in the YaeT complex in *yaeT6* cells.

Immunoblot analysis demonstrated that the presence of the *yaeT6* allele does not change the cellular steady-state levels of YfgL when compared to those found in *yaeT+* cells in either the *imp+* or *imp4213* backgrounds (Figure 3, panel a). In addition, when His-tagged YfgL was purified in a Ni²⁺ column, the amount of YaeT, NlpB, and YfiO that co-purified was the same in *yaeT+* and *yaeT6* cells (Figure 3, panel b). Therefore, suppression by the *yaeT6* allele is not caused by the disso-

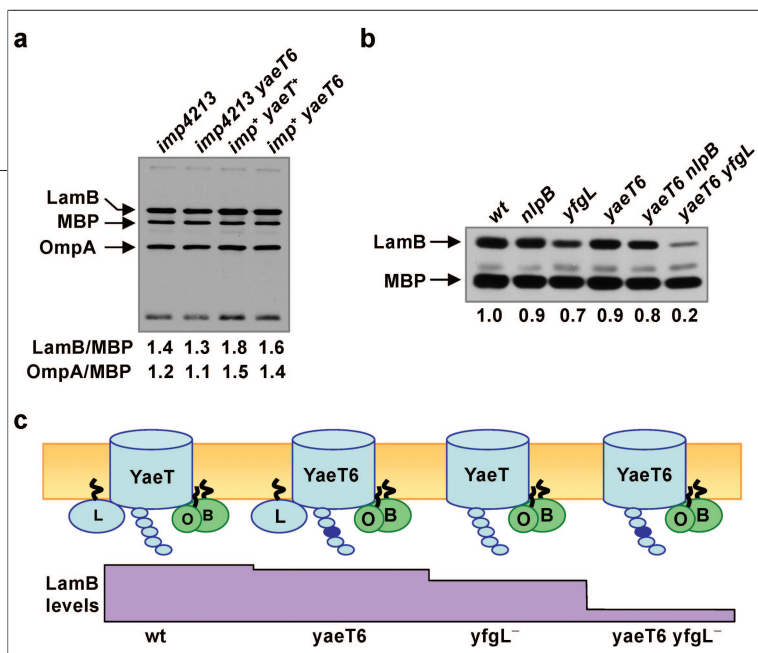


Figure 4. Effect of the *yaeT6* allele in OMP levels. a) Levels of OMPs LamB and OmpA and the periplasmic protein MBP were analyzed by immunoblot, and their relative levels are shown below the lanes. The presence of the *yaeT6* allele causes a slight reduction in OMPs in both *imp4213* and wild-type *imp* backgrounds. The strains used from left to right are NR698, NR810, NR842, and NR850, and their respective relevant genotypes are shown above the lanes. b) All single mutants of the YaeT complex cause a reduction in the OMP levels. The effects of the *yaeT6* and *nlpB* mutations appear additive. However, the *yaeT6* and *yfgL* alleles exhibit a synthetic phenotype, since the reduction in OMP levels in the double mutant is much greater than the sum of the effect of each individual allele. The strains used from left to right are NR842, NR846, NR847, NR850, NR851, and NR853, and their respective relevant genotypes are shown above the lanes. c) Scheme representing the composition of the YaeT complex and the relative levels of LamB in different mutant backgrounds (from data shown in panel b).

yfgL. First, *imp4213* and *yfgL* are a conditional synthetic lethal pair (*i.e.* under certain growth conditions, the double mutant is not viable). In this case, an *imp4213 yfgL* double mutant is viable only under slow growth conditions (3). Second, loss-of-function alleles of *yfgL* decrease the permeability of the OM of *imp4213* cells, but they increase it in *imp+* cells (3).

Unlike *yfgL* and *imp4213*, the *yaeT6* and *imp4213* alleles are not a synthetic lethal pair under any condition tested (Table 2); introduction of the *yaeT6* allele into *imp4213* cells can be easily accomplished under normal growth conditions, and it does not require additional mutations. Furthermore, the *imp4213 yaeT6* double mutant can grow as well as its parent *imp4213* in all conditions tested, as long as the growth medium does not contain small toxic molecules. Thus, while *yfgL* and *imp4213* are a synthetic conditional lethal pair, *yaeT6* and *imp4213* are not, corroborating that YfgL is functional in *yaeT6* cells. Moreover, these findings also indicate that slowed growth is not required for conferring resistance to bile salts to *imp4213* cells.

After replacing the *imp4213* allele with wild-type *imp*, we assayed the barrier function of the OM in a strain carrying only the *yaeT6* allele by determining its sensitivity to various antibiotics. The presence of the *yaeT6* allele increases the sensitivity (*i.e.*, permeability) to all compounds tested to levels similar to those of a strain carrying a *yfgL* null allele (Table 3). Thus, in an *imp+* strain, *yaeT6* compromises the integrity of the OM, which manifests as increased permeability to small molecules, but in *imp4213* cells, the *yaeT6* allele decreases the permeability of their OM. This synthetic interaction between *imp* and *yaeT6* is similar to that reported previously between *imp* and *yfgL* (3) (Table 2). These synthetic phenotypes reflect the fact that Imp and the YaeT–lipoprotein complex perform distinct but related functions in OM assembly (4, 6, 10).

We previously described synthetic growth phenotypes between pairs of alleles of *yfgL*, *nlpB*, and *yfiO* that supported the biochemical data showing that all three lipoproteins were part of a functional complex (4). In a similar vein, although cells carrying the *yaeT6* allele do not exhibit any growth defects, introduction of either a null allele in *yfgL* or *nlpB* or a partial loss-of-function mutant allele in *yfiO* led to double mutants that are to various extents mucoid, heterogeneous in size, and slow-growing. The least severe defects were observed in the *yaeT6 nlpB* mutant, while the most severe defects resulted in the *yaeT6 yfiO* strain. In fact, the growth defects in the *yaeT6 yfiO* strain were so severe that we could not work with this strain. This perhaps is a consequence of the fact that both YaeT and YfiO are essential proteins in *E. coli* (4, 16, 17).

We also found that the severity of the growth defects in the *yaeT6 nlpB* and *yaeT6 yfgL* double mutants correlated with the effects that these allele combinations have on OMP levels and membrane integrity. As described above, the *yaeT6* allele causes a mild reduction in LamB (Figure 4). This decrease in OMP levels was similar to that of an *nlpB* null allele (Figure 4, panel b), and in the *yaeT6 nlpB* double mutant, the effects of these mutations appeared additive (Figure 4, panel b). In contrast, the levels of LamB were greatly reduced in the *yaeT6 yfgL::kan* double mutant when compared to each single mutant (Figure 4, panels b and c). This change in LamB levels is not likely the result of decreased synthesis, because the levels of the MBP, which is another member of the maltose regulon, were not affected (Figure 4, panel b). Likewise, the presence

TABLE 2. Summary of phenotypes discussed in this study

Phenotypes ^a	<i>yaeT6</i>	<i>yfgL8</i>	<i>nlpB</i>	<i>yfiO::kan</i> ^b
Conditional synthetic lethality with <i>imp4213</i>	–	+	–	+
Suppressor of <i>imp4213</i>	+	++	–	–
Decreased OMP levels	+	++	+	++
Increased OM permeability in <i>imp</i> ⁺	+	++	±	++
Synthetic phenotypes with mutant alleles in <i>yaeT</i> , <i>yfgL</i> , <i>yfiO</i> , or <i>nlpB</i>	+	+	+	+

^aWhen appropriate, the relative strengths of phenotypes are indicated with “+” and “–” signs. As indicated in the text, some of these phenotypes are taken from previous studies (3, 4). ^b*yfiO::kan* is a partial loss-of-function allele (4).

of the *yaeT6* allele increased the defects in membrane integrity caused by *nlpB* and *yfgL* null alleles; again, the increase in sensitivity to various antibiotics appeared additive in the double mutants (Table 3). Thus, there is a correlation between growth defects, decrease in OMP levels, and membrane integrity. These defects are milder in the single mutants and, when the double mutants are compared, they are less severe in the *yaeT6 nlpB* double mutant than in the *yaeT6 yfgL* double mutant. This is in agreement with previous findings showing that mutations in *yfgL* confer stronger phenotypes than *nlpB* mutations (4).

All these data indicate that, although YaeT, YfgL, YfiO, and NlpB are part of a complex, these proteins are not functionally equivalent. Both YaeT and YfiO are essential, while YfgL and NlpB are not (4, 16). Combining

proper OM integrity requires an appropriate balance between LPS and OMP assembly (3, 4). This could explain why mutations in *yfgL* could decrease membrane permeability in *imp4213* cells but increase it in *imp*⁺ cells. In the first case, balance is restored, while in the latter, balance is offset in the opposite direction. At first glance, finding that *yaeT6* is a suppressor of *imp4213* appears to further support this model. However, a more detailed comparison between *yfgL* and *yaeT6* reveals that this model is not correct, and it provides insight into the mechanism of suppression.

The fact that a *yfgL*[–] allele causes a greater reduction in OMP levels than the one caused by *yaeT6* seems to correlate with the former allele being a stronger suppressor of *imp4213* than the latter (*i.e.*, only *yfgL* increases resistance to chlorobiphenyl vancomycin).

mutant alleles of *yaeT* and *yfiO* causes the most severe phenotypes, and the loss of NlpB causes milder phenotypes than the loss of YfgL. Moreover, although both *yaeT6* and *yfgL* suppress the sensitivity to MacConkey agar conferred by *imp4213*, these two alleles exhibit different interactions with *imp4213* and different degrees of suppression (Table 2).

Previously, we considered the possibility that

TABLE 3. Sensitivity of *imp*⁺ *yaeT6* strains to antibiotics

Strain	Zone of inhibition (mm) ^a			
	Bac	Nov	Em	Rif
<i>yadG::cam</i> ^b	<6	(8)	(9)	10
<i>yadG::cam yaeT6</i>	8	10	13	13
<i>yadG::cam yfgL::kan</i>	9	12	14	13
<i>yadG::cam nlpB::kan</i>	(8)	(9)	8	8 (11)
<i>yadG::cam yaeT6 yfgL::kan</i>	12 (21)	15	10	16 (22)
<i>yadG::cam yaeT6 nlpB::kan</i>	13	13	13 (15)	13 (21)

^aThe number in parenthesis indicate that the zone was not totally clear but hazy. Bac = bacitracin, Nov = novobiocin, Em = erythromycin, and Rif = rifampin. ^bThe *yadG* allele does not alter the sensitivity to toxic compounds. It is used as a marker linked to *yaeT6*.

The genes identified by chemical conditionality specify components of a multiprotein complex required for the assembly of OM β -barrel proteins.

However, mutant alleles in the genes for two other components of the YaeT–lipoprotein complex, NlpB and YfiO, also decrease OMPs to levels comparable to *yaeT6* and *yfgL*, respectively, yet they do not suppress any of the defects caused by *imp4213* (4). Thus, the reduction of OMP levels alone cannot explain suppression of *imp4213*. It is worth noting, however, that the severity of reduction in OMP levels caused by both suppressing and nonsuppressing mutations in the genes encoding the members of the YaeT–lipoprotein complex correlates strongly with increased permeability caused by these mutations in *imp*⁺ cells. In wild-type (*imp*⁺) cells, decreases in OMP levels alone can explain the increased permeability caused by defects in the YaeT–lipoprotein complex (see below).

Probing OM Composition through Chemical

Specificity. The genes identified by chemical conditionality, *yfgL* and *yaeT*, both specify components of a multiprotein complex required for the assembly of OM β -barrel proteins (4). Since loss of YfgL function suppresses *imp4213* (3), it is not surprising that a mutant form of another member of the YfgL complex, namely, YaeT6, can also suppress *imp4213* defects. What we do not understand is why defects in the machinery that assembles OMPs suppress a mutation, *imp4213*, which causes defects in LPS assembly.

Normally, the OM of *E. coli* is an asymmetric bilayer where the LPS resides in the outer leaflet and phospholipids in the inner leaflet (18). LPS and phospholipids do not mix (19, 20), so exclusion of phospholipids from the outer leaflet allows LPS molecules to become highly compacted, and this makes the OM an effective barrier against both hydrophobic and large hydrophilic compounds; small hydrophilic molecules (<600 Da) can enter the cell through porins (21, 22).

Alterations to LPS, caused either by mutation or the addition of chelators such as EDTA, allow phospholipids to flip into the outer leaflet of the OM (23–27). Because these two lipids do not mix, the misplaced phospholipids segregate into domains forming patches of phospholipid bilayer (19, 20, 25, 26). These phospholipid bilayer domains allow passage of toxic molecules across the OM in two ways (see Supplementary Figure 1). First, hydrophobic molecules can now enter the cell by diffusion through these patches. In addition, because the boundaries between LPS and phospholipids are disordered, transient “cracks” may allow the passage of not only hydrophobic molecules but also

large hydrophilic compounds such as vancomycin (28). Moreover, we suggest that hydrophobic and amphipathic molecules like chlorobiphenyl vancomycin may partition into these phospholipid patches, increasing their local concentration in the phospholipid bilayer and, therefore, also near the cracks. In fact, increasing the hydrophobicity of vancomycin derivatives by the addition of hydrophobic substituents, such as chlorobiphenyl groups, increases their binding to phospholipid bilayers (29). We propose that the increase in local concentration of these molecules likely increases their passage through the cracks.

There is general agreement that the increased permeability caused by *imp4213* is the result of the presence of phospholipids in the outer leaflet (see Supplementary Figure 1) (3, 28). A model has been proposed to explain the reason *yfgL* suppresses *imp4213*, conferring resistance to chlorobiphenyl vancomycin but not vancomycin (28). According to this model, both molecules enter *imp4213* cells through the transient cracks in the OM. However, because of its aromatic, hydrophobic component, chlorobiphenyl vancomycin enters as micelles, while the hydrophilic vancomycin enters as a soluble monomer. The lowering of OMP levels in the OM of *imp4213* cells caused by *yfgL* mutations results in narrower OM cracks that exclude the large chlorobiphenyl vancomycin micelles but not vancomycin.

We agree that the basis for the differential suppression of chlorobiphenyl vancomycin but not vancomycin sensitivity by the *yfgL* mutation reflects the difference in the hydrophobicity of these two compounds. However, we disagree with the micelle model. Although strongly affected by solution components, the concentration of chlorobiphenyl vancomycin that we use is at least 10-fold below the concentration at which it dimerizes (30). More importantly, this model cannot explain our results using purified bile salts.

Cells carrying *imp4213* are more sensitive to dihydroxy bile salts, such as sodium deoxycholate and sodium chenodeoxycholate, than to their trihydroxy counterpart, sodium cholate. Both the *yfgL*[−] and the *yaeT6* suppressors increase resistance to the dihydroxy bile salts without altering the susceptibility to the trihydroxy bile salt. The aggregation number, which reflects the size of micelles, is lower for sodium chenodeoxycholate than for sodium cholate (at concentrations >8 mg mL^{−1}), but higher for sodium deoxycholate (31). If the relative size of micelles and of OM cracks deter-

mined resistance, as proposed in the micelle model (28), then we would not predict that both sodium deoxycholate and sodium chenodeoxycholate behave similarly, yet they do.

We argue that the most relevant difference between the dihydroxy bile salts and sodium cholate is hydrophobicity. The presence of an additional hydroxyl group in sodium cholate results in a lower partition coefficient ($X \text{ Log } P$: 2.818) than for sodium deoxycholate and sodium chenodeoxycholate ($X \text{ Log } P$: 3.992) (13). As a result, the dihydroxy bile salts partition to membranes more efficiently than sodium cholate. Moreover, *in vitro*, the rate of bilayer flip-flop is much higher for the dihydroxy bile salts than for sodium cholate (32). We believe that these differences form the basis for suppressor specificity.

We suggest that both the *yfgL*⁻ and *yaeT6* mutations cause a reduction in the phospholipid content of the outer leaflet of *imp4213* cells, albeit to different levels (see Supplementary Figure 1). This reduction would affect sensitivity to molecules differentially according to their hydrophobicity. Reducing the phospholipid content in the outer leaflet reduces the surface area of phospholipid bilayer patches and, thereby, the local concentration at the cell surface of hydrophobic and amphipathic compounds that partition to these patches. Consequently, this reduction decreases the diffusion of small hydrophobic compounds, such as the dihydroxy bile salts through phospholipids patches, and of amphipathic compounds such as chlorobiphenyl vancomycin, through phospholipids patches or membrane cracks. This decrease in the size of phospholipids patches would have a more modest effect on the sensitivity to hydrophilic molecules like vancomycin. Reductions in the amount of outer leaflet phospholipids will cause a larger decrease in surface area of the patches than in their perimeter (cracks). This may explain why decreasing phospholipid content in the outer leaflet causes a more dramatic decrease in dihydroxy bile salt sensitivity than in chlorobiphenyl vancomycin sensitivity. In other words, we propose that the continuum of suppressors is simply a reflection of the phospholipid

content of the outer leaflet, which affects penetration of hydrophobic and amphipathic molecules through the OM. Since these strains still have a permeable OM, we cannot accurately measure the phospholipid content in the outer leaflet of the OM, as the probes used routinely to measure phospholipids can enter the periplasmic space and thereby also label the inner membrane phospholipids.

We think that phospholipid content of the outer leaflet also explains the permeability defects caused by mutations in the genes specifying the YaeT–lipoprotein complex in *imp*⁺ strains. As noted above, there is a direct correlation between the increase in permeability and the defect in OMP assembly caused by these mutations. Since phospholipids fill the void in the outer leaflet created by the lack of OMPs in the OM (26), the greater the defect, the higher the amount of phospholipid bilayer in the OM.

How the *yfgL*⁻ and *yaeT6* mutations may reduce the phospholipid content of the outer leaflet of *imp4213* cells is not known. In fact, we know almost nothing about the mechanism(s) of phospholipid transport to, and assembly, in the OM. It is possible that the *yfgL*⁻ and *yaeT6* mutations suppress *imp4213* by directly affecting this mechanism(s). Although we do not understand how the YaeT complex assembles OMPs in the OM, one of the models proposes that YaeT is a scaffold for OMPs insertion into the OM at the YaeT–lipid interface (33). This model requires rearrangement of OM lipids for OMP insertion as does the “molten discs” model for *in vitro* insertion of OMPs in lipid bilayers (34). Therefore, it is plausible that the YaeT complex could coordinate OM lipid rearrangement to mediate OMP insertion in the OM, and this could explain why the *yfgL*⁻ and *yaeT6* alleles affect phospholipid levels in the outer leaflet of the OM. Alternatively, these suppressor alleles could decrease the amount of phospholipids in the outer leaflet indirectly, for example, by inducing a stress response or by increasing the LPS-assembly activity of Imp4213. Until we fully understand how Imp, YaeT, and YfgL function in OM biogenesis, we cannot distinguish between these models.

METHODS

Bacterial Strains and Growth Conditions. Strains used are listed (see Supplementary Table 1). Luria-Bertani (LB) broth and agar, lactose MacConkey agar, and M63 minimal agar were pre-

pared as described previously (35). All liquid cultures were grown under aeration at 37 °C, and their growth was monitored by measuring the optical density at 600 nm (OD_{600}). When needed, tetracycline and kanamycin were used at a concentra-

tion of 25 $\mu\text{g mL}^{-1}$. Chloramphenicol and ampicillin were used at a concentration of 15 and 10 $\mu\text{g mL}^{-1}$, respectively, for strains carrying the *imp4213* allele, and of 20 and 125 $\mu\text{g mL}^{-1}$, respectively, for those carrying the *imp* wild-type allele. When necessary, strains were converted to *ara*⁺ by selecting on arabinose minimal media. In complementation tests, expression of *yaeT* was accomplished by using arabinose MacConkey plates.

Bile Salt Resistance Selection and Genetic Mapping. Spontaneous bile salt-resistant mutants were obtained by selecting for growth of NR701 on lactose MacConkey agar (3). P1 transduction was used for strain construction and genetic mapping as described previously (35). To map the *yaeT6* allele, we used bacteriophage P1 mapping as follows. We generated a P1 lysate from a pool of mutants carrying randomly inserted mini-Tn*cam* cassettes in the chromosome of our wild-type strain MC4100 (36). This pool of mutants was used as a donor of wild-type alleles in P1 transductions where the recipient was the bile salt-resistant *imp4213* suppressor strain. We then screened for transductants that had lost the ability to grow on MacConkey agar that was conferred by the suppressor mutation and found that a *yadG::cam* allele was ~20% linked to the suppressor mutation. We demonstrated that the suppressor mutation linked to *yadG* is sufficient to confer bile salt resistance to *imp4213* cells, since we could confer bile salt resistance when we backcrossed the suppressor mutation-linked *yadG::cam* insertion (donor) with a strain only carrying the *imp4213* allele (recipient).

Detailed mapping using additional markers in the *yadG* region showed that the suppressor mutation mapped to minute 4.2 of the chromosome. DNA sequence of this region revealed that the suppressor mutation is a 6-bp insertion (agaaac) between codons 218 and 219 of *yaeT*, so the suppressor allele was named *yaeT6*.

Sensitivity to Toxic Compounds. Sensitivity to bile salts was assayed by the ability or inability to grow onto lactose MacConkey agar. Sensitivity to bacitracin, novobiocin, erythromycin, and rifampin was measured using BBL Sensi-Discs Antimicrobial Susceptibility Test Discs (BBL) as previously described (3) and reported as the diameter (in mm) of the zone of inhibition of growth around each 6-mm disc. Data shown is representative of three independent experiments. Sensitivity to chlorobiphenyl vancomycin and vancomycin of strains carrying the *imp4213* allele was determined by assessing growth on LB agar containing either compound at a final concentration of 4 $\mu\text{g mL}^{-1}$ for chlorobiphenyl vancomycin or 2 $\mu\text{g mL}^{-1}$ for vancomycin.

To measure the MICs to purified bile salts, all sodium cholate, sodium deoxycholate, and sodium chenodeoxycholate (Sigma Aldrich) were dissolved into water and added to the appropriate wells of a 96-well plate containing a 1:1000 dilution of overnight culture in LB broth. After 1:2 serial dilutions were performed, the 96-well plate was incubated overnight at 37 °C. Growth was determined by OD₆₀₀. Data shown are representative of three independent experiments.

Immunoblot Analysis. One-milliliter samples were taken from overnight LB broth cultures. For standardization, samples were pelleted (16,000g for 5 min) and resuspended in a volume (in mL) of SDS sample buffer equal to OD₆₀₀/10. Prior to electrophoresis, samples were boiled for 10 min, and equal volumes were loaded onto 10% polyacrylamide (w/v) gels containing SDS (37). Proteins were transferred to nitrocellulose membranes (Schleicher & Schuell), and Western blot analysis was performed as previously described (3). When appropriate, polyclonal sera raised against YfgL, LamB, and MBP (from our laboratory collection) were used as primary antibodies at a dilution of 1:7000; 1:10,000; and 1:20,000; respectively. These antisera also recognize OmpA. Donkey anti-rabbit IgG horseradish peroxidase conjugate (Amersham Pharmacia Biotech) was used as secondary antibody at a 1:8000 dilution. For visualization of bands, the

ECL antibody detection kit (Amersham Pharmacia Biotech) and Biomax film (Kodak) were used. When indicated, ImageJ 1.34s software was used to quantify the relative intensity of the bands.

Purification of the YaeT Complex. Cells were grown in 0.5 L of LB broth to OD₆₀₀ ~0.6 and harvested by centrifugation at 5000g for 10 min. Cells were lysed in 5 mL of BugBuster reagent (Novagen) in the presence of lysozyme (100 $\mu\text{g mL}^{-1}$) and DNase I (50 $\mu\text{g mL}^{-1}$) by shaking at RT for 20 min. After ultracentrifugation at 100,000g for 30 min (Beckman SL7), the supernatant was transferred to a new test tube, and imidazole was added to 20 mM. The entire cell lysate was loaded twice onto a column packed with 0.5 mL of NTA-Ni resin (Qiagen) that had been equilibrated with 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.1% Triton X-100 (v/v), and 20 mM imidazole. The column was washed with 10 mL of equilibration buffer and eluted with 5 mL of 20 mM Tris-HCl, pH 7.4, and 200 mM imidazole. The eluted sample was concentrated in an Amicon Ultra device (Millipore) by centrifugation at 5000g for 30 min. Ten microliters of the concentrated sample was used for SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and proteins were visualized using Coomassie blue stain. Their identity was confirmed either by immunoblot analysis or mass spectrophotometry.

Acknowledgment: We thank the Kahne and Silhavy laboratory members, especially J. Malinverni, for their helpful discussions. This work was supported by Grants GM66174 (D.K.) and GM34821 (T.J.S.) from the National Institute of General Medical Sciences.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited, *Microbiol. Mol. Biol. Rev.* 67, 593–656.
- Ruiz, N., Kahne, D., and Silhavy, T. J. (2006) Advances in understanding bacterial outer-membrane biogenesis, *Nat. Rev. Microbiol.* 4, 57–66.
- Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T. J. (2005) Chemical conditionality: a genetic strategy to probe organelle assembly, *Cell* 121, 307–317.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., and Kahne, D. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*, *Cell* 121, 235–245.
- Sampson, B. A., Misra, R., and Benson, S. A. (1989) Identification and characterization of a new gene of *Escherichia coli* K-12 involved in outer membrane permeability, *Genetics* 122, 491–501.
- Bos, M. P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004) Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface, *Proc. Natl. Acad. Sci. U.S.A.* 101, 9417–9422.
- Braun, M., and Silhavy, T. J. (2002) Imp/OstA is required for cell envelope biogenesis in *Escherichia coli*, *Mol. Microbiol.* 45, 1289–12302.
- Eggert, U. S., Ruiz, N., Falcone, B. V., Branstrom, A. A., Goldman, R. C., Silhavy, T. J., and Kahne, D. (2001) Genetic basis for activity differences between vancomycin and glycolipid derivatives of vancomycin, *Science* 294, 361–364.
- Levinson, G., and Gutman, G. A. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution, *Mol. Biol. Evol.* 4, 203–221.
- Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., and Tommassen, J. (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly, *Science* 299, 262–265.
- Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A. (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins, *Trends Biochem. Sci.* 28, 523–526.

12. MacConkey, A. T. (1905) Lactose-fermenting bacteria in faeces, *J. Hyg.* **5**, 333–379.
13. Wang, R., Fu, Y., and Lai, L. (1997) A new atom-additive method for calculating partition coefficients, *J. Chem. Inf. Comput. Sci.* **37**, 615–621.
14. Werner, J., and Misra, R. (2005) YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*, *Mol. Microbiol.* **57**, 1450–1459.
15. Doerfler, W. T., and Raetz, C. R. (2005) Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant, *J. Biol. Chem.* **280**, 27679–27687.
16. Onufiyk, C., Crouch, M. L., Fang, F. C., and Gross, C. A. (2005) Characterization of six lipoproteins in the σ^E regulon, *J. Bacteriol.* **187**, 4552–4561.
17. Malinverni, J., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R., and Silhavy, T. J. (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*, *Mol. Microbiol.*, **61**, 151–164.
18. Kamio, Y., and Nikaïdo, H. (1976) Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium, *Biochemistry* **15**, 2561–2570.
19. Lasch, P., Schultz, C. P., and Naumann, D. (1998) The influence of poly-(L-lysine) and porin on the domain structure of mixed vesicles composed of lipopolysaccharide and phospholipid: an infrared spectroscopic study, *Biophys. J.* **75**, 840–852.
20. Takeuchi, Y., and Nikaïdo, H. (1981) Persistence of segregated phospholipid domains in phospholipid–lipopolysaccharide mixed bilayers: studies with spin-labeled phospholipids, *Biochemistry* **20**, 523–529.
21. Nakae, T., and Nikaïdo, H. (1975) Outer membrane as a diffusion barrier in *Salmonella typhimurium*. Penetration of oligo- and polysaccharides into isolated outer membrane vesicles and cells with degraded peptidoglycan layer, *J. Biol. Chem.* **250**, 7359–7365.
22. Decad, G. M., and Nikaïdo, H. (1976) Outer membrane of gram-negative bacteria. XII. Molecular-sieving function of cell wall, *J. Bacteriol.* **128**, 325–336.
23. Leive, L. (1965) Release of lipopolysaccharide by EDTA treatment of *E. coli*, *Biochem. Biophys. Res. Commun.* **21**, 290–296.
24. Leive, L. (1974) The barrier function of the gram-negative envelope, *Ann. N.Y. Acad. Sci.* **235**, 109–129.
25. Nikaïdo, H., and Vaara, M. (1985) Molecular basis of bacterial outer membrane permeability, *Microbiol. Rev.* **49**, 1–32.
26. Vaara, M. (1993) Antibiotic-supersusceptible mutants of *Escherichia coli* and *Salmonella typhimurium*, *Antimicrob. Agents Chemother.* **37**, 2255–2260.
27. Jia, W., Zoelby, A. E., Petruzzello, T. N., Jayabalasingham, B., Seyedirashti, S., and Bishop, R. E. (2004) Lipid trafficking controls endotoxin acylation in outer membranes of *Escherichia coli*, *J. Biol. Chem.* **279**, 44966–44975.
28. Nikaïdo, H. (2005) Restoring permeability barrier function to outer membrane, *Chem. Biol.* **12**, 507–509.
29. Allen, N. E., and LeTourneau, D. L. (1997) The role of hydrophobic side chains as determinants of antibacterial activity of semisynthetic glycopeptide antibiotics, *J. Antibiot.* **50**, 677–684.
30. LeTourneau, D. L., and Allen, N. E. (1997) Use of capillary electrophoresis to measure dimerization of glycopeptide antibiotics, *Anal. Biochem.* **246**, 62–66.
31. Ninomiya, R., Matsuoka, K., and Moroi, Y. (2003) Micelle formation of sodium chenodeoxycholate and solubilization into the micelles: comparison with other unconjugated bile salts, *Biochim. Biophys. Acta* **1634**, 116–125.
32. Kamp, F., Hamilton, J. A., and Westerhoff, H. V. (1993) Movement of fatty acids, fatty acid analogs, and bile acids across phospholipid bilayers, *Biochemistry* **32**, 11074–11085.
33. Gentle, I. E., Burri, L., and Lithgow, T. (2005) Molecular architecture and function of the Omp85 family of proteins, *Mol. Microbiol.* **58**, 1216–1225.
34. Tamm, L. K., Hong, H., and Liang, B. (2004) Folding and assembly of beta-barrel membrane proteins, *Biochim. Biophys. Acta* **1666**, 250–263.
35. Silhavy, T. J., Berman, M. L., and Enquist, L. W. (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
36. Kleckner, N., Bender, J., and Gottesman, S. (1991) Uses of transposons with emphasis on Tn10, *Methods Enzymol.* **204**, 139–180.
37. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* **227**, 680–685.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stiltz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Teaching a Chemical Biologist

The rising interest in interdisciplinary sciences such as chemical biology has generated a unique set of challenges for educators. How do we stimulate interest in multiple areas of science without overwhelming the students? Do we start teaching students chemical biology when they are undergraduates or wait until graduate school? Which science—chemistry or biology—should serve as the foundation for more in-depth studies in a graduate program? How do we take students with diverse scientific training and help them appreciate both the chemistry and the biology of cellular processes? Institutions around the world have met these challenges by adding new courses, revamping existing ones, or completely changing their undergraduate and graduate curricula. Starting in this issue of *ACS Chemical Biology* and continuing next month, we place chemical biology education In Focus in a series of commentaries from scientists who are successfully changing how we train the next generation of students. We expect that these articles will provide insight, guidance, and fresh ideas for other educators who are developing chemical biology courses or programs at their institutions.

We begin our In Focus series with a commentary on the Chemical Biology for Sophomores! lecture course and lab. This course, funded by a Howard Hughes Medical Institute Professors Award to Alanna Schepartz (Yale University), began in 2004. Joshua Kritzer, who helped teach the course with Schepartz, notes that “training the next generation of chemical biologists thus requires an undergraduate education that values a wide exposure to a variety of interrelated fields, from cancer evolution to robotics, from immunology to image analysis.” To do this, the Yale course relies heavily on inquiry-based instruction, a teaching technique in which students construct their own knowledge by solving problems in an open-ended, self-directed manner rather than through a prescribed set of exercises for which the outcome is known. The Yale lecture course and lab use many inquiry-based techniques, such as conducting in-class discussions on specific research papers and having each student write a review on a topic of interest to him or her. Kritzer notes in his commentary (p 411) that integrating inquiry-based techniques in the lecture course and lab allows the professor to teach the fundamental concepts in chemistry and biology that the students can use to explore the innovative areas of chemical biology.

The University of Dortmund in Germany, founded as an engineering school, is taking a different and bold approach to chemical biology education by restructuring its B.Sc., M.Sc., and Ph.D. curricula. These changes have transformed the traditional German *Diplom* in order to promote student exchange within the international community of chemical biologists. The university has capitalized on its close ties and collaborations with the Max Planck Institute for Molecular Physiology to facilitate the integration of chemistry and biology into one set of courses. As Arndt, Niemeyer, and Waldmann discuss in their commentary (p 407), “In essence, any program like ours must make sure that true benefits are offered, both for the undergraduate and graduate students and for the faculty.” With this in mind, the Dortmund B.Sc. curriculum provides a solid foundation in chemistry that includes basic molecular and cellular biology as well as bioorganic chemistry. The M.Sc. program prepares the students for job opportunities or for continued doctoral studies by including courses with more problem-solving exercises and student presentations. Students who wish to pursue a Ph.D. can apply to the new International Max Planck Research School in Chemical Biology, whose objective

is “to advance the application of chemistry to biochemistry and cellular biology on the graduate level by fostering emerging scientists from biochemistry, biophysics, and chemistry.” Arndt and colleagues also discuss how they overcame the practical and intellectual pitfalls that they faced when they redesigned the curriculum, and highlight how these extensive changes have benefited the students and faculty.

These are just two examples of the wide array of innovative choices we can make to effectively teach chemical biology. Successful programs find ways to appeal to students from diverse scientific backgrounds, providing them with ample opportunity to tailor their education to their particular interests without compromising their basic understanding of chemistry and biology. The challenges are unique to each program but can be resolved with creative pedagogical changes. Successful solutions for one institution can be adapted and shaped to fit another program's set of challenges.

Next month, we will focus on other programs (the University of Michigan, Vanderbilt University, the University of Wisconsin at Madison, McGill University in Canada, Keio University in Japan, and Tokyo Medical and Dental University in Japan) that have revamped their graduate curricula to meet their needs. We recognize that many universities are offering a wide range of diverse undergraduate and graduate programs in chemical biology (see our WIKI for a growing list). We invite you to write an In Focus piece: describe what approaches your institution has taken to teach and engage students interested in multidisciplinary science. We also invite you to participate in our Wiki discussion on education. We look forward to publishing more pieces in this series about training the next generation of chemical biologists.



Evelyn Jabri
Executive Editor

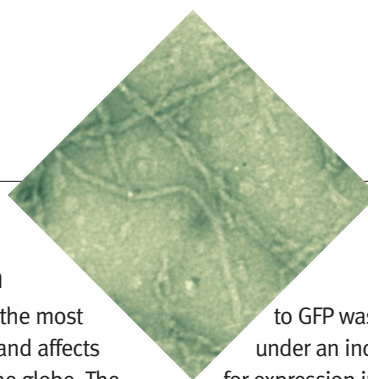
ACS Chemical Biology Strives to Serve Its Authors and Readers

At *ACS Chemical Biology*, we continue to improve the value of our print and online journal to our authors, readers, and the entire chemical biology community. We strive to offer features and functionalities to our authors to highlight their work and share it with a broader audience. Commencing with this issue, we will introduce a new section in the print journal and offer enhanced links to related content on our website.

In our ongoing efforts to promote junior members of the chemical biology community, we have developed a new section called ‘Introducing our Authors’. This new part of the journal will appear next to the ‘In this Issue’ section and will highlight the lead author of the manuscripts published in *ACS Chemical Biology*. We provide a brief background of the scientists and their personal perspectives and insights into what it is like to work on their projects. Our readers will now be able to put a face with a name and meet new members of the chemical biology community.

We're always looking for improved ways to alert our readers to relevant and interesting content. In addition to the notes in our Table of Contents, we now list related documents at the bottom of our web pages. We expect these links to enhance the discoverability of our content and ensure that your manuscripts are widely read and cited.

In this ISSUE



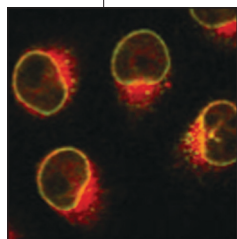
Arrested Development

Mitosis is a highly complex process in which a cell duplicates its genome and divides it between two daughter cells. Several small-molecule inhibitors of mitosis exist; in addition to their potential as anticancer agents, they have been valuable molecular tools to help scientists understand this complicated process. However, small-molecule inhibitors for many biological activities essential for mitosis have not been identified. Using a chemical genetics screen, Rundle *et al.* (p 443) discover a plant metabolite that causes prolonged mitotic arrest with an unusual phenotype.

The compound, *ent*-15-oxokaurenoic acid (EKA), is a diterpenoid that when incubated with cells halts mitosis at a stage resembling

prometaphase, where condensed chromosomes are scattered across abnormally formed mitotic spindles. The authors determined that EKA blocks chromosome movement by preventing association of centromeric protein E, a kinesin-like motor protein involved

in chromosome congression, with kinetochores, the protein structures that link chromosomes to microtubules from the mitotic spindle. In an effort to understand how EKA treatment brings about this phenotype, the researchers synthesized a cell-permeable, biotinylated EKA derivative and used it in affinity chromatography experiments to identify proteins that interact with EKA. Six EKA-binding proteins were isolated, including RanBP2, a protein whose depletion by small interfering RNA causes cells to arrest at prometaphase. Immunofluorescence experiments revealed that a substantial fraction of EKA colocalizes with RanBP2 at the nuclear envelope in interphase cells and at the base of mitotic spindles in mitotic cells. The similar phenotypes elicited by EKA treatment and RanBP2 depletion strongly suggest that EKA causes mitotic arrest by inhibiting RanBP2 function.



Inhibiting Aggregation

Alzheimer's disease (AD) is the most common form of dementia and affects millions of people around the globe. The molecular basis for the disease is under intense investigation, and substantial evidence suggests that aggregation of the peptide A β 42, a cleavage product of the amyloid precursor protein, plays a key role in the pathology of the disease. Identification of clinically useful inhibitors of A β 42 oligomerization has been hindered by the structural complexity of A β fibrils and the lack of effective screens for A β aggregation inhibitors. Now, Kim *et al.* (p 461; see the accompanying Point of View by Gazit) present a high-throughput fluorescence-based screen for small-molecule inhibitors of A β 42 aggregation.

The screen is cleverly designed to couple inhibition of A β 42 aggregation with an increase in fluorescence. A protein construct containing the A β 42 peptide fused

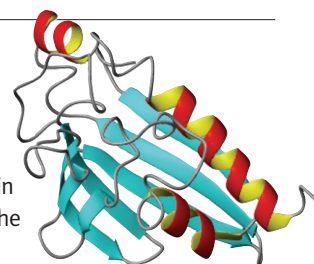
to GFP was created and placed under an inducible promoter for expression in *Escherichia coli*.

In the absence of an inhibitor, protein expression results in A β 42 aggregation, which prevents GFP from folding into its native, fluorescent structure. A library of ~1000 molecules based on a triazine scaffold was screened, and compounds that caused an increase in fluorescence were identified as potential A β 42 aggregation inhibitors. Biochemical aggregation experiments and electron microscopy analysis of one of the hits confirmed that the compound prevented A β 42 fibril formation; this provided independent validation of the assay. In addition to facilitating the search for potential drugs for AD, this innovative screen could be applied to finding therapeutics for other aggregation-associated diseases, such as Huntington's and prion diseases.

Sacrificial Proteins

Bacteria have a notorious talent for developing resistance to antibiotics. A particularly inventive method for evading the activity of a structural subclass of the enediyne antibiotics was recently discovered, wherein a bacterial protein is sacrificed in the process of inactivating the antibiotic. The naturally occurring enediyne antibiotics exert their activity through oxidative damage to DNA, but when bacteria send in the resistance protein CalC, the enediyne wields its destructive forces on the protein instead, sheltering the DNA while CalC is proteolyzed. Singh *et al.* (p 451) now elucidate the NMR structure of CalC alone as well as bound to the enediyne calicheamicin (CLM), exposing insightful molecular details of the self-sacrifice mechanism of enediyne resistance.

Analysis of the closest structural homologues revealed that CalC is a member of the steroidogenic acute regulatory related lipid transfer (START) protein superfamily. Despite the wide range of organisms and cellular activities in which START proteins function, they all contain a conserved hydrophobic binding cavity that dictates substrate specificity. CLM is strategically positioned in this cavity in CalC such that the DNA binding portion of the molecule is buried and the enediyne warhead is precisely situated to abstract the α -hydrogen of CalC Gly113. The proteolytic self-sacrificing process is thus triggered. Intriguingly, the broad distribution of START domains located throughout human cells present the possibility that proteins containing START domains may also facilitate the intracellular transport of CLM to its site of action, the nucleus.



Published online August 18, 2006

10.1021/cb6003405 CCC: \$33.50

© 2006 by American Chemical Society

Introducing our AUTHORS



Natalie T. Rundle

Current position: Maternity leave

Education: University of Toronto, B.S., Honors, 1994; McMaster University, M.S., 1997; University of British Columbia, Ph.D., 2003, with Prof. Michel Roberge

Postdoctoral work: Queensland Institute of Medical Research, Brisbane, Australia, 2004–2005, with Prof. Martin Lavin

Nonscientific interests: Outdoor fitness, travel, music, literature

A major interest of mine is the use of cell-based screening of natural products to identify new chemicals that affect cell-cycle progression. I find it attractive that these chemicals can serve as tools to study biological mechanisms and that they may have therapeutic potential. My work has identified a structural class of compounds that arrest cells at an early stage of mitosis. It was an exciting step forward for us to be able to chemically modify the compound described in this paper in a way that allowed the identification of its binding partners and the visualization of its binding sites within the cell. This paper illustrates a chemical genetics approach to studying the complex pathways in mitosis specifically, but the strategy may be applied to protein target identification in other systems as well. (Read Rundle's article on p 443.)

The molecular cascade leading to Alzheimer's disease (AD) initiates with the misfolding, oligomerization, and aggregation of the Alzheimer's β -amyloid ($A\beta$) peptides. Inhibition of this pathway is an attractive target for intervention, but finding compounds that will inhibit aggregation is challenging. To meet this challenge, we developed a high-throughput screen by fusing $A\beta$ to GFP. In this fusion, the folding and fluorescence of GFP are blocked by $A\beta$ misfolding and aggregation. Compounds that inhibit $A\beta$ misfolding and aggregation relieve this blockage, enable GFP folding, and are readily identified as green fluorescent "hits". Because the screen is sensitive, fast, and inexpensive, we expect that it will facilitate rapid progress in the search for lead compounds that can be developed into drugs that prevent AD. (Read Kim's article on p 461.)

Current position: Princeton University, Department of Chemistry, Ph.D. candidate with Prof. Michael H. Hecht

Education: Seoul National University, Department of Chemistry, B.S., 1998

Nonscientific interests: Swimming, sci-fi



Woojin Kim



Shanteri Singh

Current position: University of Wisconsin, Madison, School of Pharmacy, Department of Pharmaceutical Sciences, Assistant Scientist in Prof. Jon S. Thorson's group

Education: Tata Institute of Fundamental Research, Mumbai, India, Ph.D., 1996, with Prof. R. V. Hosur

Postdoctoral work: Memorial Sloan-Kettering Cancer Center (MSKCC), New York, 1998, with Prof. Dinshaw J. Patel; Universiteit Utrecht, Bijvoet Center for Biomolecular Research, the Netherlands, 2003, with Prof. Rob Kaptein; University of Wisconsin, Madison, Center for Eukaryotic Structural Genomics (CESG), 2005, with Prof. John L. Markley

Nonscientific interests: I love spending time with my teenage daughter and keeping up with her day-to-day life

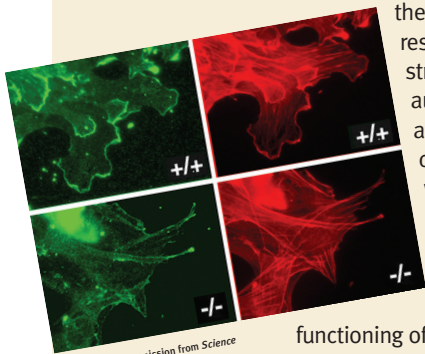
Most of my work focuses on the structure determination of biological macromolecules using NMR spectroscopy. Presently, I am working on protein engineering to tailor enzymes involved in biosynthetic pathways. The structural studies of CalC and its relationship to steroidogenic acute regulatory protein related transfer (START) domain proteins could implicate a new role for CalC. Apart from structural studies, I am also interested in using NMR as a tool to understand enzyme mechanisms, protein engineering, and diversification of natural products. (Read Singh's article on p 451.)

Spotlight

Actin Arginylated

Protein arginylation is a post-translational modification in which an arginine residue is transferred to the N-terminus of a protein by the enzyme arginine-transfer RNA protein transferase (Ate1). Ate1 knockout mice are embryonic lethal, having severe defects in cardiovascular development and angiogenesis, but the molecular basis of and the proteins affected by post-translational arginylation have remained ambiguous for >40 years. Toward understanding the biological role of N-terminal arginylation, Karakozova *et al.* (*Science* 2006, 313, 192–196) decipher the biochemical and cellular effects of β -actin arginylation.

A combination of 2D gel electrophoresis and mass spectrometry on samples derived from embryonic fibroblasts was used to confirm that β -actin is arginylated *in vivo*. To determine how the physical properties and biological function of actin are affected by arginylation, the authors compared embryonic fibroblasts from wild-type mice and mice deficient in Ate1 (Ate1^{-/-}). *In vitro* biochemical examination revealed that β -actin derived from Ate1^{-/-} cells was as stable and interacted with the same profile of proteins as β -actin from wild-type cells. However, in contrast with the single β -actin filaments that are formed in normal cells, actin from Ate1^{-/-} cells formed filamentous aggregates, hindering the ability of Ate1^{-/-} cells to move as proficiently as wild-type cells. In addition, Ate1^{-/-} cells had defects in spreading, lamella formation, and intracellular localization of β -actin. On



Reprinted with permission from *Science*

the basis of these results and the crystal structure of actin, the authors propose that arginylation of actin coats the filaments with a positive charge that prevents aggregation, contributing to proper functioning of the protein. The importance of Ate1 in embryonic

development highlights the need to understand the biological role of Ate1 function and the consequences of perturbing this protein. These results not only shed light on the role of N-terminal arginylation in actin function but also pave the way toward understanding the global role of N-terminal arginylation. **EG**

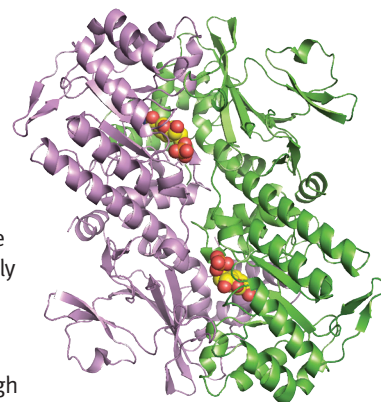
A Light in the NAMPRase Tunnel

Nicotinamide phosphoribosyltransferase (NAMPRase) is an important enzyme in the biosynthesis of nicotinamide adenine dinucleotide (NAD⁺), a molecule intimately involved in biochemical redox reactions during vital processes such as glycolysis and the citric acid cycle. Interestingly, depletion of NAD⁺ levels in tumors through inhibition of NAMPRase activity has demonstrated encouraging anticancer effects.

Because NAMPRase catalyzes the conversion of free nicotinamide to nicotinamide mononucleotide (NMN), which is a key step in the salvage pathway of NAD⁺, inhibitors of the enzyme may have potential as cancer drugs. Khan *et al.* (*Nat. Struct. Mol. Biol.* 2006, 13, 582–588) and Wang *et al.* (*Nat. Struct. Mol. Biol.* 2006, 13, 661–662) now report the crystal structures of free NAMPRase, NAMPRase bound to NMN, and NAMPRase bound to the inhibitor FK866. The structures provide insights into the substrate specificity and the mechanism of the enzyme and jumpstart the rational design of novel NAMPRase inhibitors.

The structure of NAMPRase revealed that it belongs to the dimeric class of type II phosphoribosyltransferases, which include nicotinic acid phosphoribosyltransferase (NAPRTase) and quinolinic acid phosphoribosyltransferase (QAPRTase). The proteins can each be organized into three domains composed of a mixture of β -strands and α -helices, and all three proteins possess an extensive dimer interface. However, it is a few key structural differences among these enzymes that ultimately expose the basis for their substrate specificity. NAMPRase is quite a bit larger (~100 amino acids) than either NAPRTase or QAPRTase, and distribution of these additional residues over the structure, along with differences in domain orientations, has a dramatic impact on the active-site environment of NAMPRase.

Structures of NAMPRase bound to NMN and FK866 revealed that the active site of the enzyme is located at the dimer interface. In fact, the nicotinamide ring of NMN participates in π -stacking interactions with a phenylalanine from one monomer unit and a tyrosine from the other monomer unit. The basis for the substrate specificity centers at an aspartate residue, which is not present in NAPRTase or QAPRTase and which takes part in a direct hydrogen bond with the amide group of NMN. The importance of the aspartic acid residue in defining the substrate specificity of NAMPRase was also confirmed by mutagenesis and kinetic studies. The significance of the dimerization of NAMPRase is reinforced upon examination of the binding of the inhibitor FK866. At the dimer interface, FK866 binds in a tunnel, with some resemblance to the binding of NMN. Notably, structural and kinetic data indicate that FK866 is a tight-binding competitive inhibitor of NAMPRase, in contrast to previous reports that FK866 inhibits NAMPRase *via* a noncompetitive mechanism. The unique presence of the tunnel confers specificity of FK866 for NAMPRase over NAPRTase and QAPRTase, because FK866 is exquisitely shaped to partake in favorable interactions upon slithering into place. The information gained from these structures will contribute significantly to furthering our understanding of the enzyme's mechanism and role in biology. **EG**



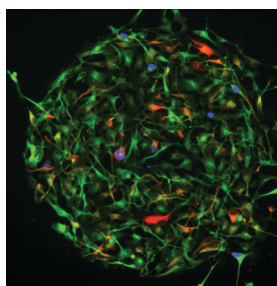
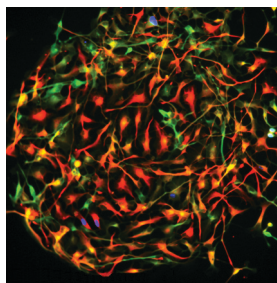
Reprinted with permission from *Nature Structural & Molecular Biology*

Understanding Fate

The process of cell differentiation is driven by complex spatial and temporal signaling mechanisms. Systematic exploration of the molecular factors that contribute to the fate of a cell will help scientists navigate the murky waters of cellular differentiation. To this end, Soen *et al.* (*Mol. Syst. Biol.*, published online July 4, 2006, doi:10.1038/msb4100076) present a microarray-based method for investigating the phenotypic effects of exposing neural precursor cells to different combinations of extracellular signaling molecules.

A microarray was generated of defined combinations of 13 recombinant signaling molecules, many of which have been implicated in neuronal cell differentiation. Bipotent neural cells capable of differentiating into neurons or glial cells were allowed to attach to the microarray surface and were incubated under conditions favorable for differentiation. Analysis

of neural and glial cell differentiation markers on each cell enabled the effects of the molecular microenvironments to be assessed. The researchers determined that different combinations of signaling molecules resulted in four distinguishable outcomes relating to the differentiation state of the cells. Whereas certain mixtures of signaling factors promoted differentiation toward glial cells, others nudged cells toward becoming neurons. Interestingly, some combinations appeared to decrease both differentiation markers, in essence retracting the cells into an “undifferentiated-like” state that coincided



Reprinted with permission from *Molecular Systems Biology*

with an increased proliferative phenotype. Still other mixtures increased both differentiation markers, with the cells classified as being in an indeterminate state of differentiation. Analysis of the relationships within mixtures of signaling factors pointed to additional subtleties, including sometimes unexpected dose–response and kinetic profiles and the ability of certain molecules to have dominant effects over others. This powerful method can be adapted to the investigation of additional molecular factors with a variety of cell types, enhancing our understanding of the molecular environment involved in cell differentiation and progressing the exciting prospect of manipulating the fate of a cell. **EG**

Worming Our Way into New Antibiotics

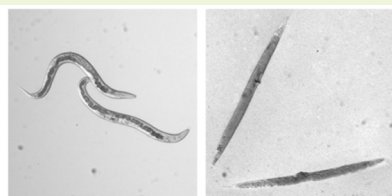
The growing number of infections caused by bacteria resistant to known antibiotics is a worldwide health concern. However, the critical need for new antibiotics has been plagued by the limitations of traditional screens. Typical screens are unable to recognize toxic molecules, compounds with poor pharmacokinetic properties, or molecules that cannot penetrate the multidrug-resistance barrier of Gram-negative bacteria. In addition, most *in vitro* screens barely resemble the biological systems they are attempting to replicate, and this calls into question their relevance. Moy *et al.* (*PNAS* 2006, 103, 10414–10419) now report an innovative, high-throughput, live-animal antibiotic screen using the nematode *Caenorhabditis elegans* and the human pathogenic bacteria *Enterococcus faecalis*.

E. faecalis is a human opportunistic bacterium that, like many human bacterial pathogens, also infects the nematode intestinal tract. When *C. elegans* are infected with *E. faecalis*, half the worms die within 5 days, but antibiotic treatment upon infection

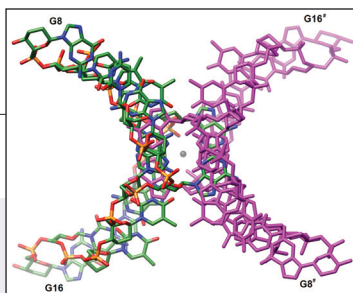
can rescue the worms from death. On the basis of this system, a screen was developed to identify novel antimicrobials that could cure nematodes infected with *E. faecalis*. Infected worms were transferred to a liquid medium in 96-well plates, and 6000 synthetic molecules and 1136 natural product extracts were

tested for their ability to cure the infection. Visual inspection with a dissecting microscope easily distinguished live worms, which adopt a sinusoidal posture, from dead worms, which become straight and rigid because of bloating from the *E. faecalis* cells. Eighteen of the small molecules and nine of the extracts were found to promote survival of infected

worms. This screen has a number of advantages over traditional antibiotic discovery screens, including the ability to identify prodrugs, compounds that target virulence factors, and molecules that enhance the host’s defense system. In addition, the assay selects for nontoxic compounds that are effective *in vivo*. This live-animal screen presents an intriguing new method for antibiotic discovery. **EG**



Reprinted with permission from the *Proceedings of the National Academy of Sciences*



Reprinted with permission from the *Journal of the American Chemical Society*

Rationalizing the Ribose

Many polymers in biology assemble from monomers that display unifying chemical properties. Proteins are assembled from amino acids that all share L-chirality, whereas DNA and RNA are sugar-coated strictly with pentose in the backbone. Such stringent choices have remained a ponderous point for biologists and chemists who envision founder macromolecules emerging from a prebiotic chemical soup. A new study by Egli *et al.* (*J. Am. Chem. Soc.*, ASAP Article 10.1021/ja062548x S0002-7863(06)02548-0) explored this theme by asking DNA to trade in its standard ribose for a hexose sugar. The group synthesized a hexose-based nucleic acid, termed homo-DNA. Although at first glance the functional groups and geometry looked rather similar to DNA, a high-resolution view of homo-DNA demonstrated radical differences. The X-ray crystal structure of a double-stranded octamer revealed base-pairing and helical properties that are quite foreign to the textbook rules for DNA. The duplex resembles a slowly twisting ribbon rather than the tight coil of DNA. The steps between each base pair varied considerably, and the intrastrand base stacking found in nature's double helix was completely lacking. Some likenesses were observed, such as cross-strand base stacking and the antiparallel architecture, but the elegant uniformity that DNA uses to store genetic information was largely absent. The researchers postulate that stable base-pairing systems are highly unlikely with hexose-based nucleic acids, and this might explain why nature chose pentose over hexose. This study is also particularly interesting because of the techniques used to solve the structure of the duplex. An old friend to protein crystallographers, selenium, was used in the form of a phosphoroselenoate in the homo-DNA backbone. These compounds are usually too reactive for the time scale of crystal growth, but in this case, the researchers miraculously timed their synthesis, crystallization, and data collection to make this unique structure possible. **JU**

Derepressing Antidepressants

Histone methylation is one of several modifications of chromatin structure that play a key role in the regulation of gene expression. BHC110, an enzyme found in a number of multiprotein complexes involved in nucleosome modification, is capable of demethylating histone H3 lysine 4 (H3K4) and consequently causes repression of gene expression. BHC110 shares sequence homology with monoamine oxidase (MAO) enzymes, which are targets of several antidepressant drugs. Lee *et al.* (*Chem. Biol.* 2006, 13, 563–567) now report that certain MAO inhibitors are also potent inhibitors of BHC110 and that cells exposed to these inhibitors exhibit transcriptional derepression of BHC110 target genes.

Three selective and three nonselective MAO inhibitors were tested for their ability to inhibit histone and nucleosome demethylation *in vitro* by recombinant BHC110, and two of the nonselective inhibitors showed dose-dependent activity against

the enzyme. The most active compound, tranlycypromine (brand name Pamate), had an IC_{50} of $<2 \mu M$, which notably is 10-fold less than the IC_{50} of the drug against MAO enzymes. Tranlycypromine was next tested for its ability to inhibit histone demethylation in live cells. The transcriptional activity of two BHC110 target genes, *Oct4* and *Egr1*, was analyzed in response to tranlycypromine exposure. Quantitative reverse-transcriptase PCR and chromatin immunoprecipitation experiments in embryonic carcinoma cells revealed that tranlycypromine treatment results both in derepression of *Oct4* and *Egr1* gene expression and in enhanced global histone methylation levels. This discovery offers new insights into the mechanisms of some antidepressant medications and provides additional tools for exploring the role of histone demethylation in important molecular and cellular processes, such as gene expression, cellular differentiation, and oncogenesis. **EG**

Prying into Prion Mechanisms

Proteinaceous infectious particles, or prions, are unique protein pathogens thought to be responsible for several fatal diseases, including scrapie, Creutzfeldt–Jakob disease, and bovine spongiform encephalopathy (“mad cow” disease). One mysterious and remarkable characteristic of prion pathogenesis is that different conformations of the same misfolded protein produce different disease phenotypes. Tanaka *et al.* (*Nature* 2006, 442, 585–589) provide insight into this phenomenon by demonstrating that the brittleness of prion aggregates can affect the rate of prion division, ultimately leading to distinct physiological consequences.

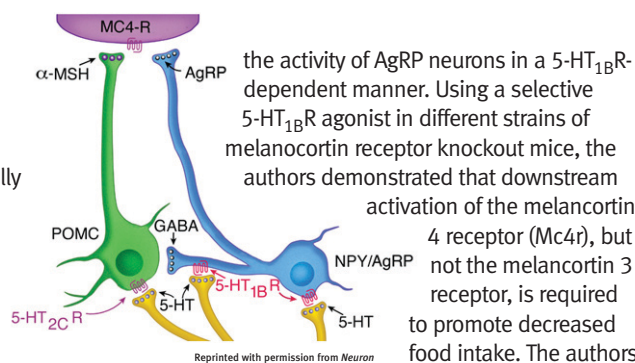
The researchers used synthetic prion forms of the yeast protein Sup35 as a model to investigate the physiological impact of different prion conformations. The color phenotype of the $[PSI^+]$ prion state, which results from Sup35 aggregation, varies depending on the physical properties of the Sup35 aggregate in the $[PSI^+]$ strain, and this provides an easily monitored system. The intrinsic fiber growth rate and the frangibility, or propensity to fragment, of three infectious amyloid conformations of the prion-forming domain of Sup35 (Sc4, Sc37, and SCS) were characterized by atomic force microscopy. Unexpectedly, it was found that the strain with the strongest phenotype, Sc4, has the slowest intrinsic growth and is the most likely to fragment. Further investigation using a variety of assays revealed that Sc4 also possesses the fastest rate of division in cells, easily compensating for the slower growth rate. In addition, the Sc4 prion particles are noticeably smaller than those found in the other strains, but the number of fibers per cell is considerably higher. Taken together, the data indicate that the strength of a prion strain phenotype is directly related to the frangibility of the infectious prion aggregate. This revelation not only demystifies a piece of the mechanism behind prion pathogenesis but also points to new strategies for restraining the infectious competence of prion aggregates. **EG**

Serotonin Weighs In

Levels of the neurotransmitter serotonin (5-HT) inversely correlate with food consumption, making some 5-HT analogues excellent weight-loss agents. However, 5-HT is involved in a host of metabolic and neurological activities that can also affect eating behavior, and this complicates the identification of the precise mechanism by which 5-HT regulates food intake. Now Heisler *et al.* (*Neuron* 2006 51, 239–249) further define the role that 5-HT plays in appetite suppression by demonstrating that the melanocortin system, a group of pituitary peptide hormones and their receptors known to be involved in a variety of biological activities, including feeding, is a critical component of the regulation of food intake by 5-HT.

5-HT is synthesized from the essential amino acid tryptophan in the brain, where it interacts with several types of 5-HT

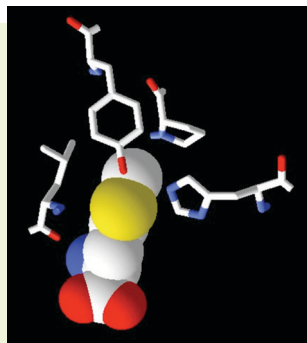
receptors, including 5-HT_{1B}R. Using transgenic mice, the authors observed that 5-HT_{1B}R receptors are anatomically positioned to regulate neurons containing the melanocortin agonist α -melanocyte-stimulating hormone (α -MSH) and the melanocortin antagonist agouti-related protein (AgRP). Both of these molecules are potent regulators of food intake. In addition, light and electron microscopy experiments revealed that 5-HT terminals are located such that both neuronal activity and release of products from the axon are likely affected by 5-HT. Electrophysiology experiments further indicated that 5-HT both increases the activity of α -MSH neurons and reduces



the activity of AgRP neurons in a 5-HT_{1B}R-dependent manner. Using a selective 5-HT_{1B}R agonist in different strains of melanocortin receptor knockout mice, the authors demonstrated that downstream

activation of the melanocortin 4 receptor (Mc4r), but not the melanocortin 3 receptor, is required to promote decreased food intake. The authors

propose a model in which reciprocal regulation of melanocortin agonist and antagonist-containing neurons, in concert with downstream activation of Mc4r, is the key pathway through which 5-HT exerts its appetite-suppressing activity. Additional molecular insights into this pathway will enhance our understanding of food regulation and may facilitate the development of more-effective weight-loss agents. **EG**



Reprinted with permission from the *Proceedings of the National Academy of Science*

Cell Surface Mutants

The incorporation of non-canonical amino acids into recombinant proteins enables the generation of innovative tools with which to manipulate molecular and cellular function. Aminoacyl-transfer RNA synthetases (aaRS), the enzymes that ligate specific amino acids to their cognate tRNAs, can be mutated

to permit incorporation of noncanonical amino acids into proteins without affecting wild-type aaRS activity. The use of mutant aaRS offers unique control over recombinant protein production, but the generation of effective mutants can be a tedious process. Link *et al.* (*PNAS* 2006, 103, 10180–10185) now describe a high-throughput, flow-cytometry-based method for identifying mutant aaRS that efficiently incorporates non-canonical amino acids into target proteins.

On the basis of previous studies, the authors chose a variant of the *Escherichia coli* outer membrane protein C (OmpC) for

incorporation of the unnatural amino acid azidonorleucine (ANL) by the *E. coli* aaRS methionyl-tRNA synthetase (MetRS). The authors selected four sites for mutagenesis within MetRS by examining the crystal structure and identifying the residues most critical for methionine binding, and they used a modified PCR gene assembly process to generate a saturation mutagenesis MetRS library. After transformation of the library into bacteria, cells expressing recombinant OmpC containing ANL were covalently labeled *via* reaction of the azide of ANL with a molecule containing a biotinylated cyclooctyne functionality. Staining with fluorescent avidin followed by flow cytometry and cell sorting analysis led to the identification of three mutant MetRS proteins. Remarkably, all three contained the same leucine to glycine (L13G) mutation, and a protein containing this single mutation was subsequently generated and found to be the most efficient of the MetRS mutants evaluated for ANL incorporation. These results demonstrate the power of rapid identification of mutated aaRS for the expression of cell surface proteins possessing unique reactivity, and this methodology can easily be expanded to the generation of additional mutant aaRS with other noncanonical amino acids. **EG**

Chemical Biology Education at Dortmund: A Joint Endeavor with a Max Planck Institute

Hans-Dieter Arndt^{†,*}, Christof M. Niemeyer[†], and Herbert Waldmann^{†,*}

[†]Universität Dortmund, Fachbereich Chemie, Otto-Hahn-Strasse 6, D-44221 Dortmund, Germany, ^{*}Max-Planck-Institut für molekulare Physiologie, Otto-Hahn-Strasse 11, D-44227 Dortmund, Germany

During the last few decades, biology has been transformed from a science that relies on strongly descriptive and phenomenological correlation into a molecular science. Biological processes can be traced back to chemical reactions and depend on the properties and interactions of molecules.

Biology is now tightly intertwined with chemistry. This new interfacial discipline, now commonly referred to as chemical biology, has had a rejuvenating impact on chemical and biological research. The academic training and teaching curricula of these neighboring disciplines must adapt to meet the demands and expectations of a new generation of students.

Addressing this demand will require new lecture courses that combine a focused chemistry curriculum with the most important lessons of biology and biochemistry. However, without additional practical training, a truly interdisciplinary education (a prerequisite for interdisciplinary research) cannot be achieved.

Chemical Biology Education in Dortmund. On the basis of this analysis, the faculty of the department of chemistry at the University of Dortmund created a curriculum alternative to that of traditional chemistry. The close ties and collaborations with the Max Planck Institute (MPI) for Molecular Physiology, which is located on the same campus, facilitated the process. This fundamental biological competence compensated for the lack of a biology department at

the University of Dortmund, which was founded as an engineering school. Hence, this curriculum is a very successful example of collaborations between a research institute and a university.

Moreover, there was academic and political pressure to transform the traditional German *Diplom* as a qualifying degree into a two-stage system of B.S. and M.S. degrees, in order to promote student exchange with the European countries and abroad. This new B.S. degree would encompass broad training over 3 yr and ideally would qualify students for their first jobs. The M.S. phase would then offer advanced training for higher qualification and be a prerequisite for doctoral programs. Both demands accordingly shaped the new program in chemical biology at Dortmund.

Bachelor's of Science in Chemical Biology. Every chemist knows that molecules have specific forms, properties, and reactivity and that they can be tailored in boundless ways. Chemists typically are trained about the world of small molecules, whereas biochemistry and biology mostly deal with large macromolecules. Frequently, curricula create the perception that rules of chemical reactivity no longer apply to large molecules and that learning about biomacromolecules and small molecules necessitates separate scientific approaches. However, perpetuating



The Dortmund University Campus

Image courtesy of MPI Dortmund

*Corresponding author,
hans-dieter.arndt@
mpi-dortmund.mpg.de.

Published online August 22, 2006

10.1021/cb600320r CCC: \$33.50

© 2006 by American Chemical Society

Chemical biology students are adaptable: those with good results in the chemical subjects also tend to display good performance in biology.

Box 1. Bachelor's of Science in Chemical Biology

Overall size: ~250 students

Applicants admitted per year: ~90

Official start: fall 2003, annual enrollment

Duration: 3 yr

Lecture courses and problem-solving classes: general and analytical chemistry, organic chemistry, inorganic chemistry, physics, mathematics, physical chemistry, biophysics, bioorganic and bioinorganic chemistry, biochemistry and molecular biology, cell biology, microbiology, and toxicology

Laboratory courses: inorganic chemistry, organic chemistry, physical chemistry, bioorganic chemistry, biochemistry, microbiology, and cell biology

One elective subject: *e.g.*, economics, chemical engineering, biotechnology, organic synthesis

Bachelor's thesis: 13 wk full time, thesis, public thesis defense

Language of teaching: German (thesis optional in English)

such separation will not prepare students for a thorough scientific understanding or for the next wave of innovation. It is in the beginning of our educations that chemistry should define concepts and ideas. Many of us find it very difficult to rethink our molecular understanding later in our educational process.

Hence, the B.S. course (Box 1) designed in Dortmund has its roots in chemistry. It was not intended to reinvent the biochemistry curriculum but rather to offer chemistry with a distinct new focus. The elements of the classic curriculum that illuminate an understanding of molecules were kept. However, at the expense of more specialized chemistry topics (*e.g.*, organometallic chemistry, technical chemistry, or a physics lab course), basics were incorporated, including cell biology, microbiology, biochemistry, molecular biology, and bioorganic chemistry.

Each of these subjects is accompanied by an introductory lab course, allowing the students to put their growing knowledge into practice as early as possible. In Dortmund, a cross-disciplinary practical course in chemical biology has been established. It features student experiments in organic chemistry, biochemistry, cell biology, biophysical chemistry, and microarray technologies, both at

the undergraduate and graduate levels. In our experience, these lab courses prove extremely important: becoming acquainted with the distinct setups of chemical and biological experiments simultaneously is one of the keys to stimulating interdisciplinary thinking later on. (A book describing the individual experiments is available (1). We are very open to suggestions and additional experiments for upcoming editions.)

For practical reasons, not all chemistry lecture courses and exams could be redesigned from scratch; together with the additional biological focus, an intense learning atmosphere resulted naturally. It is interest-

ing to note that our chemical biology students are adaptable and are gaining a broad understanding: those with good results in the chemical subjects also tend to display good performance in biology, and *vice versa*. Currently, the first round of students is finishing their B.S. thesis work and, as far as we can judge, with very good results.

Master's of Science in Chemical Biology.

This program (Box 2) should prepare the student for either job opportunities on the intermediate research level or careers in nonresearch-focused enterprises (insurance, marketing, law, *etc.*). If a research-oriented career is the goal, the students will mature in the program and prepare for doctoral studies. Most lectures offered by the faculty, therefore, concentrate on current topics in chemical biology and include problem-solving classes, seminars, and student presentations. Because external and international applicants may enter the program, this focused approach accommodates a diverse group of students. Furthermore, many courses are elective, and this promotes early building of a scientific focus. Advanced lab courses have been integrated to create familiarity with complex experimentation and instrumentation, and this is further deepened in research internships in labs of their choice. For the master's thesis, the independent research capabilities de-

Box 2. Master's of Science in Chemical Biology

Overall size: ~80 students

Applicants admitted per year: ~40

Official start: fall 2004, biannual enrollment

Duration: 2 yr

Lecture courses and problem-solving classes: chemical biology, nanobiotechnology; electives in medicinal chemistry, chemical genetics, special nucleic acid chemistry, biophysics, biomembranes, synthetic proteins, and more

Laboratory courses: advanced bioorganic chemistry, bioinorganic chemistry, nanobiotechnology, biophysics; compulsory internships in research laboratories

Master's thesis: 6 mo full time, thesis, public thesis defense

Language of teaching: German (English optional)

Box 3. IMPRS Chemical Biology Program

Joint program with the Ruhr-University Bochum located on the University of Dortmund campus

Overall size: ~30

Applicants admitted per year: ~10

Official start: January 2002

Annual start of the Ph.D. program: October 1 (from 2003 onward); exceptionally qualified candidates may be accepted throughout the year

Duration: 3 yr

Components: lectures series in chemical biology (2 h per wk), research-oriented practical courses (~3 courses, each 1 wk), research seminar (1 h per wk), a workshop in presentation techniques, and an annual scientific retreat of the IMPRS-CB (2–3 d) to promote exchange

Electives: laboratory rotation, summer schools

Language of teaching and Ph.D. thesis: English (German optional)

velop, while the students tackle a specific research assignment over a period of 6 mo. If the student's performance is more than satisfactory, the degree will qualify him or her for doctoral studies in Germany, other European countries, and most countries overseas.

International Max Planck Research School for Chemical Biology. Doctoral studies in Germany are generally taken up on the initiative of students and are based on individual decisions. Graduate schools are not common. Therefore, in addition to the doctoral programs in chemistry and biology at the local universities, a combined effort of the MPI of Molecular Physiology, the University of Dortmund, and the Ruhr-University Bochum led in 2001 to the founding of the International Max Planck Research School in Chemical Biology (IMPRS-CB, Box 3). The objective of the IMPRS-CB is to advance the application of chemistry to biochemistry and cellular biology at the graduate level by fostering emerging scientists from biochemistry, biophysics, and chemistry.

A training program for the students was designed to supplement the typical doctoral curriculum in order to accomplish this goal. It is composed of an advanced lecture

program and practical training. Specific lectures from bioorganic and bioinorganic chemistry, biophysics, biochemistry, nanobiotechnology, structural biology, molecular and cell biology, physiology, and bioinformatics are offered. Frequently, outside instructors are invited to contribute their in-depth expertise to the lecture course program.

The objective of the IMPRS-CB is to train excellent, independent researchers. Hence, a multidisciplinary committee reviews the applications and invites candidates to in-house interviews, giving them the opportunity to discuss their previous research topics and meet faculty members. Students admitted to the Ph.D. program then choose one of the participating research groups and begin conducting experimental work for their Ph.D. thesis immediately. The project for the Ph.D. thesis is guided by the chosen faculty adviser.

Challenges. When a curriculum is redesigned and a complete new study plan is offered, practical as well as intellectual pitfalls can emerge. First, a higher student workload generally cannot be avoided when they are learning two disciplines simultaneously. Sometimes, two or three demand-

ing, yet unrelated, courses must be taken concurrently.

Second, in our experience, the term “biology” present in the label of our curriculum attracts a much different and more diverse student body than that seen in chemistry or biochemistry programs. This student body sometimes tends to underestimate the “chemistry” in the program title. If a program leans toward chemistry (as our B.S. does), this issue must be clarified at the very beginning, so that students do not enroll with false expectations.

Third, in teaching as well as in exams, we as faculty must develop a sense of perspective. Most of us were trained in a particular discipline and only later did our curiosity drive us to learn about the other. It is not easy to adjust our level of expectation of what even gifted students can deliver when they study two essential subjects in parallel. Great care must be taken that in our enthusiasm we do not perpetuate what we want to overcome: rejection of one subject or the other!

Another challenge is the sometimes conservative hiring policy of some companies that may favor graduates with a “clean” résumé (“We do not hire hybrids”); students may base their decisions and choices on these expectations. But students are inevitably driven to focus, and their core competencies will grow. We expect the leaders of tomorrow to be in command of both chemical and biological contexts, and a chemical biology program will build a true sense for both. Success will convince.

Of course, an education in chemical biology may not cater to every need, and there are good reasons to keep the traditional chemistry and/or biology options. In essence, any program like ours must make sure that true benefits are offered, both for the undergraduate and graduate students, and for the faculty. Toward this end, we see students develop new research interests, aligning not with the dry wisdom of textbooks but with the opportunities of the

future; they find new connectivity beyond known patterns. They get immersed with concepts from very different fields, giving them more room to develop into a direction that suits their particular characters. They learn to communicate between the disciplines from the very first day, yet to focus when necessary: it is here that fruition can be expected in the near future.

On the other hand, faculty must rethink its approach and define common interests. The sometimes strict border between the synthetic and the investigative philosophies of our sciences must be softened: chemical biology needs both. This vision clearly also meets the demands of modern exploratory biology, which is constantly developing new synthetic traits. The field will benefit greatly from the knowledge of chemistry (which is synthetic by its very nature) instilled in this new generation of promising researchers.

Conclusion. When our program was conceived at the beginning of the new millennium and subsequently implemented, it was the only one of its kind in Germany and expectations were rather mixed. At that time, enrollment of students in chemistry had declined almost everywhere in Germany to a level rarely seen before. However, our unique offer of chemical biology was not only welcomed but instantaneously taken

by so many prospective students that local-access restrictions have been in effect since 2005 in order to maintain quality. In the Dortmund IMPRS-CB graduate program, applications regularly outnumber the stipends by >15-fold, and the first graduates are leaving this year, with more than one career offer in their bags. We very much hope that the strong demand carries on an

that our distinctive program continues to attract high-caliber undergraduate and graduate students from Germany and abroad (Box 4).

REFERENCE

1. Waldmann, H., and Janning, P. (2004) *Chemical Biology—A Practical Course*, 1st ed., Wiley-VCH, Weinheim, Germany.

Box 4. Contacts and Links

General and Student Information

Dekanat, FB Chemie der Universität Dortmund, Otto-Hahn-Str. 6, D-44221 Dortmund, Germany
www.chembiol-dortmund.de/english
raimund.leibold@uni-dortmund.de

Research Opportunities in Chemical Biology

Chemistry Department at the University of Dortmund,
www.chemie.uni-dortmund.de
 Max Plank Institute for Molecular Physiology,
www.mpi-dortmund.mpg.de
 Chemical Genomics Center,
www.cgc.mpg.de

IMPRS

IMPRS for Chemical Biology, Otto-Hahn-Str. 11, D-44227 Dortmund, Germany
www.imprs-cb.mpg.de/index.html

Specific Questions

hans-dieter.ardt@mpi-dortmund.mpg.de
christof.niemeyer@uni-dortmund.de
herbert.waldmann@mpi-dortmund.mpg.de

When Undergraduates Ask “Why,” Chemical Biology Answers

Joshua A. Kritzer*

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142

Universities face many challenges in educating the next generation of scientists, including how to implement inquiry-based teaching methods and how to prepare students for careers in chemical biology. New initiatives are demonstrating that the breadth and innovation inherent to chemical biology can be used to bring more student-led inquiry into the classroom and the teaching laboratory.

As any exasperated parent of a five-year-old can tell you, children are born with an innate desire to ask questions. “Why doesn’t the ocean dry up?” “Where do plants go in the winter?” “What are germs, and why do they make us sick?” When students reach the undergraduate level, they often rediscover the childlike excitement of broad academic freedom. The prospect of choosing their own classes, majors, and electives forces students to consider their own interests and their own questions about the world. Recent programs aimed at reforming undergraduate science education have taken into account this inquisitive energy. Inquiry-based teaching has emerged as a pillar of well-publicized initiatives such as the National Resource Council’s *BIO 2010* report (1), the National Academies’ Summer Institutes on Undergraduate Education in Biology (www.academiessummerinstitute.org), and the Howard Hughes Medical Institute’s (HHMI) Professors Program (www.hhmi.org/research/professors). By allowing undergraduates to follow open-ended research projects, often in groups with the help of a mentor, these programs let under-

graduates experience firsthand the excitement of scientific research.

The questions of how to implement inquiry-based teaching dovetail with concerns on how to train the next generation of chemical biologists (2). Chemical biology poses unique problems for education because it requires breadth of knowledge as well as depth; even scientists with great in-depth understanding of their own fields can run up against obstacles when pursuing a lead in chemical biology. For instance, a cell biology postdoctoral fellow who never took an undergraduate statistics course might suddenly find himself awash in microarray data. Or a chemistry grad student might have spent years synthesizing a family of natural products only to be left to her own devices to demonstrate their relevance in biological assays. Training the next generation of chemical biologists thus requires an undergraduate education that values a wide exposure to a variety of inter-related fields, from cancer evolution to robotics or from immunology to image analysis. But how can undergraduates gain exposure to diverse subject matter without becoming jacks of all trades and masters of none?

Inquiry-based approaches are a natural way to broaden the experience of future chemical biologists without completely

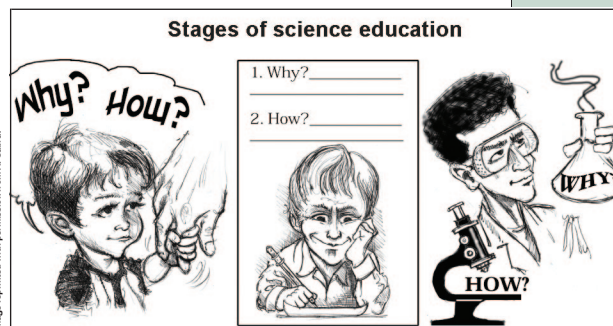


Image reprinted with permission from A. Cabral

*Corresponding author,
kritzer@wi.mit.edu.

Published online August 18, 2006

10.1021/cb600319s CCC: \$33.50

© 2006 by American Chemical Society

Chemical biology showed me how the reactions I learned in organic chemistry could be used in fascinating ways to explore biology.

Box 1. Student Feedback.

The Chemical Biology for Sophomores! lecture and lab courses have been revised, updated, and tweaked each year since their launch in 2004. A key part of this development process is feedback from students. The following are excerpts from anonymous end-of-course student questionnaires from the first 2 years of the courses.

On the Lecture Course

"I enjoyed the fact that we focused mainly on primary literature with Prof. Schepartz guiding us on the scientific story behind the papers. This was a unique thing to experience, especially as a sophomore."

"I surprised myself when I ended up liking the class—as a biology major, I had just finished organic chemistry the semester before and was looking forward to never taking another chemistry class again. However, chemical biology showed me how the reactions I learned in organic chemistry could be used in fascinating ways to explore biology. I would recommend this course if you really enjoy biology and wish to see how extremely relevant chemistry can be to the field."

"Great chance for non-chemistry majors to really see the interdisciplinary connections."

On the Lab Course

"It gave me a sense of what research is like without a commitment to a lab . . . a great place to decide whether research is for you."

The Chemical Biology Laboratory "is infinitely more interesting than standard laboratory courses, in which specific lab techniques are performed not as means to an end, but as ends in and of themselves. [This lab] is unique in that it gives a student a reason for performing the techniques, namely, to solve an actual research problem."

"I learned that you never know where trying to answer a seemingly simple question may take you. Scientific research is unpredictable."

rewriting existing curricula. By structuring an early-level course to allow undergraduates to pursue problems where they lead (rather than in a predetermined direction), educators can encourage students to explore fields on the borders of their chosen majors. Thus, despite the limited span of, say, a year-long lecture and laboratory course, students can gain experience in a variety of fields in a fluid way. Keeping both lecture and laboratory courses unforced and often self-directed provides a refreshing and realistic view of science and gives the student an appreciation for the interdisciplinary nature of modern

scientific research. Also, because chemical biology's most impressive work is still ongoing, the student comes away feeling that he or she has contributed substantively to chemical biology as a whole (and with the chance to publish their laboratory findings).

The Chemical Biology for Sophomores! lecture course and laboratory, first implemented at Yale University in 2004 and funded by an HHMI Professors Award to Alanna Schepartz (Yale University), exemplifies how chemical biology can make the most of inquiry-based teaching and *vice versa* (www.schepartzlab.yale.edu/

chembiol.sopho.org/ and www.hhmi.org/news/professor-schepartz.html). The lecture course offers detailed introductions to diverse topics. It uses current primary literature as the source text and casts the professor in the role of guide to explore the most dynamic and innovative sections of chemical biology. Similar survey formats may be more typical of graduate-level courses, but the science is not dumbed down or glossed over for sophomore-level undergraduates. Instead, the interdisciplinary nature of chemical biology allows for the teaching of basic science along with the cutting-edge applications. For instance, the course begins with Merrifield's original papers on solid-phase peptide synthesis (3, 4) and immediately dives into case studies on applications of intein chemistry and direct labeling of proteins *in vivo* (5–7). During this first unit, students are exposed to basic biochemistry (the names and structures of the 20 amino acids, and methods of protein expression in bacteria) as well as basic bioorganic chemistry (the strategies and mechanisms of protecting group chemistry, amide bond formation, and bio-orthogonal reactions). In this way, the focus on chemical biology frees professors to explore truly exciting, present-day fields of study while still teaching nuts-and-bolts basics.

Other open-ended techniques were used to complement this lecture strategy to enhance the atmosphere of exploration. Students were encouraged to think out loud about key papers in discussions led by open-ended questions. Every 2–3 weeks, case studies were used to reinforce the basic science behind the applications while exposing the students to innovative chemical biology being performed by leaders of the field. Students were asked to compose a review article or original research proposal on a topic of their own choosing rather than take a final exam. Individual students found that they could manage such an assignment (normally expected only of graduate students) because they had the freedom to

remain in their intellectual comfort zone while focusing on a question they were personally interested in. Altogether, these inquiry-based techniques have been successful because they take advantage of the unique scientific breadth and innovative characteristics of chemical biology.

The laboratory-based portion of Chemical Biology for Sophomores! also uses the field's strengths to advance inquiry among undergraduates. Graduate-student mentors lead small teams of undergraduates in semester-long research projects, from making buffers to writing up the results and everything in between. The projects are limited in scope and streamlined by the graduate students to fit into the allotted time frame but are not diluted in their import or rigor. For example, one project involved a phage display selection seeking to optimize miniature protein inhibitors of the oncoprotein hDM2 (8). The undergraduates performed techniques as diverse as Kunkel mutagenesis (9), phage preparation, bio-panning, and fluorescence polarization, all while getting an up-close view of how molecular evolution methods work (or fail to work). These topics and techniques span diverse areas of molecular biology, virology, biochemistry, and biophysics but were integrated into a seamless scientific story by the overarching goal of developing novel hDM2 inhibitors (10). Other projects have involved CD and NMR spectroscopy, synthesis and screening of synthetic libraries, cell-culture experiments, and live-cell fluorescence imaging. A short science-writing workshop is included in the laboratory course, and at the end of the semester each sophomore writes a preliminary manuscript with background, methods, results, and discussion sections. In the above example, the project was completed by the mentor and the work was published (11), ultimately giving the students both a first publication and a career-altering experience.

Educating the next generation of scientists to take advantage of the increasingly interdisciplinary nature of research will require major shifts in strategy, including broader implementation of inquiry-based teaching methods. New programs such as Chemical Biology for Sophomores! are demonstrating that the breadth and innovation inherent to chemical biology can facilitate student-led inquiry in the classroom and the teaching laboratory. In return, undergraduates are demonstrating that, in parallel to in-depth study of a chosen major, they have the interest and ability to pursue a range of topics at the chemistry–biology interface.

REFERENCES

1. National Research Council Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century. *Bio 2010. Transforming Undergraduate Education for Future Research Biologists*, Washington, DC, 2003.
2. Godwin, H. A., and Davis, B. L. (2005) Teaching undergraduates at the interface of chemistry and biology: challenges and opportunities, *Nat. Chem. Biol.* **1**, 176–179.
3. Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154.
4. Merrifield, R. B. (1964) Solid phase peptide synthesis. II. The synthesis of bradykinin. *J. Am. Chem. Soc.* **86**, 304–305.
5. Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins, *Nat. Methods* **3**, 429–438.
6. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat. Chem. Biol.* **1**, 13–21.
7. Chen, I., and Ting, A. Y. (2005) Site-specific labeling of proteins with small molecules in live cells, *Curr. Opin. Biotechnol.* **16**, 35–40.
8. Momand, J., Wu, H. H., and Dasgupta, G. (2000) MDM2—master regulator of the p53 tumor suppressor protein, *Gene* **242**, 15–29.
9. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488–492.
10. Chene, P. (2003) Inhibiting the p53-MDM2 interaction: an important target for cancer therapy, *Nat. Rev. Cancer* **3**, 102–109.
11. Kritzer, J. A., Zutshi, R., Cheah, M., Ran, F. A., Webman, R., Wongjirad, T. M., and Schepartz, A. (2006) Miniature protein inhibitors of the p53-hDM2 interaction, *ChemBioChem* **7**, 29–31.

Spinal Cord Injuries: Solving the Enigma



www.yanla.com

About 400,000 Americans are living with spinal-cord injuries, according to the Christopher Reeve Foundation. Modern medicine has thus far offered no reliable treatments to regrow neurons or halt the auxiliary nerve damage and death that typically accompanies the initial injury. However, spinal-cord injuries offer an intriguing challenge to researchers interested in combining the disciplines of biology and chemistry. One such investigator is Molly Shoichet of the University of Toronto (U of T). For the past decade, Shoichet and her colleagues have worked to develop novel materials that could protect surviving neurons or spur regrowth after a neurological injury. Though she and other researchers have far to go before individuals can overcome their sometimes devastating disabilities, her approach offers a promising new way to combat this type of injury.

Combining Interests

Shoichet was born in Toronto in 1965 to parents who worked in business: her mother managed a picture-frame-manufacturing business, and her father ran an executive charter company. “My parents were always wonderfully supportive of the different things I’ve wanted to do,” says Shoichet. “They didn’t necessarily agree with all of them, but they supported my own choices.”

That encouraging attitude led her to pursue scientific research, an interest she began cultivating in high school. A particularly energetic teacher “made chemistry fun,” Shoichet says. Rather than use memorization to tackle the subject, the instructor led Shoichet and her classmates through numerous hands-on experiments. “What makes science fun now is discovery, and there was plenty of discovery in the

way he taught class,” she says. Shoichet notes that this approach also probably played a hand in one of her two brothers’ career choices. Brian Shoichet, who studied under the same high school chemistry teacher, is now a professor of pharmaceutical chemistry at the University of California, San Francisco.

By the time she graduated from high school, Molly Shoichet had decided to follow in Brian’s footsteps by attending Massachusetts Institute of Technology (MIT) in Cambridge. Early on, she chose to major in chemistry. Shoichet notes that chemistry offered a hidden advantage over other majors—the program required few classes, so she could fill her time pursuing other interests, such as biology. Taking classes in the life sciences sparked a curiosity for medicine, she says. However, many of her other chemistry classes—especially those in polymer chemistry—ignited her interest in basic research.

In a few settings, she was able to combine those two interests. For example, during her senior year, she worked with MIT materials scientist Ionnis Yannis on a new project he started to create polymers for nerve repair. However, by the time she graduated from college in 1987, Shoichet was unsure whether she wanted to ultimately pursue research or practice medicine. Unable to decide, she applied to both medical school and graduate school. That fall, she enrolled in a graduate program in polymer science at the University of Massachusetts, Amherst (UMass). “I thought I knew what a doctor did, but not what a scientist did,” Shoichet says. “Part of my rationale with pursuing a graduate degree was to try to understand that.” For the next 2 years, she deferred her acceptance to medical school, in the end deciding that basic research was her preferred path.

Profiles provide insights into the lives, backgrounds, career paths, and futures of scientists who serve as Experts on *ACS Chemical Biology’s* online Ask the Expert feature. Readers are encouraged to submit questions to the Experts at www.acschemicalbiology.org. The editors will post the most interesting exchanges on the website.

Published online August 22, 2006

10.1021/cb600332x CCC: \$33.50

© 2006 by American Chemical Society

After arriving at UMass, Shoichet was intent on pursuing basic polymer research. She found a valuable mentor in Tom McCarthy. Under McCarthy's guidance, Shoichet investigated different ways to modify polymer surfaces without affecting the bulk structure. She focused on fluoropolymers because they are inherently chemically inert, allowing the surface to be chemically modified with limited affect on the morphology or topology of the surface. Shoichet introduced carboxylic acids to fluoropolymer films (1) and then used these charged surfaces to control the adsorption of oppositely charged poly(L-lysine) (2). An important offshoot of this work was determining how changes in surface properties affected the behavior of cells cultured on polymer films—work that would guide her future career. She and McCarthy published several papers based on this research (3–7), and she completed her doctoral degree in 1992.

Industry and Innovation

As her time at UMass wound down, Shoichet was determined to find a job that combined medicine and chemistry. She pursued several strategies to track down the best opportunity. Contacting everyone on a list of speakers from a recent Gordon Conference eventually led her to CytoTherapeutics, based in Providence, RI. The company's main focus was encapsulated cell therapy: using cells, which were encased in polymer capsules to protect them from immune rejection, to deliver biologics to treat injuries and diseases. CytoTherapeutics, which is now a part of the Palo Alto-based Stem Cells, Inc., focused on treating neurological disorders such as Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis (also known as Lou Gehrig's disease).

"In conventional drug delivery, we're limited to how much drug we can give at one time, and drugs lose their bioactivity

over time. The idea [at CytoTherapeutics] was that if we deliver cells, they're constantly producing new therapeutic molecules. With that approach, we should have a ready supply of therapeutic molecules for a long time," she explains.

Shoichet found her new job "a wonderful opportunity," she says. Though CytoTherapeutics used polymer chemistry to create the capsules that surrounded cells, few polymer scientists worked at the company when she arrived. That gave her ample openings to carve a niche at the company: she worked with neuroscientists to evaluate the materials that would promote long-term survival and functioning of cells encapsulated within the hollow fiber membranes (8–10).

After nearly 3 years, Shoichet left CytoTherapeutics for personal reasons—her husband, Kevin Bartus, had finished his master's degree in business administration at Harvard University and secured a job with a consulting company in Toronto. This gave Shoichet and her husband an opportunity to move back near her family in Toronto. She took away a valuable lesson from her time at CytoTherapeutics. "I noticed that biologists would just take materials off the shelf and make them work," she says. "What became clear to me is that if I could understand the design criteria that they needed materials for, I could make materials specifically for biological and medical applications instead of just taking something off the shelf."

Seeking new opportunities once she and her husband relocated, Shoichet initially looked for industry jobs similar to the one she left. However, a prospect that intrigued her more was working at the U of T. Shoichet learned about an annual grant program awarded by the Natural Sciences and Engineering Research Council of Canada geared toward women interested in pursuing research in the physical sciences and engineering. Over the next

few months, she prepared a proposal that combined her interest in medicine with her growing expertise in polymer science. "I realized that spinal-cord injury was an area in which I could apply my knowledge," she says. "It's a big problem that's unanswered, and I thought that people in universities should work on these big problems."

Shoichet proposed using polymers to create nerve guidance channels for protecting and regrowing damaged neurons. She won the grant and was immediately offered a position at the U of T. She quickly became aware of the intricacies involved in treating spinal-cord injuries. "Every year, I understand a bit more how complicated it will be to come up with a solution," she says. "I know now why it's such a niche area."

The nervous system has a variety of mechanisms that seem to actively fight repair after injury, Shoichet explains. A cascade of cellular events releases biomolecules that cause damaged axons to degenerate further. Later, a glial scar walls off damaged tissue at the injury site. These complications make treating spinal-cord injuries difficult even immediately after the trauma occurs, and more difficult as time goes on.

Polymer Cure?

However, since she arrived at the U of T in 1995, she has focused her work on several approaches that offer renewed hope for treating spinal-cord injuries. For example, Shoichet and graduate students in her lab have collaborated closely with U of T neurosurgeon Charles Tator; they have developed an injectable hydrogel that could deliver neuron-saving drugs directly to the site of an injury.

For the past several years, Shoichet and her colleagues have tested prototype polymer gels on rat models of compression injuries, in which the spinal cord is

“What became clear is that if I could understand the design criteria needed, I could make materials specifically for biological and medical applications instead of just taking something off the shelf.”

crushed. The biggest initial concern when testing their new method was safety, says Shoichet. Hydrogels have never been used to treat spinal-cord injuries in this way. It was unknown whether this approach could have detrimental effects, such as lingering in the spinal column and blocking fluid flow, which can lead to increased pressure in the brain.

In 2003, she and her colleagues published the first paper showing the safety and feasibility of this method to localize drug delivery (11, 12). This year, Shoichet's lab provided evidence that a new hydrogel formulation provides some protection to neurons on its own, without additional drugs. “While we were really looking for safety, we saw that the new gel had its own benefits,” she says (13). She and her colleagues are currently investigating which drugs and biomolecules might further encourage cell survival and limit degeneration after compression injuries.

She and her colleagues are also pursuing methods to regrow severed axons by creating nerve guidance channels, or polymer tubes, that encourage axon regeneration by providing a permissive environment. Here, Shoichet and her team are testing their prototypes on a more severe injury model, fully transected rat spinal cords. Working again with Tator and also with U of T neural-stem-cell biologist Cindi Morshead, Shoichet and her lab are currently investigating strategies to add various factors to the channels to encourage axon regrowth. For example, the researchers are embedding signaling molecules that stimulate axon growth during development into polymer nerve guidance channels.

In ongoing studies, Shoichet's lab is investigating biomimetic approaches for axonal guidance. These strategies include creating scaffolds that guide axons with immobilized concentration gradients of neurotrophic factors (14, 15) and scaffolds that guide axons with cell adhesive peptides, within a 3D, photochemically patterned hydrogel (16, 17).

Recently, Shoichet's lab has turned its focus to solving another biomedical challenge—crafting materials for cancer-drug delivery. “It's a new area for us. We'd developed some really interesting polymers” in the course of developing the hydrogels and nerve guidance channels, says Shoichet. “We thought that the ideas we had in developing those polymers and some of their properties would be interesting for targeted cancer-drug delivery.”

Much of the lab's work in this area, adds Shoichet, has centered on copolymers that they created with a random distribution of hydrophilic and hydrophobic components that are also charged. She and her colleagues found that these components can self-aggregate to a diameter of <100 nm, an ideal size for cancer-drug delivery. Shoichet's team is now experimenting with modifying the surface of these polymers with antibodies that could help direct them to their targets.

Constantly developing new ways to apply her knowledge of polymer chemistry to medical problems keeps Shoichet on her toes—and will keep her busy for the rest of her career, she anticipates. “We're always trying to learn and be innovative and expand our horizons and purview,” she says. “We'll always be changing and advancing knowledge, and hopefully moving towards new therapeutic applications.”

—Christen Brownlee, Science Writer

REFERENCES

1. Shoichet, M. S., and McCarthy, T. J. (1991) Convenient syntheses of carboxylic acid functionalized fluoropolymer surfaces, *Macromolecules* 24, 982–986.
2. Shoichet, M. S., and McCarthy, T. J. (1991) Surface modification of poly(tetrafluoroethylene-co-hexafluoropropylene) film by adsorption of poly(L-lysine) from aqueous solution, *Macromolecules* 24, 1441–1442.
3. Zawodzinski, T. A., Jr., Gottesfeld, S., Shoichet, M. S., McCarthy, T. J. (1993) The contact angle between water and the surface of perfluorosulfonic acid membranes, *J. Appl. Electrochem.* 23, 86–88.
4. Bee, T. G., Cross, E. M., Dias, A. J., Lee, K.-W., Shoichet, M. S., McCarthy, T. J. (1992) Control of wettability of polymers using organic surface chemistry, *J. Adhes. Sci. Technol.* 6, 719–731.
5. Shoichet, M. S., and McCarthy, T. J. (1991) Poly(L-lysine) adsorption to fluoropolymer films, in *Materials Synthesis Based on Biological Processes* (Alper, M., Rieke, P. C., Frankel, R., Calvert, P. D., Tirrell, D. A., Eds.) Materials Research Society Symposia Proceedings Vol. 218, pp 57–68. Materials Research Society, Pittsburgh, PA.
6. McCarthy, T. J., Bee, T., Brennan, J., Cross, E., Franchina, N. L., Dias, A. J., Lee, K.-W., and Shoichet, M. S. (1995) Polymer surface modification, in *Organic Thin Films and Surfaces: Directions for the Nineties, Volume 1* (Ulman, A., Ed.), Academic Press, San Diego.
7. Bee, T. G., Dias, A. J., Franchina, N. L., Kolb, B. U., Lee, K.-L., Patton, P. A., Shoichet, M. S., and McCarthy, T. J. (1993) Surface chemistry of chemically resistant polymers, in *Polymer Surfaces and Interfaces II*, pp 1–25, John Wiley and Sons, New York.
8. Shoichet, M. S., Winn, S. R., Athavale, S., Harris, J. M., and Gentile, F. T. (1994) Poly(ethyleneoxide)-grafted thermoplastic membranes for use as cellular hybrid bio-artificial organs in the central nervous system, *Biotechnol. Bioeng.* 43, 563–571.
9. Shoichet, M. S. and Rein, D. H. (1996) *In vivo* biostability of a polymeric hollow fiber membrane for cell encapsulation, *Biomaterials* 17, 285–290.
10. Shoichet, M. S., Li, R. H., White, M. L., and Winn, S. R. (1996) Stability of hydrogels used in cell encapsulation: An *in vitro* comparison of alginate and agarose, *Biotechnol. Bioeng.* 50, 374–381.
11. Jimenez Hamann, M. C., Tsai, E. C., Tator, C. H., and Shoichet, M. S. (2003) Novel intrathecal delivery system for treatment of spinal cord injury, *Exp. Neurol.* 182, 300–309.
12. Jimenez Hamann, M. C., Tator, C. H., and Shoichet, M. S. (2005) Injectable intrathecal delivery system for localized administration of EGF and FGF-2 to the injured rat spinal cord, *Exp. Neurol.* 194, 106–119.
13. Gupta, D., Tator, C. H., and Shoichet, M. S. (2006) Fast-gelling injectable blend of hyaluronan and methylcellulose for intrathecal, localized drug delivery to the injured spinal cord, *Biomaterials* 27, 2370–2379.
14. Kapur, T. A., and Shoichet, M. S. (2004) Immobilized concentration gradients of nerve growth factor guide neurite outgrowth, *J. Biomed. Mater. Res.* 68A, 235–243.
15. Moore, K., MacSween, M., and Shoichet, M. S. (2006) Immobilized concentration gradients of neurotrophic factors guide neurite outgrowth of primary neurons in macroporous scaffolds, *Tissue Eng.* 27, 505–518.
16. Luo, Y., and Shoichet, M. S. (2004) A photolabile hydrogel for guided 3-D growth and migration, *Nat. Mater.* 3, 249–253.
17. Musoke, P., and Shoichet, M. S. (2006) Anisotropic three-dimensional peptide channels guide neurite outgrowth within a biodegradable hydrogel matrix, *Biomed. Mater.* 1, 162–169.

From Green Bacteria to Human Dementia: A Novel Model for Discovering Amyloid Assembly Inhibitors

Ehud Gazit*

Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Tel Aviv 69978, Israel

The formation of amyloid assemblies, which are large protein aggregates that share biophysical, biochemical, and ultrastructural properties, is associated with >20 human disorders (1, 2). Despite the amyloid-forming proteins' different origins and the lack of any simple homology among them, in all cases fine 7–10 nm fibrils are observed that have a predominant β -sheet structure and a strong birefringence upon staining with Congo red. One of the most notable properties of amyloid fibrils is their effect on the fluorescence of small aromatic dyes such as thioflavin. When the dye binds to mature amyloid fibrils, its fluorescence markedly increases. Therefore, the ability to inhibit amyloid fibril formation, as reflected by the thioflavin fluorescence, is based on the use of a major assay for the high-throughput screening of inhibitors.

Amyloid fibril formation has been associated with some of the most common and grave degenerative disorders, including Alzheimer's disease (AD), Parkinson's disease (PD), and type 2 diabetes. However, despite continuous efforts by academic groups and the pharmaceutical industry, currently no approved therapeutic agents can control the process of amyloid formation and reverse the degenerative symptoms observed in these disorders. Despite the screening of large chemical libraries that may contain hundreds or thousands of compounds that use the thioflavin assay, as mentioned above, no direct amyloid

assembly inhibitors have been developed into clinical therapeutic agents. The drugs available merely treat the symptoms rather than halt or reverse the course of the disease.

One of the reasons for the lack of approved efficacious agents that can inhibit amyloid formation and improve the prognoses of patients suffering from amyloid diseases may be because of the exact molecular character of the pathological species. In recent years, accumulating evidence has indicated that soluble amyloid assemblies rather than the mature amyloid plaques may actually be the pathological culprits that result in the degenerative phenomena (Figure 1) (3–13).

This rather novel notion is most established in the case of AD. Researchers demonstrated that soluble amyloid oligomers specifically affect memory-related functions, manifested by impairment of the long-term potentiation (LTP), a cerebral synaptic activity that is associated with learning and memorizing processes (4–6). The shift in the paradigm allows us to better understand some of the apparent inconsistencies in the "amyloid hypothesis". This includes the lack of signs of dementia in some people who possess large levels of amyloid deposits. In one famous case, a nun, Sister Mary, had very good cognitive abilities before her death at 101, despite having a very significant amount of amyloid deposits (14). The soluble oligomer notion also provides an explanation for the observation of LTP and

ABSTRACT The formation of amyloid assemblies is associated with major human disorders. Yet no therapeutic agents presently exist to control this process. In a recent paper, a new bacterial system is described that uses a fusion of the Alzheimer's disease β -amyloid polypeptide to the GFP. The assay detects the formation of small, soluble amyloid intermediates associated with degenerative diseases. This assay allows the researchers to use high-throughput screening methods to find inhibitors of the formation of amyloid assemblies.

*Corresponding author,
ehudg@post.tau.ac.il.

Published online August 18, 2006

10.1021/cb600328c CCC: \$33.50

© 2006 by American Chemical Society

Fluorescence could be detected only if an inhibitor allows the correct folding and if the fusion protein exists in a nonaggregative state.

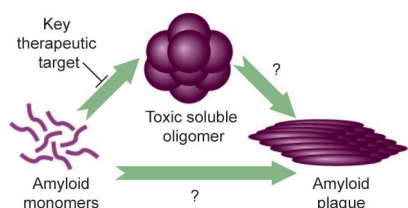


Figure 1. Soluble oligomers appear to be a major pathological element that facilitates the cognitive impairment associated with AD. It is not clear whether the formation of the oligomers is on-pathway or off-pathway in terms of the formation of the mature plaques. The sensitivity of a bacterial system to the formation of small aggregates will allow the identification of novel inhibitors to elucidate this key biochemical step.

cognitive deficiencies in AD model mice well before any amyloid deposits could be detected (13).

Specific, stable, soluble, dodecamer amyloid species of ~56 kDa were independently isolated and characterized by two different groups (15, 16). These species are extremely stable in solution and do not disintegrate in denaturing SDS polyacrylamide gel electrophoresis (15). These stable species affect the LTP activity in rat brain slices and live rodents and cause memory impairment upon intracranial injection into rats (15, 16). Most intriguingly, the very same dodecamer soluble species, which were termed A β *56, were found in the brains of AD patients and therefore can serve as novel biomarkers for the onset of AD long before the large fibrils can be observed by any imaging technique. The importance of the early soluble oligomers may be even greater given the many hints emerging that these assemblies may actually play a role in a condition known as mild cognitive impairment (MCI). The MCI disorder involves minor memory loss but not total dementia, is much more common than AD, and may affect >20 million Americans. Yet most likely, MCI may actually lead to the dreadful AD dementia.

The article by Kim *et al.* (17) on page 461 of this issue of *ACS Chemical Biology* offers a new way to select for amyloid formation inhibitors by using a bacterial screen, with special emphasis on early soluble assemblies (Figure 2). The assay is based on the expression of the AD β -amyloid (A β) polypeptide fused to a GFP in *Escherichia coli* bacteria. The authors chose to fuse the A β 42 polypeptide, an A β variant that contains 42 amino acids and is considered to be the most important pathological form (compared with the shorter A β 40 peptide). The assay is based on a reduction in the level of fluorescence that is observed upon the aggregation of the fused A β 42 protein in the bacterium, as was previously reported by the same group (18). This occurs because of the slow formation of the fluorophore *in vivo*. Therefore, the fluorescence of the fusion protein depends on its correct folding and solubility. Fluorescence could be detected only if an inhibitor allows the correct folding and if the fusion protein exists in a nonaggregative state. The assay eliminates false-positive hits, which can prevent aggregation, but it allows proteins to fold correctly because of their effect on β -sheet structures in general. This is because inhibition of aggregation is only one essential factor needed to obtain a fluorescence signal. The other crucial element is the correct folding of the GFP protein. Therefore, a positive fluorescence signal can be observed only if the two events occur (17).

A major advantage of the assay is the fact that a decrease in fluorescence occurs already upon the formation of small oligomers, which are so important for the pathology of AD and other amyloid diseases, as previously mentioned. The bacterial assay also has other inherent advantages, because the bacteria represent a physiological environment under molecular crowding conditions and other effects that are difficult to replicate in a test tube. Furthermore, the bacterial assay allows researchers to study protein stability in a pro-

teolytic environment and their ability to cross biological membranes. The latter property is more relevant to intracellular protein-aggregation phenomena, as described below, and less so to extracellular AD assemblies. In order not to miss agents that are effective inhibitors of A β aggregation but do not traverse well across biological membranes, the authors are also developing a cell-free assay that will be based on *in vitro* transcription and translation of the fusion protein.

The *E. coli* system appears to be very cost-effective. The authors used a simple, 96-well-plate rapid test to screen a library of ~1000 compounds. Members of the triazine chemical scaffold were added to the bacteria that expressed the fusion protein, and a conventional plate reader was used to determine the level of fluorescence in each well. The authors reported that the assay allowed them to distinguish between similar chemical entities and thus allows the pharmacophore definition of the identified molecules. An interesting observation that was reported is the similarity between the identified active compounds and one developed by Selkoe *et al.* (19) to control the formation of soluble oligomers.

The novel bacterial assay, as described in the new article (18), also represents a major

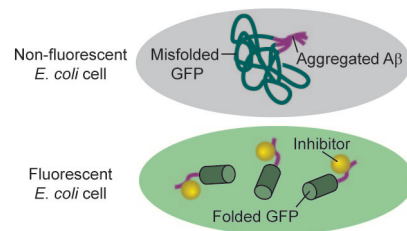


Figure 2. The bacterial systems used to select aggregation inhibitors. Fluorescence could be observed only if the GFP–A β fusion protein maintains its soluble structure and if correct folding of the GFP moiety occurs. The assay is sensitive enough for the aggregation of a small number of monomers. Thus, inhibitors of the early nucleation step may be identified.

advantage given some of the limitations of synthetic A β polypeptides. The synthetic A β polypeptide is relatively difficult to synthesize and thus very costly, a significant limitation to how much screening can be done. Furthermore, because of its exceptional tendency to aggregate, it is almost impossible to attain a synthetic preparation of A β that does not include any pre-nucleated seeds. The number may vary between batches, but some seed contamination always occurs. This not only results in large variability between different experiments but also may impair the ability to select for the most important inhibitors, those of the early oligomers that may represent the nucleation seeds. Biologically expressed A β can also be easily used to study mutated forms of A β , such as altered polypeptides associated with the early onset of the disease. Actually, some of the mutations in the A β polypeptide, most notably the Arctic mutation that was identified in families in northern Sweden, are known to result in a greater tendency to form prefibrillar amyloid assemblies as well as associated LTP impairment (20).

The current work paves the way for new developments in the field that will involve other amyloid-associated diseases as well as protein-aggregation disorders in general. The most challenging ones are those associated with intracellular protein-aggregation events. The author discusses Huntington's disease, which involves the aggregation of polyglutamine-rich huntingtin protein and the prion disorders, which may be hereditary, spontaneous, or infective encephalopathies. The latter form is the well-known bovine spongiform encephalopathy or "mad cow" disease. Two other key diseases are PD, in which the formation of intracellular amyloid Lewy bodies is linked to motor dysfunction, and amyotrophic lateral sclerosis (commonly known as Lou Gehrig's disease), which is associated with the intracellular aggregation of the superoxide dismutase enzyme.

REFERENCES

1. Rochet, J. C., and Lansbury, P. T., Jr. (2000) Amyloid fibrillogenesis: themes and variations, *Curr. Opin. Struct. Biol.* **10**, 60–68.
2. Sacchettini, J. C., and Kelly, J. W. (2002) Therapeutic strategies for human amyloid diseases, *Nat. Rev. Drug. Discovery* **1**, 267–275.
3. Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T., Jr. (2002) Neurodegenerative disease: amyloid pores from pathogenic mutations, *Nature* **418**, 291.
4. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation *in vivo*, *Nature* **416**, 535–539.
5. Klein, W. L. (2002) Abeta toxicity in Alzheimer's disease: globular oligomers (ADDLs) as new vaccine and drug targets, *Neurochem. Int.* **41**, 345–352.
6. Kirkitadze, M. D., Bitan, G., and Teplow, D. B. (2002) Paradigm shifts in Alzheimer's disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies, *J. Neurosci. Res.* **69**, 567–577.
7. Porat, Y., Kolusheva, S., Jelinek, R., and Gazit, E. (2003) The human islet amyloid polypeptide forms transient membrane-active prefibrillar assemblies, *Biochemistry* **42**, 10971–10977.
8. Kaye, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis, *Science* **300**, 486–489.
9. Williams, A. D., Segal, M., Chen, M., Kheterpal, I., Geva, M., Berthelie, V., Kaleta, D. T., Cook, K. D., and Wetzel, R. (2005) Structural properties of Abeta protofibrils stabilized by a small molecule, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7115–7120.
10. Gazit, E. (2004) The role of prefibrillar assemblies in the pathogenesis of amyloid diseases, *Drugs Future* **29**, 613–619.
11. Bitan, G., Fradinger, E. A., Spring, S. M., and Teplow, D. B. (2005) Neurotoxic protein oligomers—what you see is not always what you get, *Amyloid* **12**, 88–95.
12. Watson, D., Castano, E., Kokjohn, T. A., Kuo, Y. M., Lyubchenko, Y., Pinsky, D., Connolly, E. S., Jr., Esh, C., Luehrs, D. C., Stine, W. B., Rowse, L. M., Emmerling, M. R., and Roher, A. E. (2005) Physicochemical characteristics of soluble oligomeric Abeta and their pathologic role in Alzheimer's disease, *Neurol. Res.* **27**, 869–881.
13. Jacobsen, J. S., Wu, C. C., Redwine, J. M., Comery, T. A., Arias, R., Bowlby, M., Martone, R., Morrison, J. H., Pangalos, M. N., Reinhart, P. H., and Bloom, F. E. (2006) Early-onset behavioral and synaptic deficits in a mouse model of Alzheimer's disease, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5161–5166.
14. Snowdon, D. A. (1997) Aging and Alzheimer's disease: lessons from the Nun Study, *Gerontologist* **37**, 150–156.
15. Barghorn, S., Nimmrich, V., Striebinger, A., Krantz, C., Keller, P., Janson, B., Bahr, M., Schmidt, M., Bitner, R. S., Harlan, J., Barlow, E., Ebert, U., and Hillen, H. (2006) Globular amyloid beta-peptide oligomer—a homogenous and stable neuropathological protein in Alzheimer's disease, *J. Neurochem.* **95**, 834–847.
16. Lesne, S., Koh, M. T., Kotilinek, L., Kaye, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006) A specific amyloid-beta protein assembly in the brain impairs memory, *Nature* **440**, 352–357.
17. Kim, W., Kim, Y., Min, J., Kim, D. J., Chang, Y.-T., and Hecht, M. H. A. (2006) High-Throughput Screen for Compounds That Inhibit Aggregation of the Alzheimer's Peptide, *ACS Chem. Biol.* **1**, 461–469.
18. Kim, W., and Hecht, M. H. (2005) Sequence determinants of enhanced amyloidogenicity of Alzheimer A β 42 peptide relative to A β 40, *J. Biol. Chem.* **280**, 35069–35076.
19. Walsh, D. M., Townsend, M., Podlisy, M. B., Shankar, G. M., Fadeeva, J. V., Agnaf, O. E., Hartley, D. M., and Selkoe, D. J. (2005) Certain inhibitors of synthetic amyloid beta-peptide (Abeta) fibrillogenesis block oligomerization of natural Abeta and thereby rescue long-term potentiation, *J. Neurosci.* **25**, 2455–2462.
20. Klyubin, I., Walsh, D. M., Cullen, W. K., Fadeeva, J. V., Anwyl, R., and Selkoe, D. J. (2004) Soluble Arctic amyloid beta protein inhibits hippocampal long-term potentiation *in vivo*, *Eur. J. Neurosci.* **19**, 2839–2846.

Rescuing Defective Vesicular Trafficking Protects against α -Synuclein Toxicity in Cellular and Animal Models of Parkinson's Disease

Hilal A. Lashuel^{†,*} and Harald Hirling[‡]

[†]Laboratory of Molecular Neurobiology and Neuroproteomics, and [‡]Laboratory of Cellular Neurobiology, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Station 15, CH-1015 Lausanne, Switzerland

ABSTRACT Studies in yeast are providing critical insights into the mechanisms of neurodegeneration in Parkinson's disease (PD). A recent study shows that disruption of vesicular trafficking between the endoplasmic reticulum (ER) and the Golgi, caused by the overexpression and/or aggregation of α -synuclein, is linked to degeneration of dopamine neurons. Overexpression of proteins that are known to enhance ER-to-Golgi transport rescue defective trafficking in yeast, worm, fly, and cellular models of PD.

The failure of proteins to fold correctly or to remain folded is the primary cause of several systemic and neurodegenerative diseases that affect a significant portion of the world's population (1). To protect against aberrant folding, living organisms have evolved efficient protein synthesis and quality-control machinery. This system relies on close cooperation between the chaperone and the protein-degradation machinery to ensure control over proper folding, targeting, and degradation of proteins.

Protein folding occurs in the cytosol or within the lumen of the endoplasmic reticulum (ER). Proteins that are destined to go through the secretory pathway must pass through a series of quality-control checkpoints to ensure their correct processing and targeting to the extracellular space, the plasma membrane, or their final destination within the cell (Figure 1). The first checkpoint along this pathway resides in the ER, where soluble proteins are translated into the ER lumen and transmembrane proteins are translated and integrated into the ER membrane. Concomitant to translation, a specialized set of chaperones and the quality-control machinery of the ER ensure correct folding. In addition, post-translational modifications such as disulfide bonds or N-linked glycosylations are introduced. Once a protein is properly folded and modified and has passed the quality-control

process, it is directed to a specialized ER exit site, where it is integrated into transport vesicles that bud from the ER membrane. These vesicles are anterogradely transported to the Golgi, where they dock and fuse with the membrane of the cis-Golgi (see ref 2 for a detailed description of trafficking at the ER and Golgi apparatus). The major function of the Golgi is to fine-tune the added sugar residues and to sort the different cargo proteins into particular vesicles to be transported along distinguished trafficking pathways. These include routes to the plasma membrane for secretion and insertion of surface proteins or to other intracellular compartments such as the endosomal/lysosomal system. Each transport step also engages in a retrograde transport activity. This is particularly important for transport between the ER and the Golgi in order to recapture components of the vesicle-trafficking machinery and ER resident proteins.

However, this process is not perfect, and many newly synthesized proteins misfold and rapidly degrade. Proteins that fail to fold properly are retrotranslocated at the level of the ER to the cytosol for degradation by the ubiquitin proteasome system (UPS). If the ER folding machinery and the UPS cannot keep up with protein misfolding, the accumulation and/or aggregation of misfolded proteins induces cellular stress by multiple mechanisms. The result is cellular dysfunction.

*Corresponding author,
hilal.lashuel@epfl.ch.

Published online August 18, 2006

10.1021/cb600331e CCC: \$33.50

© 2006 by American Chemical Society

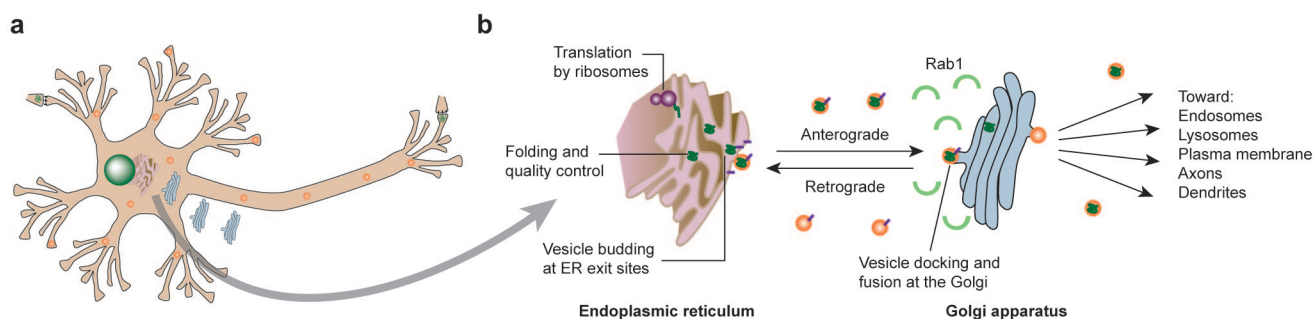


Figure 1. Vesicular trafficking in neurons. a) The extended morphology of neurons showing the cell body, dendrites, axons, and two synapses that are the contact sites with upstream or downstream cells in the network. b) Like all eukaryotic cells, neurons possess a set of organelles that form the secretory pathway. Protein synthesis takes place on ribosomes that associate with the ER membrane. Chaperones assist in proper folding of these newly translated polypeptide chains. A complex quality-control machinery that recognizes misfolded proteins allows only correctly folded proteins to reach ER exit sites, where transport vesicles containing the cargo proteins are formed. These vesicles are transported to the cis-side of the Golgi, where they dock and fuse with its membrane. This process is dependent on the small GTPase Rab1. After further protein modification, cargo proteins are packaged into specific vesicles that take on diverse transport routes in the cell.

tion, the initiation of ER-induced apoptosis, and ultimately disease manifestation.

Several misfolding diseases are caused by mutations that result in the loss of protein function due to improper folding, trafficking, and/or enhanced intracellular degradation by the UPS (*e.g.*, cystic fibrosis, sickle-cell anemia, α -1-antitrypsin deficiency, familial hypercholesterolemia, and some forms of cancer) (1). If the rate of protein misfolding is faster than that of degradation because of mutations and/or impaired quality-control machinery within the cell, then the misfolded proteins accumulate and self-associate to form highly ordered β -sheet-rich toxic aggregates. The presence of aggregates of misfolded protein in the form of intracellular inclusions or extracellular deposits in the vicinity of dying neurons is a defining hallmark of several neurodegenerative diseases (NDDs), including Alzheimer's disease (AD), Parkinson's disease (PD), amyotrophic lateral sclerosis, and polyglutamine and prion diseases (3).

Increasing evidence from neuropathologic, genetic, animal modeling, biochemical, and biophysical sources points toward protein misfolding and aggregation as the primary cause of several NDDs. However, the exact mechanisms by which these pro-

cesses cause neurodegeneration and cell death remain a subject of intense investigation and debate. Studies on cellular and animal models of protein-aggregation diseases suggest that the pathogenesis is complicated, and it is likely that neurodegeneration occurs by more than one mechanism. Oxidative stress, membrane disruption, ER stress, altered chaperone activity, impairment of the UPS, mitochondrial deficit, transcriptional dysregulation, axonal transport abnormalities, and Golgi fragmentation are all possible consequences of protein aggregation and are thought to play key roles in the initiation and/or progression of neurodegeneration.

Yeast Sheds Light on Neurodegeneration in PD. The discovery of disease-associated mutations in the genes encoding the aggregating proteins inspired the development of genetic animal and cellular models as tools for understanding the relationship between protein aggregation and disease. The existing genetic models of protein aggregation diseases are all based on the massive overexpression of the gene coding for the wild-type protein or disease-associated mutants in mouse, rat, *Drosophila*, and *Caenorhabditis elegans*. These models recapitulate some features of the disease, but none has

been shown to reproduce the complete disease phenotype observed in humans.

Baker's yeast (*Saccharomyces cerevisiae*), a single-celled organism with <6000 genes, was once thought to be too simple for modeling complex pathologies of the nervous system, but it is now emerging as a powerful tool for modeling NDDs. Despite significant differences between yeast cells and neurons, many of the basic cellular processes, such as protein folding and the quality-control machinery, are conserved in both eukaryotic cells. In a recent study published in *Science Express*, Antony Cooper (University of Missouri, Kansas City), Susan Lindquist's team (Massachusetts Institute of Technology), and colleagues from several other research groups (4) took advantage of these similarities by using a yeast model to elucidate a new molecular mechanism underlying the pathogenesis of PD. This neurodegenerative movement disorder is characterized by the loss of dopamine (DA) neurons from the *substantia nigra* (SN) and the formation of intraneuronal proteinaceous inclusions, referred to as Lewy bodies (LBs).

α -Synuclein Aggregation and Defective Trafficking. Lindquist and colleagues (4, 5) created a yeast model of PD based on

Colocalization of Ypt1p with α -synuclein suggests that α -synuclein toxicity may involve the sequestration of proteins that play a critical role in the ER-to-Golgi transport.

increased expression of the presynaptic protein α -synuclein, the primary constituent of LBs. Expression of mutant α -synuclein or the wild-type protein is sufficient to cause familial PD (6–9). α -Synuclein aggregation and fibrillogenesis are also implicated in the pathogenesis of several NDDs, including AD, multiple-system atrophy, dementia with LBs, Down syndrome, and neurodegeneration with brain iron accumulation, collectively referred to as synucleinopathies (10).

In the yeast system, the expression of α -synuclein (wild-type or disease-associated mutant A53T) can be tightly experimentally regulated and its effects monitored in real time; thus, detection of the early events involved in α -synuclein toxicity is possible (4). Within the first 4–8 h, the presence of α -synuclein aggregates and increased ER stress were observed to coincide with growth arrest and loss of cell viability. ER and proteasome-specific substrates were used to show that the expression of α -synuclein does not affect the general proteasome activity. However, it significantly impairs the turnover of substrates for which degradation requires trafficking from the ER to the Golgi as well as the transport of proteins that traffic through this pathway. Detailed dissection of the early events occurring during the first 4 h demonstrated that the first detectable defects in cell growth coincide with the impairment of vesicular transport from the ER to the Golgi and occur before the induction of ER stress.

Rescuing Defective Vesicular Trafficking.

Next, Lindquist and colleagues performed a complementation screen for modifiers of α -synuclein toxicity, which identified 34 genes that suppressed α -synuclein toxicity and 20 genes that increased it. Many suppressor genes that were specific for α -synuclein toxicity encode proteins that are also involved in ER-to-Golgi transport. If toxicity from increased α -synuclein levels and/or aggregation occurs through disruption of the ER-to-Golgi transport machinery, then promoting the forward transport from the ER to

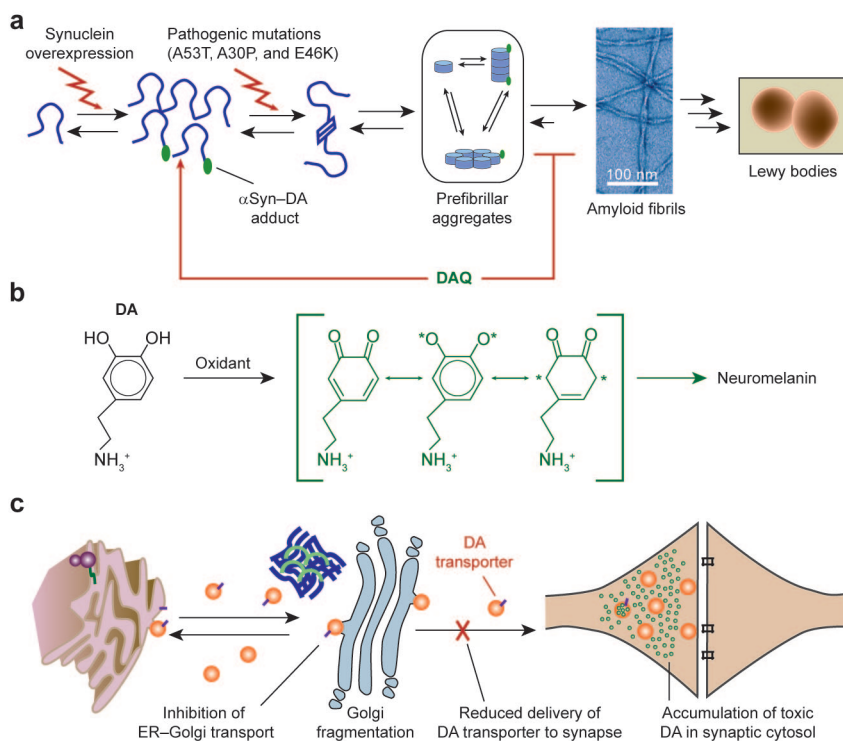


Figure 2. Potential toxic mechanisms linking protein aggregation, defective trafficking, and selective degeneration of DA neurons in PD. a) Schematic depiction of the current understanding of the aggregation pathway of α -synuclein based on *in vitro* biophysical studies. b) DA oxidation results in the production of reactive oxygen species and quinone and semiquinone intermediates, all of which are highly cytotoxic. To protect against the toxic properties of DA metabolites, the majority of DA is stored in vesicles before its release. c) The study by Cooper *et al.* shows that α -synuclein aggregates interfere with a Rab1-dependent step of ER-to-Golgi transport. This could lead to fragmentation of the Golgi due to an imbalance of incoming and outgoing vesicles, reduced targeting of DA transporters to synaptic vesicles in the presynaptic terminal, and, as a consequence, reduced uptake of DA into vesicles and an accumulation of DA in the cytosol. Formation of toxic DAQ intermediates has been shown to covalently modify α -synuclein and enhance α -synuclein toxicity through the kinetic stabilization of toxic prefibrillar aggregates.

the Golgi should reverse α -synuclein toxicity. Indeed, this was the case. This was particularly striking with Ypt1p and Rab1 (the human homologue of Ypt1p). In yeast, *Drosophila*, *C. elegans*, and rat DA neurons, the overexpression of Ypt1p/Rab1 resulted in significant reduction of α -synuclein-induced neurodegeneration.

The small GTPases of the Rab family play essential roles in vesicle docking and fusion (Figure 2, panel c) (11). Rab1 has been

shown to be specifically involved in ER-to-Golgi trafficking and the docking of ER-derived transport vesicles at the Golgi membrane. This suggests that the transport step affected by α -synuclein toxicity is vesicle docking/fusion at the Golgi membrane. The colocalization of Ypt1p with α -synuclein in cytosolic inclusion suggests that α -synuclein toxicity may involve the sequestration of proteins that play a critical role in the ER-to-Golgi transport and, eventually, the

disruption of the secretory pathway. Further biochemical and biophysical characterization of the α -synuclein aggregates could provide important insights into the mechanism of α -synuclein toxicity in yeast. Whether α -synuclein has a physiological role in ER-to-Golgi trafficking or transport of synaptic vesicles remains to be determined.

α -Synuclein-Induced Disruption of ER-to-Golgi Transport. Given the essential role of the Golgi apparatus in the processing and targeting of proteins through the secretory pathway, any disruption at this level is likely to have detrimental consequences for the function of the cell. Furthermore, a delicate balance between anterograde and retrograde membrane traffic through the Golgi is critical to avoid its fragmentation. A link between α -synuclein aggregation and Golgi fragmentation is supported by previous findings demonstrating that formation of α -synuclein aggregates, particularly prefibrillar aggregates that precede LB formation (Figure 2, panel a), causes fragmentation of the Golgi in cellular models of synucleinopathies (12) as well as in nigral neurons of PD patients (13). Further studies are required to elucidate the exact mechanisms by which α -synuclein aggregation disrupts ER-to-Golgi transport and cause Golgi fragmentation.

Defective Vesicular Trafficking and DA Neurons. Although α -synuclein is an abundant protein in different parts of the brain (up to 1% of total proteins), α -synuclein aggregation in PD occurs primarily in DA neurons of the SN. The wide distribution of α -synuclein in the brain suggests that α -synuclein on its own cannot explain the selective degeneration of DA neurons in PD.

Overexpression of human α -synuclein in transgenic flies (14) or specifically in the SN of rats (15) and primates (16) results in selective DA neuronal death and in the formation of α -synuclein-containing inclusions. Overexpression of PD-linked α -synuclein mutations in human mesencephalic cell lines leads to an impaired storage and secretion of DA, causing an increase in cytosolic DA and

enhanced oxidative stress (17, 18). Xu *et al.* (19) reported that blocking DA synthesis in cultured DA neurons prevents α -synuclein toxicity, consistent with toxicity being mediated by interactions between the two molecules. These observations are consistent with the known hypersensitivity of DA neurons that express high levels of cytoplasmic DA (*e.g.*, the SN rather than the ventral tegmental area) to cell death in PD.

The selective vulnerability of DA neurons of the SN to α -synuclein toxicity in PD may be related to the toxicity and increased concentration of cytoplasmic DA in these cells; this suggests that improper packaging, secretion, and/or oxidation of DA might explain the selective degeneration of DA neurons (17, 18). DA oxidation and formation of DA orthoquinone (DAQ) *in vitro* covalently modifies α -synuclein and results in kinetic stabilization of toxic α -synuclein aggregates (20). Therefore, the simultaneous contribution of several factors, including α -synuclein oligomerization, DA metabolism, and oxidative stress, might be required for selective degeneration of DA in the SN of PD brains. Disruption of vesicular trafficking is likely to hasten cell death by simultaneously increasing the levels of cytosolic DA and enhancing α -synuclein aggregation.

The findings by Cooper and colleagues offer new insight into the mechanisms by which α -synuclein overexpression and/or aggregation interferes with normal function and viability of neurons and reveal new targets for therapeutic intervention in PD and related synucleinopathies. In addition to being a good model with which to study genetic diseases, the yeast system is also demonstrated by these researchers to offer an excellent platform on which to screen for druglike molecules as modifiers of α -synuclein function(s), aggregation, and toxicity. The identification of specific pharmacological agents that suppress α -synuclein-induced ER-to-Golgi trafficking defects and prevent or reverse neurodegeneration in mouse models of PD as well as in clinical studies is the ulti-

mate proof of the therapeutic potential of these findings.

Acknowledgment: We thank Ruth Luthi-Carter for critical reading of the article.

REFERENCES

- Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein-misfolding diseases, *Nature* 426, 905–909.
- Lee, M. C., Miller, E. A., Goldberg, J., Orci, L., and Schekman, R. (2004) Bi-directional protein transport between the ER and Golgi, *Annu. Rev. Cell Dev. Biol.* 20, 87–123.
- Ross, C. A., and Poirier, M. A. (2005) Opinion: What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol.* 6, 891–898.
- Cooper, A. A., Gitler, A. D., Cashikar, A., Haynes, C. M., Hill, K. J., Bhullar, B., Liu, K., Xu, K., Strathearn, K. E., Liu, F., Cao, S., Caldwell, K. A., Caldwell, G. A., Marsischky, G., Kolodner, R. D., Labaer, J., Rochet, J. C., Bonini, N. M., and Lindquist, S. (2006) Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models, *Science* 313, 324–328.
- Outeiro, T. F., and Lindquist, S. (2003) Yeast cells provide insight into alpha-synuclein biology and pathobiology, *Science* 302, 1772–1775.
- Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E. S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W. G., Lazzarini, A. M., Duvoisin, R. C., Di Iorio, G., Golbe, L. I., and Nussbaum, R. L. (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease, *Science* 276, 2045–2047.
- Kruger, R., Kuhn, W., Muller, T., Woitalla, D., Graeber, M., Kosel, S., Przuntek, H., Epplen, J. T., Schols, L., and Riess, O. (1998) Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease, *Nat. Genet.* 18, 106–108.
- Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., Lincoln, S., Crawley, A., Hanson, M., Maraganore, D., Adler, C., Cookson, M. R., Muentzer, M., Baptista, M., Miller, D., Blancato, J., Hardy, J., and Gwinn-Hardy, K. (2003) Alpha-synuclein locus triplication causes Parkinson's disease, *Science* 302, 841.
- Zarranz, J. J., Alegre, J., Gomez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atares, B., Llorens, V., Tortosa, E. G., Del Ser, T., Munoz, D. G., and De Yébenes, J. G. (2004) The new mutation, E46K, of alpha-synuclein causes parkinson and Lewy body dementia, *Ann. Neurol.* 55, 164–173.
- Trojanowski, J. Q., and Lee, V. M. (2003) Parkinson's disease and related alpha-synucleinopathies are brain amyloidoses, *Ann. N.Y. Acad. Sci.* 991, 107–110.
- Zerial, M., and McBride, H. (2001) Rab proteins as membrane organizers, *Nat. Rev. Mol. Cell Biol.* 2, 107–117.

12. Gosavi, N., Lee, H. J., Lee, J. S., Patel, S., and Lee, S. J. (2002) Golgi fragmentation occurs in the cells with prefibrillar alpha-synuclein aggregates and precedes the formation of fibrillar inclusion, *J. Biol. Chem.* *277*, 48984–48992.
13. Fujita, Y., Ohama, E., Takatama, M., Al-Sarraj, S., and Okamoto, K. (2006) Fragmentation of Golgi apparatus of nigral neurons with alpha-synuclein-positive inclusions in patients with Parkinson's disease, *Acta Neuropathol.* DOI: 10.1007/s00401-006-0114-4.
14. Feany, M. B., and Bender, W. W. (2000) A Drosophila model of Parkinson's disease. *Nature* *404*, 394–398.
15. Lo Bianco, C., Ridet, J. L., Schneider, B. L., Deglon, N., and Aebischer, P. (2002) Alpha-synucleinopathy and selective dopaminergic neuron loss in a rat lentiviral-based model of Parkinson's disease, *Proc. Natl. Acad. Sci. U.S.A.* *99*, 10813–10818.
16. Kirik, D., Annett, L. E., Burger, C., Muzyczka, N., Mandel, R. J., and Bjorklund, A. (2003) Nigrostriatal alpha-synucleinopathy induced by viral vector-mediated overexpression of human alpha-synuclein: a new primate model of Parkinson's disease, *Proc. Natl. Acad. Sci. U.S.A.* *100*, 2884–2889.
17. Lotharius, J., and Brundin, P. (2002) Pathogenesis of Parkinson's disease: dopamine, vesicles and alpha-synuclein, *Nat. Rev. Neurosci.* *3*, 932–942.
18. Lotharius, J., and Brundin, P. (2002) Impaired dopamine storage resulting from alpha-synuclein mutations may contribute to the pathogenesis of Parkinson's disease, *Hum. Mol. Genet.* *11*, 2395–2407.
19. Xu, J., Kao, S. Y., Lee, F. J., Song, W., Jin, L. W., and Yankner, B. A. (2002) Dopamine-dependent neurotoxicity of alpha-synuclein: a mechanism for selective neurodegeneration in Parkinson disease, *Nat. Med.* *8*, 600–606.
20. Conway, K. A., Rochet, J. C., Bieganski, R. M., and Lansbury, P. T., Jr. (2001) Kinetic stabilization of the alpha-synuclein protofibril by a dopamine-alpha-synuclein adduct, *Science* *294*, 1346–1349.

Lighting Up the Nascent Cell Wall

Wilfred A. van der Donk*

Department of Chemistry, University of Illinois at Urbana–Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

The manner by which eubacteria assemble their cell walls has been intensely investigated for more than half a century, primarily because this process is targeted by many important antibiotics, including the β -lactams, vancomycin, fosfomycin, nisin, and bacitracin (1). The eubacterial cell wall consists of layers of peptidoglycan (PG) polymer (also termed murein) made up of alternating disaccharides composed of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) (Figure 1) that are cross-linked through short peptides attached to the lactyl group of each MurNAc (2). This cross-linking by penicillin-binding proteins (PBPs) generates a netlike, 3D structure that provides the cell with the strength to withstand intracellular pressures of several atmospheres. Whereas the cytoplasmic biosynthesis of lipid II, the precursor for the polymerization of PG, is well understood (3), the mechanism by which this monomer is inserted into existing PG in a growing or dividing cell is still largely unknown. For the details of PG biosynthesis to be understood, knowledge about the subcellular localization of the biosynthetic machinery and its substrates is essential. In a recent issue of *Proceedings of the National Academy of Sciences, U.S.A.*, Walker and co-workers (4) provide a powerful and general tool for investigating the spatial distribution of sites of nascent PG biosynthesis.

Although the molecular details are not known, ramoplanin binds in the region of MurNAc pyrophosphate found both in lipid II and in the reducing end of the growing glycan polymer (5), where new lipid II mol-

ecules are inserted through a transglycosylation reaction catalyzed by the high-molecular-weight PBPs (Figure 1). Walker *et al.* (4) used microscopy and a series of fluorescently labeled ramoplanin analogues to visualize the sites of new PG biosynthesis. The staining patterns observed showed fluorescent patches at the hemispherical cell poles and at new cell-division septa at midcell of the rod-shaped bacterium *Bacillus subtilis*. In addition, helical staining patterns were observed along the cylindrical side wall of the cell (Figure 2). A similar helical pattern along the longitudinal cell axis had been reported previously by Daniel and Errington (6), who used fluorescently labeled vancomycin. This molecule binds to the *D*-Ala-*D*-Ala segment present in lipid II as well as along PG chains that have not been fully cross-linked by transpeptidases nor hydrolyzed by *D*,*D*-carboxypeptidases (Figure 1) (2). On the basis of the respective recognition sites for ramoplanin and vancomycin, one might have expected that the former would be more specific for staining initiation sites of nascent PG biosynthesis; however, similar (although not identical) patterns were in fact observed when either probe was used (4). One advantage of ramoplanin staining over vancomycin staining is that presumably because of its higher affinity for its target, lower concentrations of reagents could be used to visualize the coiled patterns. Because both molecules are antibiotics, the use of lower concentrations could be important to ensure that the reagents themselves do not influence the distribution of sites of nascent PG biosynthesis. The observation that cell poles are

ABSTRACT Many antibiotics target the assembly of the cell wall of eubacteria, a netlike 3D structure composed of layers of peptidoglycan (PG). Very little is known about how the lipid precursor of PG, lipid II, is inserted into the existing cell wall in a growing and dividing cell. A new study provides a powerful tool for investigating this insertion process and opens the door to understanding the mechanism of eubacterial cell wall biogenesis.

*Corresponding author,
vddonk@uiuc.edu.

Published online August 18, 2006

10.1021/cb600308w CCC: \$33.50

© 2006 by American Chemical Society

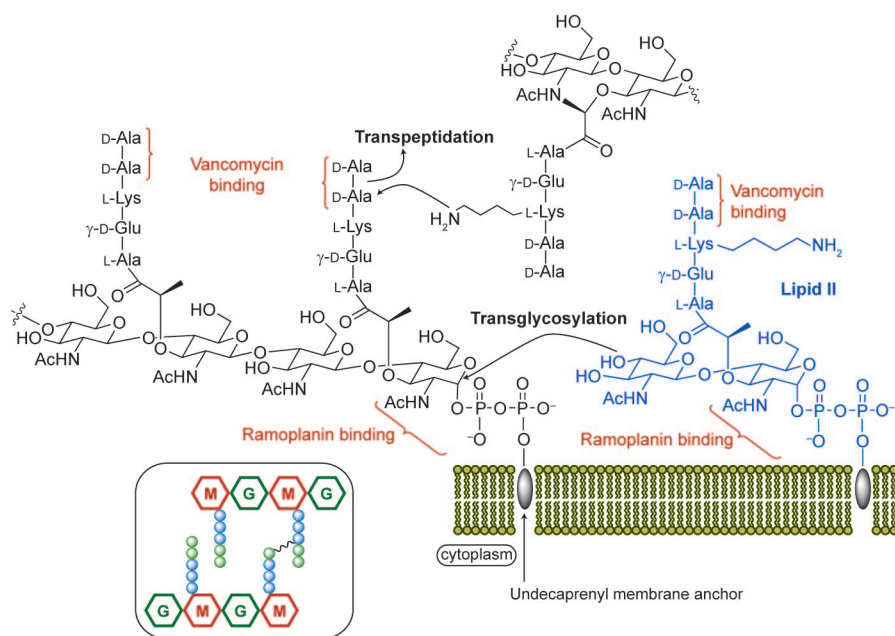


Figure 1. Schematic representation of new PG biosynthesis. The structure of lipid II is shown in blue. It is inserted into existing PG via a transglycosylation reaction. After incorporation, the side-chain amine of a lysine (or in certain cases, a diaminopimelic acid) of one nascent chain is cross-linked with another PG chain by a transpeptidation reaction in which the terminal D-Ala of one of the MurNAc-linked pentapeptides is displaced. The binding sites of vancomycin, which primarily inhibits transpeptidation, and ramoplanin, which inhibits transglycosylation, are indicated. Inset: Nascent PG after one transpeptidation cross-link. M = MurNAc, G = GlcNAc, blue spheres are the first three amino acids of pentapeptide, and green spheres represent D-Ala-D-Ala. In mature PG, the terminal D-Ala can be removed as a result of transpeptidation or alternatively by proteolysis by a D,D-carboxypeptidase.

stained by ramoplanin, with its selectivity for binding to initiation sites, suggests that the poles are not inert and that PG biosynthesis persists.

The detection of helical localization of the biosynthetic substrates for PG biosynthesis is very intriguing, especially because it had been long thought that PG biosynthesis occurred in a dispersed fashion along the cylindrical cell of rod-shaped bacteria. It raises the million-dollar question: what positions the substrate lipid II and/or the biosynthetic machinery along spirals wrapping the cylindrical cell membrane? Interestingly, the helical distribution of sites of new PG biosynthesis comes at a time when more and more studies report similar sublocalizations of proteins in bacteria. Recent years have seen the identification of bacterial homologues of the major eukaryotic cytoskeletal proteins. MreB is the bacterial homologue of actin, FtsZ is the homologue of tubulin, and Crescentin is the homologue of intermediate filament (7). All three are distributed at certain times during bacterial cell cycles in helicoid patterns. Because the

shape of the PG layer is a critical determinant of bacterial morphology (2), many mutants identified in various nonspherical bacteria on the basis of a change of shape (*rod mutants*) have provided candidate proteins that may determine the spatial organization of the machinery for new PG biosynthesis. Among these mutants are genes in the *mre* gene cluster (murein cluster e). Like actin, MreB assembles into filaments that form large fibrous spirals in the cytoplasm, just under the cell membrane of the rod-shaped bacteria *B. subtilis* (8, 9) and *Escherichia coli* (10), as well as at the start of the cell cycle of the crescent-shaped organism *Caulobacter crescentus* (11, 12). Visualization of these structures by immunofluorescence microscopy or GFP-fusion imaging demonstrates that in *B. subtilis* and *C. crescentus* the spiral consists of three or four turns along the length of the cell, whereas in *E. coli* the helix forms one or two turns. Interestingly, Pbp2, whose gene is in the same operon as *mreB*, is also organized along helical bands encircling the cell in *C. crescentus*. This spiral arrangement of Pbp2 is

dependent on MreB, and Pbp2 coimmunoprecipitated with several other PBPs; this suggests it is part of a multienzyme PG biosynthetic complex (11). Bacterial cell-wall biosynthesis in rod-shaped bacteria is thought to take place in two stages, one involving cell division at the septum and a subsequent stage in which elongation of the side wall takes place (2). Pbp2 is essential for the latter process and is a membrane protein with a cytoplasmic and periplasmic domain that catalyzes cross-linking of PG by transpeptidation. Collectively, these data suggest a model in which the intracellular MreB helix might position the Pbp2 spiral, thus serving as a spatial and temporal cytoskeletal scaffold directing PG biosynthesis during cell elongation. However, several other studies suggest a more complex picture. The helical positioning of fluorescently labeled vancomycin reporting on the location of nascent PG biosynthesis was *not* dependent on MreB (6). Furthermore, during the cell cycle of *C. crescentus*, the coiled distribution of MreB dynamically changed to a midcell localization at the onset of cell divi-

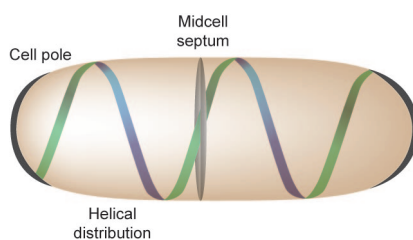


Figure 2. Schematic representation of the sites of nascent PG biosynthesis as revealed by vancomycin and ramoplanin staining patterns. When 3D deconvolution microscopy was used, bands were observed at the cell poles and the division septa and patterns consistent with helices were also seen. These spirals run along the cylindrical part of the rod-shaped cell. This figure is not representative of the intensities of the staining or the details of the patterns, which vary according to conditions. For the actual fluorescent microscopy images, see refs 4 and 6.

sion, whereas Pbp2 remained organized in spirals (11, 13).

If the biosynthetic complex for PG biosynthesis does not appear to track along MreB filaments, how does it then become organized along helices during cell elongation? The membrane proteins MreC and MreD are essential for determining a rodlike shape and also form helical cables around the longitudinal axes of various bacteria (13–15). Moreover, MreC is required for the spiral distribution of PG precursors in *B. subtilis* (15). Interestingly, simultaneous visualization of cytosolic MreB and periplasmic MreC in *C. crescentus* showed that their helices do not colocalize and also that MreC helix formation is independent of MreB (13). Thus, transfer of the information about localization from the cytoplasm to the periplasm does not appear to be achieved through a direct physical interaction between these two protein helices. On the other hand, the mere fact that the two coils do not overlay suggests that some form of communication must occur. The Pbp2 helix did partially overlap with the MreC spiral, and in the absence of functional MreC or MreB, newly synthesized Pbp2 mislocalized to the divi-

sion plane. This mislocalization was shown to be dependent on the tubulin homologue FtsZ (13), which forms a ring at midcell in the earliest known event of cell division and then recruits all other proteins associated with cytokinesis (16). In cells containing chemically inactivated MreB and that were depleted of FtsZ, Pbp2 once again was localized along helical patterns. To explain these observations, Theriot *et al.* (13) recently hypothesized that MreB prevents Pbp2 accumulation at the divisional site and that MreC actively promotes helical localization along the cylindrical side wall. When either is absent or inactivated, mislocalization of Pbp2 occurs. The reason that FtsZ depletion counteracts MreB inactivation in this model is that in the absence of FtsZ no PG precursors accumulate at midcell that can recruit Pbp2.

Unlike *C. crescentus*, many bacteria have >1 MreB homologue. *B. subtilis* contains three such proteins: MreB, Mbl (MreB-like), and MreBH. Like MreB, Mbl is important in determining cell shape and forms spiral patterns along the cylindrical part of the cell (8). The importance of these Mbl fibers for positioning new PG biosynthesis during cell elongation is currently under debate. Daniel and Errington (6) reported that the helical staining pattern with fluorescein-labeled vancomycin was abolished in *mbl* null mutants. On the other hand, Walker and co-workers (4) used either labeled vancomycin or ramoplanin to show a qualitatively similar pattern of helicoid staining of wild-type and *mbl*⁻ cells. The reasons for this discrepancy are not known at present.

To date most studies have looked at localization of Pbp2 by fluorescence imaging and at transpeptidation and transglycosylation sites using fluorescently labeled vancomycin and ramoplanin. However, PBPs themselves need not necessarily track along cytoskeletal filaments; in principle only one component of the multienzyme PG biosynthetic complex must be properly positioned for the entire complex to be correctly local-

ized. In the spherical bacterium *Staphylococcus aureus*, Pbp2 appears to be recruited to sites of new PG synthesis (division site) by the presence of its substrate (17). Thus, other means for proper spatial and/or temporal control over PG biosynthesis in rod-shaped bacteria could involve coordinated localization of lipid II in the outer leaflet of the cytoplasmic membrane. At present, the translocase that exports intracellularly synthesized lipid II to the outside of the membrane remains unidentified; hence, its subcellular localization is unknown. Another possibility suggested by Walker and co-workers (4) is based on the observed helical bands of proteins of the general secretory (Sec) machinery (18). The PBPs are exported by the Sec pathway, and hence the observed helical localization of SecA and SecY may play a role in determining the timing and place of PBP membrane insertion and therefore new PG biosynthesis.

In closing, the long-held view of bacteria as bags of relatively uniformly distributed biomolecules enclosed by cell walls and membranes was abandoned some time ago with the discovery of bacterial cell cycles and polarity. Now the mechanisms by which subcellular localization is governed are starting to emerge. In one example discussed here, recent years have seen major developments in our understanding of how the spatial and temporal coordination of PG biosynthesis is achieved. Although many questions still remain, given the rapid pace of progress in the area, the advances in the spatial and temporal resolution of fluorescence imaging, and the development of small molecules as additional tools, the near future is likely to see major new discoveries.

REFERENCES

- Walsh, C. T. (2003) *Antibiotics: Actions, Origins, Resistance* pp 1–49, ASM Press, Washington DC.
- Höltje, J. V. (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*, *Microbiol. Mol. Biol. Rev.* 62, 181–203.

3. van Heijenoort, J. (2001) Recent advances in the formation of the bacterial peptidoglycan monomer unit, *Nat. Prod. Rep.* **18**, 503–519.
4. Tiyanont, K., Doan, T., Lazarus, M. B., Fang, X., Rudner, D. Z., and Walker, S. (2006) Imaging peptidoglycan biosynthesis in *Bacillus subtilis* with fluorescent antibiotics, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11033–11038.
5. Walker, S., Chen, L., Hu, Y., Rew, Y., Shin, D., and Boger, D. L. (2005) Chemistry and biology of ramoplanin: a lipoglycopeptide with potent antibiotic activity, *Chem. Rev.* **105**, 449–476.
6. Daniel, R. A., and Errington, J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell, *Cell* **113**, 767–776.
7. Møller-Jensen, J., and Löwe, J. (2005) Increasing complexity of the bacterial cytoskeleton, *Curr. Opin. Cell Biol.* **17**, 75–81.
8. Jones, L. J., Carballido-Lopez, R., and Errington, J. (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*, *Cell* **104**, 913–922.
9. van den Ent, F., Amos, L. A., and Löwe, J. (2001) Prokaryotic origin of the actin cytoskeleton, *Nature* **413**, 39–44.
10. Shih, Y. L., Le, T., and Rothfield, L. (2003) Division site selection in *Escherichia coli* involves dynamic redistribution of Min proteins within coiled structures that extend between the two cell poles, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7865–7870.
11. Figge, R. M., Divakaruni, A. V., and Gober, J. W. (2004) MreB, the cell shape-determining bacterial actin homologue, co-ordinates cell wall morphogenesis in *Caulobacter crescentus*, *Mol. Microbiol.* **51**, 1321–1332.
12. Gitai, Z., Dye, N., and Shapiro, L. (2004) An actin-like gene can determine cell polarity in bacteria, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8643–8648.
13. Dye, N. A., Pincus, Z., Theriot, J. A., Shapiro, L., and Gitai, Z. (2005) Two independent spiral structures control cell shape in *Caulobacter*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18608–18613.
14. Divakaruni, A. V., Loo, R. R., Xie, Y., Loo, J. A., and Gober, J. W. (2005) The cell-shape protein MreC interacts with extracytoplasmic proteins including cell wall assembly complexes in *Caulobacter crescentus*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18602–18607.
15. Leaver, M., and Errington, J. (2005) Roles for MreC and MreD proteins in helical growth of the cylindrical cell wall in *Bacillus subtilis*, *Mol. Microbiol.* **57**, 1196–1209.
16. Errington, J., Daniel, R. A., and Scheffers, D. J. (2003) Cytokinesis in bacteria, *Microbiol. Mol. Biol. Rev.* **67**, 52–65.
17. Pinho, M. G., and Errington, J. (2005) Recruitment of penicillin-binding protein PBP2 to the division site of *Staphylococcus aureus* is dependent on its transpeptidation substrates, *Mol. Microbiol.* **55**, 799–807.
18. Campo, N., Tjalsma, H., Buist, G., Stepniak, D., Meijer, M., Veenhuis, M., Westermann, M., Müller, J. P., Bron, S., Kok, J., Kuipers, O. P., and Jongbloed, J. D. (2004) Subcellular sites for bacterial protein export, *Mol. Microbiol.* **53**, 1583–1599.

Bacterial Evolution by Intelligent Design

Stephen C. Winans*

Department of Microbiology, 360A Wing Hall, Cornell University, Ithaca, New York 14853

In the past decade, interest in engineering new strains of bacteria that are optimized to carry out processes of medical, agricultural, or industrial importance has exploded. This body of research, sometimes referred to as “synthetic biology”, draws on decades of fundamental studies about the molecular biology of bacteria and seeks to exploit this knowledge to fine-tune existing types of bacterial physiology or to create whole new types of physiology. An excellent example of this research was recently published in *Nature Biotechnology* that described the isolation of a mutation in the quorum-sensing regulator LuxR of *Vibrio fischeri*. The mutation blocks the detection of the native chemical signal of *V. fischeri* but allows detection of a new signal that is not detected by the wild-type protein (1). This study will have applications in bacterial engineering and will provide insight about bacterial genetics, the detection of environmental signals, structural biology, and protein evolution.

Many types of bacteria communicate via the release and detection of diffusible chemical signals. These chemicals, which can be thought of as bacterial pheromones, stimulate diverse behaviors, including bioluminescence; the horizontal transfer of DNA; the formation of biofilms; and the production of pathogenetic factors, antibiotics, and other secondary metabolites (2). Gram-positive bacteria typically communicate by using oligopeptide signals that are detected by two-component phosphorelay proteins (3), whereas proteobacteria generally signal via acyl-homoserine lactones (AHLs). Additional classes of bacterial pheromones have also

been described (4). In addition to using species-specific signals, some groups of bacteria appear to employ a universal signal (a bacterial Esperanto) to communicate intergenerally (5).

V. fischeri is a bioluminescent marine bacterium that symbiotically colonizes various species of fish and invertebrates, which in turn exploit bacterial luminescence for a variety of purposes (5). A protein called LuxI synthesizes 3-oxo-hexanoyl-L-homoserine lactone (OHHL), an AHL-type signal molecule (6), and a protein called LuxR is the signal sensor and a signal-dependent transcriptional activator of the luciferase operon (7, 8). As a population of *V. fischeri* cells grows, the concentration of OHHL increases as a function of cell-population density. When the concentration of OHHL reaches the micromolar range, its passive efflux from the cells becomes balanced by a passive influx; therefore, its intracellular concentration increases enough to bind to LuxR. LuxR–OHHL complexes bind luciferase promoters and activate their transcription (Figure 1, panel a).

Intercellular signaling is thought to allow an estimation of population densities, a phenomenon sometimes referred to as quorum sensing (9). Certain bacterial behaviors are appropriate only if carried out simultaneously by large numbers of bacterial cells (a quorum). Bacteria are thought to use the concentration of these chemical signals as an indication of population density. However, this view of signaling is somewhat facile and teleological. One author suggested that these molecules might be released by a single bacterium to detect and

ABSTRACT In a process called quorum sensing, bacteria produce and secrete certain signaling compounds (called autoinducers) that bind to receptors on other bacteria and activate transcription of certain genes. A clever genetic selection yields a new quorum-sensing transcriptional regulator that marches to the beat of a different drummer.

*Corresponding author,
scw2@cornell.edu.

Published online August 18, 2006

10.1021/cb6003417 CCC: \$33.50

© 2006 by American Chemical Society

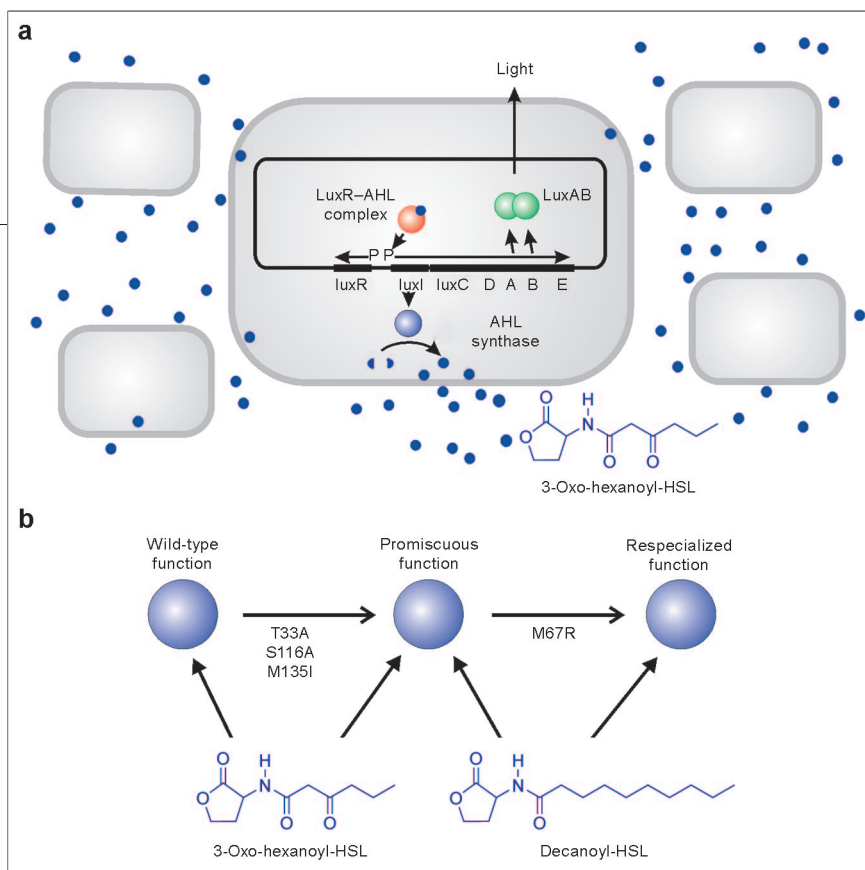


Figure 1. Communication in *V. fischeri*. **a**) Population-density-dependent expression of the luciferase operon of *V. fischeri*. LuxI proteins synthesize OHHL, which diffuses freely across the cell envelope. A combination of high cell density and a diffusion barrier causes these pheromones to accumulate and to bind to the LuxR transcription factor; this activates transcription of the luciferase operon, and bioluminescence results. **b**) Two-step evolution of a LuxR variant with altered AHL specificity. In the first step, a mutant was isolated that has a broadened specificity for AHLs; it detected both OHHL and decanoyl-HSL (as well as several other AHLs). In the second step, a mutant was isolated that detects the latter AHL but not the former.

measure diffusion barriers (10). According to this idea, these signal molecules can accumulate only if their diffusion is limited, and the bacterium might use this information, for example, to release hydrolytic enzymes or a matrix needed for biofilm formation. Both ideas are probably equally valid, because population density and diffusion barriers should both contribute to the accumulation of these signals. Equally plausible is that bacteria might require a quorum to effectively send and receive chemical signals; however, the goal of signaling may not be to measure the quorum but rather to coordinate the behavior of its members (11). In short, we should not pretend to understand the “why” of these bacterial signaling systems.

The study in *Nature Biotechnology* begins with a triple mutant in LuxR (referred to as G2E) that has the mutations T33A, S116A,

and M135I; this shows a relaxed specificity of AHL signal molecules (12). Whereas the native LuxR detects primarily OHHL, the G2E protein can detect a wide variety of similar signals. In the present study, Collins and coworkers started with the G2E mutation and used a clever sequential positive and negative selection for mutants that can detect decanoyl-homoserine lactone (HSL) but not OHHL. They constructed a *PluxI-chloramphenicol acetyl transferase (cat)* fusion to select for function in the presence of decanoyl-HSL. Expression of this fusion causes chloramphenicol resistance and was used to select for LuxR function in the presence of decanoyl-HSL. Colonies were pooled, and their DNA was extracted and introduced into a strain containing a *Plux-bit* fusion, where *bit* inhibits the activity of β -lactamase. Expression of Bit in the presence of OHHL was counterselected on a

medium that contained ampicillin. The surviving mutants all shared the same alteration, an arginine at position 67 in place of the wild-type methionine (M67R). It is unlikely that this mutation could readily have been isolated without this novel two-step selection. This LuxR variant can detect several AHLs with unsubstituted acyl groups but does not detect AHLs with 3-oxo substitutions.

Of the three alterations in mutant G2E, M135A is particularly interesting. Residue M135 lies at the same position as Ser126 of the TraR protein of *Agrobacterium tumefaciens*. The structure of TraR has been solved by X-ray crystallography (13). Ser126 of TraR lies close to Thr129, which makes a water-mediated hydrogen bond to the 3-oxo group of 3-oxo-octanoylhomoserine lactone (OOHL) (13). LuxR has a serine at position 137, which could play a similar role, and if so, the M135A mutant could destabilize this bond, thereby decreasing specificity for the 3-oxo group. Similarly, Met67 lies at the same position as Gln58 of the TraR. The β -carbon of Gln58 contacts the terminal carbon of the acyl chain of OOHL. This suggests that the β -carbon of Met67 of LuxR might hinder the binding of AHLs with acyl chains that are greater than six carbons, whereas the mutation M67R somehow relieves this block. This interpretation is consistent with the way the mutation was isolated (selection for detection of a long-chain unsubstituted AHL, and selection against detection of a short-chain 3-oxo-substituted AHL). How this mutation blocks the detection of the 3-oxo substitution is therefore puzzling. Perhaps future studies will solve this riddle.

As pointed out in the study, the isolation of the double mutant of LuxR may mimic the stepwise accumulation of mutations in nature. Several studies have shown that the first step in protein evolution often involves the acquisition of “promiscuous functions” that retain but broaden the ancestral functions (14). Later steps in evolution cause the protein to acquire “respecialized” proper-

ties. The original triple mutant caused the protein to detect a wide variety of AHLs, whereas the second mutation (M67R) restricted the detection of AHLs. Of course, the natural evolution of a new LuxR/LuxI-type regulatory system requires that both proteins coevolve such that the AHL synthase always makes a signal that the AHL receptor can detect. One AHL synthase and one AHL receptor were subjected to site-directed mutagenesis to create variants with new signal specificities. A T140A mutant of Esal (a LuxI homologue) preferentially synthesizes an unsubstituted AHL rather than a 3-oxo substituted AHL (15). Similarly, a T129A or T129V mutant of TraR detected unsubstituted AHLs with the same affinity as 3-oxo AHLs (16). Curiously, overproduction of either protein also reduces its AHL specificity. Overproduction of AHL synthases is thought to deplete the cell of the favored acyl-carrier protein substrate, and the enzyme then uses less favored ones (17). Overproduction of several AHL receptors dramatically broadens their substrate specificities, though the reason for this is not clear (18).

The creation of a new signal specificity will have interesting and unforeseen applications in synthetic biology. In previous studies, geneticists have separated signal synthesis and detection in two different bacterial strains, so that one strain detects a signal sent by another. The current study will enable reciprocal signaling, such that one strain releases a signal that is detected by a second strain, which then synthesizes a second signal that is detected only by the first strain. Reciprocal signaling could be useful in constructing oscillating circuits through negative feedback or could control the timing of sequential steps in the biosynthesis of a complex pharmaceutical. Creating similar systems with three or more signals should be possible. In addition to the construction of new variants of LuxR, many natural LuxR homologues have been discovered that detect diverse AHLs and

bind different DNA sequences. Researchers should be able to mix and match the binding domains of AHL and DNA of the various LuxR homologues to construct proteins that bind to new DNA sequences that are chosen by the intelligent bacterial designer.

Acknowledgment: Research in the author's laboratory is supported by a grant from the National Institutes of General Medical Sciences (GM42893).

REFERENCES

1. Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR, *Nat. Biotechnol.* **24**, 708–712.
2. Whitehead, N. A., Barnard, A. M., Slater, H., Simpson, N. J., and Salmond, G. P. (2001) Quorum-sensing in Gram-negative bacteria, *FEMS Microbiol. Rev.* **25**, 365–404.
3. Dunny, G. M., and Leonard, B. A. (1997) Cell-cell communication in Gram-positive bacteria, *Annu. Rev. Microbiol.* **51**, 527–564.
4. Pappas, K. M., Weingart, C. L., and Winans, S. C. (2004) Chemical communication in proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular signalling, *Mol. Microbiol.* **53**, 755–769.
5. Bassler, B. L., and Losick, R. (2006) Bacterially speaking, *Cell* **125**, 237–246.
6. Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Oppenheimer, N. J. (1981) Structural identification of autoinducer of photobacterium Fischeri luciferase, *Biochemistry* **20**, 2444–2449.
7. Engebrecht, J., and Silverman, M. (1984) Identification of genes and gene products necessary for bacterial bioluminescence, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4154–4158.
8. Engebrecht, J., Nealson, K., and Silverman, M. (1983) Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*, *Cell* **32**, 773–781.
9. Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994) Quorum sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators, *J. Bacteriol.* **176**, 269–275.
10. Redfield, R. J. (2002) Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol.* **10**, 365–370.
11. Winans, S. C., and Bassler, B. L. (2002) Mob psychology, *J. Bacteriol.* **184**, 873–883.
12. Collins, C. H., Arnold, F. H., and Leadbetter, J. R. (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones, *Mol. Microbiol.* **55**, 712–723.
13. Zhang, R. G., Pappas, T., Brace, J. L., Miller, P. C., Oulmassov, T., Molyneux, J. M., Anderson, J. C., Bashkin, J. K., Winans, S. C., and Joachimiak, A. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA, *Nature* **417**, 971–974.
14. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. McQ., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions, *Nat. Genet.* **37**, 73–76.
15. Watson, W. T., Minogue, T. D., Val, D. L., von Bodman, S. B., and Churchill, M. E. (2002) Structural basis and specificity of acyl-homoserine lactone signal production in bacterial quorum sensing, *Mol. Cell* **9**, 685–694.
16. Chai, Y., and Winans, S. C. (2004) Site-directed mutagenesis of a LuxR-type quorum-sensing transcription factor: alteration of autoinducer specificity, *Mol. Microbiol.* **51**, 765–776.
17. Gould, T. A., Herman, J., Krank, J., Murphy, R. C., and Churchill, M. E. (2006) Specificity of acyl-homoserine lactone synthases examined by mass spectrometry, *J. Bacteriol.* **188**, 773–783.
18. Zhu, J., Beaber, J. W., More, M. I., Fuqua, C., Eberhard, A., and Winans, S. C. (1998) Analogs of the autoinducer 3-oxooctanoyl-homoserine lactone strongly inhibit activity of the TraR protein of *Agrobacterium tumefaciens*, *J. Bacteriol.* **180**, 5398–5405.

Histone Citrullination by Protein Arginine Deiminase: Is Arginine Methylation a Green Light or a Roadblock?

Paul R. Thompson^{†*} and Walter Fast^{‡*}

[†]Department of Chemistry and Biochemistry, University of South Carolina, Columbia, South Carolina 29208, and [‡]Division of Medicinal Chemistry, College of Pharmacy, The University of Texas, Austin, Texas 78712

The discovery of nonribosomally encoded citrulline in proteins was first reported >40 years ago (1, 2), but the importance of this post-translational modification (PTM) to human physiology remained obscure until the 1990s, when two factors came together to bring this modification to prominence. The first factor was the discovery that rheumatoid arthritis (RA) patients produce autoantibodies targeting citrulline-containing epitopes and that these autoantibodies are a highly specific predictor of the disease (3). The second factor was the determination that histones contain citrulline residues; this finding suggested that this modification could affect gene transcription as a part of the “histone code” hypothesis (4–6). In addition, alterations of protein citrullination have been tentatively tied to the etiology of multiple sclerosis, psoriasis, glaucoma, various adenocarcinomas, and even bacterial infections by *Porphyromonas gingivalis* (7–11). The contributions of protein citrullination to some of these disease states have been reviewed elsewhere (12, 13). Recent efforts at the molecular, cellular, and whole-organism levels have begun to characterize the normal and pathophysiological roles of both protein citrullination and the enzymes responsible for catalyzing this modification: the protein (or peptidyl) arginine deiminases (PADs). In this review, we describe these efforts, placing particular emphasis on the role of one isozyme, protein arginine deiminase 4 (PAD4), in transcriptional regulation and its putative demethyliminase activity.

The PAD Family of Enzymes. The PADs, which are deiminating enzymes that hydrolyze guanidinium side chains to yield peptidylcitrulline and ammonia, belong to a larger group of guanidino-modifying enzymes called the amidinotransferase (AT) superfamily (14). Additional

ABSTRACT Protein citrullination, a once-obscure post-translational modification (PTM) of peptidylarginine, has recently become an area of significant interest because of its suspected role in human disease states, including rheumatoid arthritis and multiple sclerosis, and also because of its newfound role in gene regulation. One protein isozyme responsible for this modification, protein arginine deiminase 4 (PAD4), has also been proposed to “reverse” epigenetic histone modifications made by the protein arginine methyltransferases. Here, we review the *in vivo* and *in vitro* studies of transcriptional regulation by PAD4, evaluate conflicting evidence for its ability to use methylated peptidylarginine as a substrate, and highlight promising areas of future work. Understanding the interplay of multiple arginine PTMs is an emerging area of importance in health and disease and is a topic best addressed by novel tools in proteomics and chemical biology.

*Corresponding authors,
thompson@mail.chem.sc.edu,
waltfast@mail.utexas.edu.

Received for review June 1, 2006
and accepted July 3, 2006.

Published online August 18, 2006
10.1021/cb6002306 CCC: \$33.50

© 2006 by American Chemical Society

Intense efforts have been made to catalog the numbers and types of post-translational modifications that occur to histones . . .

members of this superfamily include the arginine deiminases (ADIs), enzymes found in both prokaryotes and the primitive eukaryote *Giardia intestinalis*, that act on nonpeptidyl arginine and that are involved in energy production; the dimethylarginine dimethylaminohydrolases (DDAHs), enzymes found in both bacteria and mammals that convert nonpeptidyl methylarginines into citrulline; the ATs, enzymes involved in both creatine and streptomycin biosynthesis; and the dihydrolases, bacterial enzymes involved in arginine catabolism that catalyze two successive hydrolytic steps. Whereas most superfamily members act on nonpeptidyl amino acids, PADs are highly specific for peptidylarginine residues and require at least one additional residue N-terminal to the site of modification (15). Although PAD and PAD-like enzymes are not universally conserved, they are present in several species, from bacteria to humans. Interestingly, the bacterial PADs are composed of only the ~40 kDa catalytic domain, whereas mammals have much larger multidomain enzymes (~75 kDa) whose activity is regulated by calcium. The extra domains in the mammalian enzymes, present in the N-terminal half of the protein, include two immunoglobulin-like domains that are proposed to mediate protein–protein interactions and/or substrate targeting (16). To date, five human PAD homologues have been identified. For historical reasons, these isozymes are designated PAD1–4 and PAD6. Human PAD4 was initially named PAD5 but was later renamed PAD4 to reflect the fact that it is a true ortholog of this isoform. PAD4 is distinguished by the insertion of a nuclear localization sequence and, in contrast to the cytoplasmic location of the other isoforms, is the only PAD localized to the cell nucleus (17).

Role of PAD4 in Transcriptional Regulation. Since the seminal finding that the histone acetylases and deacetylases (18, 19) are transcriptional coregulators, intense and continuing efforts have been made to catalog the numbers and types of PTMs that occur to histones on the premise that other histone-modifying enzymes might also influence gene transcription. Thus, the discovery that PAD4 is a nuclear enzyme that deiminates histones H2A, H3, and H4 (4) strongly suggested that it may also act as a transcriptional coregulator. A combination of proteomic techniques has determined that the *in vivo* sites of histone deimination occur at the N-terminal tails of histones H2A, H3, and H4 and, specifically, H2 Arg3, H3 Arg2, H3 Arg8, H3 Arg17, H3 Arg26, and H4 Arg3 (20–22).

Several independent studies have subsequently confirmed transcriptional coregulation by PAD4. For example, Wang *et al.* (22) transfected increasing amounts of a PAD4-encoding construct into MCF7 cells, a mammary carcinoma or breast-cancer cell line, and monitored the effects of this treatment on the transcription of an estrogen-responsive luciferase reporter construct. In these experiments, PAD4 clearly acts as a transcriptional corepressor in a dose-dependent manner. The catalytic activity of PAD4 is essential for this function because the corepressor function of a catalytically inactive mutant is significantly attenuated (22). Similarly, Cuthbert *et al.* (20) demonstrated that PAD4 fused to a zinc-finger DNA binding domain (PAD4-ZnDBD) could act as a transcriptional corepressor for the endogenous vascular endothelial growth factor-A (VEGF-A) promoter. For these experiments, the PAD4-ZnDBD was cotransfected with constructs encoding either the estrogen receptor or thyroid hormone receptor ligand binding domains fused to a different VEGF-A targeting ZnDBD, and transcription of the endogenous VEGF-A gene was monitored by real-time polymerase chain reaction (20). Importantly, the wild-type enzyme, but not a catalytically inactive mutant (a C-terminal truncation lacking amino acids 591–663), could repress the levels of transcription afforded by the thyroid hormone receptor construct, an indication that citrullination by PAD4 is essential for this effect.

Chromatin immunoprecipitation (ChIP) experiments further demonstrated a role for PAD4 in transcriptional regulation (20, 22). In these experiments, the levels of PAD4, various transcriptional cofactors, and specific histone modifications on the endogenous estrogen receptor responsive pS2 promoter were monitored at specific time points before and after the addition of estradiol. MCF7 cells were used for these experiments because PAD4 protein is expressed in this cell line in an estrogen-dependent manner (20, 22). These ChIP experiments demonstrated that PAD4 is constitutively associated with the pS2 promoter; its levels increased slightly 20–40 min after the addition of estradiol and then decreased thereafter to basal levels. Importantly, the amount of citrulline present in the N-terminal tails of histones H3 and H4 rose and then fell with similar kinetics. The levels of PAD4 and citrullinated histones on the pS2 promoter also correlated with its presumed function as a transcriptional corepressor because the levels of RNA polymerase II on this promoter were high when the

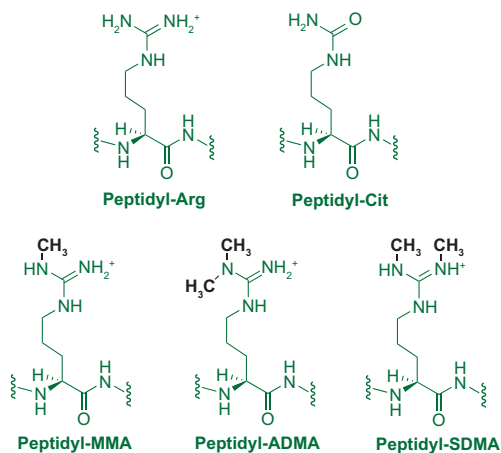


Figure 1. Arginine modifications.

levels of PAD4 and deiminated histones were low and *vice versa*. These results demonstrate that PAD4 is a transcriptional coregulator. The repressive effects of PAD4 on gene regulation are likely mediated by its ability to catalyze the deimination of specific residues present in the N-terminal tails of histones H3 and H4 (and possibly H2A) because histone modifications are generally known to affect transcription by altering the local chromatin structure either directly or indirectly *via* the recruitment of additional transcriptional coregulators (5, 23, 24). PAD4 has been shown to catalyze the citrullination of the glucocorticoid receptor interacting protein 1 binding domain in p300 (and antagonize methylation of this domain by coactivator-associated arginine methyltransferase 1, CARM1) (25). This suggests that, in addition to histone citrullination, the deimination of other transcriptional coregulators can contribute to the transcriptional corepressor function of this enzyme.

Does PAD4 Catalyze Demethyliminination? In addition to citrullination, arginine residues in the N-terminal tails of histones H3 and H4 are alternatively monomethylated and asymmetrically dimethylated on their guanidinium side chains. These methylation events are catalyzed by protein arginine methyltransferase 1 (PRMT1) and CARM1 (CARM1/PRMT4), which are type I PRMTs that catalyze the sequential formation of both N^{ω} -monomethylarginine (MMA) and asymmetric $N^{\omega},N^{\omega'}$ -dimethylarginine (ADMA) in an *S*-adenosylmethionine (SAM)-dependent manner (Figure 1). Note that PRMTs are generally classified as either type I or II enzymes. Type I

PRMTs (*e.g.*, PRMT1 and CARM1) catalyze the formation of peptidyl MMA and ADMA, whereas type II enzymes (*e.g.*, PRMTs 5 and 7) catalyze the formation of peptidyl MMA and symmetric $N^{\omega},N^{\omega'}$ -dimethyl-L-arginine (SDMA). Notably, the sites of deimination by PAD4 overlap with the sites of arginine methylation. For example, CARM1 methylates arginines 2, 17, and 26 in histone H3 (26–28), and PRMT1 methylates arginine 3 in histone H4 (29, 30). These PRMTs act as transcriptional coactivators for numerous transcription factors (*e.g.*, the estrogen receptor (31)), and their methyltransferase activity is required for this function. Generally, asymmetric dimethylation, not monomethylation, is associated with transcriptional activation (26, 27, 29, 30).

This overlap of methylation and deimination sites provided further support for the initial suggestion by Bannister *et al.* (6) that PAD4 might act to “reverse” the transcriptional coactivator function of the PRMTs by hydrolyzing methylated arginines to citrulline. PAD-catalyzed hydrolysis of methylated arginines, dubbed a “demethyliminination” reaction, would require the recognition of methylated arginines as substrates and would not truly reverse the modification but instead leave peptidylcitrulline as a final product. True reversibility by an arginine demethylase has not been discovered yet, but if PAD4 can remove the methyl mark introduced by the PRMTs, then it might serve an analogous function.

Several lines of evidence reported by Cuthbert *et al.* (20) and Wang *et al.* (22) suggest that, in addition to its established deiminase activity, PAD4 may also catalyze demethyliminination *in vivo*. For example, CHIP assays on the pS2 promoter show that the appearance of citrulline on both histone H3 and histone H4 correlates with decreasing levels of ADMA. Similarly, the levels of methylated histones H3 Arg17 and H4 Arg3 are reduced in HL 60 granulocytes, a cell line known to express PAD4, after treatment of this cell line with a calcium ionophore. The disappearance of these epitopes correlates with the appearance of citrulline on histones H3 and H4. Although these results appear to support an *in vivo* PAD4 demeth-

KEYWORDS

Activity-based protein profiling: A technique that uses small-molecule probes to detect enzyme activities in complex mixtures.

***S*-Alkylthiouonium:** A cationic *S*-alkylated thiourea functional group [(R-S(NR'₂)=NH₂)⁺] found as a covalent intermediate in the amidinotransferase superfamily.

Amidinotransferase: An enzyme capable of transferring an amidino group (–C(NH₂)=NH₂) from a donor substrate to an acceptor substrate. Also, a superfamily of enzymes that can catalyze group transfer or hydrolysis reactions at the iminium carbon of selected guanidines.

Autoantibodies: Antibodies produced by an organism, such as humans, that recognize host antigens and can elicit an inappropriate immune response.

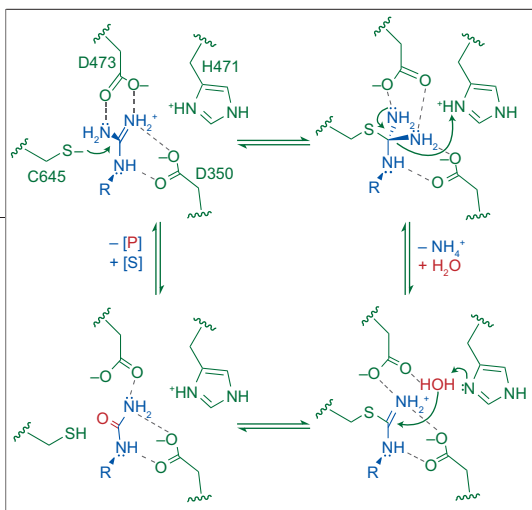


Figure 2. Proposed chemical mechanism for PAD4. See the text for details.

ylmination activity, alternative explanations exist. For example, deimination might antagonize methylation (substrate depletion), and deimination of other sites within histones H3 and H4 might prevent antibody recognition (epitope occlusion). These alternatives would not necessarily require PAD4 to deiminate methylated arginine residues.

Although the *in vivo* evidence described above supporting a physiologically relevant demethyliminase activity is compelling, the *in vitro* studies of this putative activity are less so and, in many cases, are contradictory. For example, recombinant histones H3 and H4 methylated by CARM1 and PRMT1, respectively, with [^3H]-SAM used as the methyl donor and presumably bearing MMA, were demethylated by PAD4 with the concomitant production of methylamine (22). In sharp contrast, several independent laboratories have provided compelling evidence that PAD4 cannot catalyze efficient demethyliminase *in vitro*. For example, Kearney *et al.* (15) synthesized several small-molecule and peptide substrates containing MMA, ADMA, and SDMA and then evaluated their ability to act as PAD4 substrates. While these compounds were demethylated by PAD4, the rates at which they are processed are much slower than the rates of deimination observed for the corresponding peptidylarginines, with observed rate reductions on the order of 100- to 10,000-fold (15). The *in vitro* rates of PAD4-mediated demethyliminase are ~ 1000 -fold slower than the rate at which arginine residues in histone H3 are methylated by CARM1 (28); this makes PAD4 a relatively inefficient demethyliminase. Similar results have been reported by three groups: Hidaka *et al.* (32) utilized HPLC and mass spectrometry (MS)-based assays to demonstrate that benzoylated methyl arginine derivatives are neither good substrates nor good inhibitors of PAD4; Cuthbert *et al.* (20) showed that ADMA- and SDMA-containing peptides were not demethylated by PAD4; and Wang *et al.* (22) established that histone H4-based peptides methylated by PRMT1 (and presumably bearing MMA) were poor substrates for PAD4 *in vitro*. These *in vitro* studies strongly suggest that PAD4 does not possess a physi-

ologically relevant demethyliminase activity and are in apparent conflict with the *in vivo* data.

Structure and Mechanism behind PAD Activity and Specificity. Recent structural and mechanistic studies of PAD4 and other members of the AT superfamily provide a framework for understanding how PAD4 catalyzes deimination and how this enzyme can achieve a 100- to 10,000-fold kinetic discrimination between substrates that differ by as little as one methyl group.

One possible chemical mechanism for PAD4-catalyzed deimination is proposed below (Figure 2). The substrate guanidinium is held in place by an intricate hydrogen-bonding network formed between this group and two active-site aspartyl groups (Asp350 and Asp473; PAD4 numbering is used unless otherwise indicated) that are conserved in both PADs and ADIs. These interactions help to position the guanidinium for nucleophilic attack by a conserved active-site cysteine nucleophile (Cys645 in PAD4), resulting in formation of an initial tetrahedral adduct between Cys645 and the guanidinium carbon. Collapse of this adduct to form a covalent planar S-alkylthiuronium intermediate then proceeds with the concomitant release of the first product, ammonia. Cleavage of this C–N bond likely requires concomitant protonation, and His471 is well-positioned to serve as a general acid for this step. The resulting planar S-alkylthiuronium intermediate is subsequently hydrolyzed to release the final product, citrulline. The hydrolytic water required for this step is potentially deprotonated to hydroxide by a general base such as His471, which is well-positioned to serve this function. Alternatively, the participation of ammonia (product-assisted catalysis) has been proposed to play a role in hydroxide formation because PAD4 does not possess a suitably positioned amino acid to improve the basicity of His471, whereas other AT superfamily members do possess such a residue (15). A substrate-assisted mechanism has also been proposed (33). Although detailed studies of PAD4 catalysis have been limited, key aspects of the proposed mechanism have been confirmed, including the identity and stoichiometry of the reaction products (15, 32), the incorporation of solvent ^{18}O into citrulline (15, 32), the structural conservation of the active site, and the essential nature of conserved active site residues, for example, Cys645, Asp350, Asp473, and His471, for activity (16).

This proposed mechanism is also supported by structural and mechanistic studies on PAD4 and similarities

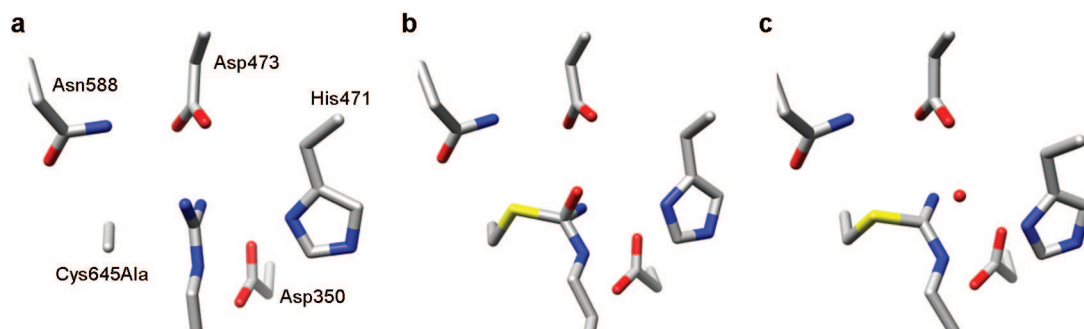


Figure 3. Structures of putative deiminase reaction intermediates. **a)** Human C645S PAD4 complexed with a protected arginine substrate mimicking the proposed Michaelis complex. **b)** *Mycoplasma arginini* ADI complexed with a proposed covalent tetrahedral reaction intermediate. **c)** *M. arginini* ADI complexed with an sp^2 -hybridized covalent reaction intermediate showing an ordered water (red sphere) poised for the subsequent hydrolytic step. PDB structures 1WDA, 1LXY, and 1S9R, respectively, were used to create the figure panels.

with other hydrolytic members of the AT superfamily, such as ADI and DDAH (33–39). For example, the extensive hydrogen-bonding network of the substrate guanidinium with Asp350 and Asp473 can easily be seen in the structure of the PAD4(C645S)–*N*α-benzoyl-L-arginine amide complex, which mimics the initial enzyme–substrate complex (Figure 3, panel a). A snapshot of the initial tetrahedral species has also been observed in structures of the ADI–arginine complex (Figure 3, panel b) (33, 34). While this tetrahedral adduct was originally assigned as a covalent complex formed by the back reaction with citrulline (34), later functional analyses suggested that this structure most likely depicts the species preceding formation of the planar thiouronium intermediate (37). The lifetime of this tetrahedral species does not appear to contribute significantly to the kinetic mechanism of ADI (38), but its structure does strongly suggest that a conserved active-site histidine is appropriately positioned to serve as a general acid to donate a proton to the leaving group. This proposed role is consistent with the elevated pK_a (7.9) assigned to His471 of PAD4 (15) and also with studies of ADI and DDAH (16, 33, 35–37). Collapse of the tetrahedral adduct and the elimination of ammonia lead to the formation of a planar sp^2 -hybridized thiouronium intermediate that is known to have a significant lifetime in the kinetic mechanisms of both ADI and DDAH (33, 35–37). Structures of the thiouronium intermediate bound to ADI are available (Figure 3, panel c) and demonstrate that the tight hydrogen-bonding network between the substrate and the two active-site aspartyl groups is conserved throughout the reaction (33, 34). The structure of

the ADI–thiouronium intermediate also reveals an ordered water molecule that is poised to attack the thiouronium intermediate, passing through a similar tetrahedral species (not shown) and generating citrulline during the hydrolytic half-reaction. Both His471 and Asp473 are well-placed to assist this step by deprotonating the water to provide a nucleophilic hydroxide, but additional studies will be required to determine the actual general base for this step. It should also be emphasized that further study will be required to determine the subtle differences in mechanism between ADI, DDAH, and PAD that result in their drastically different pH-rate profiles (15, 36, 37).

While the molecular details regarding PAD4 catalysis are still being deciphered, the structural and mechanistic data obtained to date for PAD4 and the related enzymes ADI and DDAH provide insight into how PAD4 is able to discriminate between substrates that differ by only one methyl group. First, it is unlikely that methylation contributes intrinsically to the stability of these residues toward hydrolysis. N^ω -Methylation only has small effects on the overall charge, electrophilicity, and leaving-group stability of these substrates, as demonstrated by the small variance in pK_a values (13.4–13.6) reported for guanidine, *N*-methylguanidine, and *N,N*-dimethylguanidine (40). The fact that N^ω -methylated substrates are

KEYWORDS

Citrullination: A post-translational protein modification that converts arginine side chains to citrulline. Also called deimination.

Covalent intermediate: A reaction intermediate that contains a transient covalent bond formed with an enzyme's active-site residue or with an enzyme-bound cofactor.

Deiminase: An enzyme that can hydrolyze an imine bond ($R^1R^2C=NR^3R^4$).

Demethyliminase: A deiminase capable of hydrolyzing an *N*-methylated imine substrate ($R^1R^2C=NR^3CH_3$).

Epigenetic information: Heritable information that is not directly encoded by the genome.

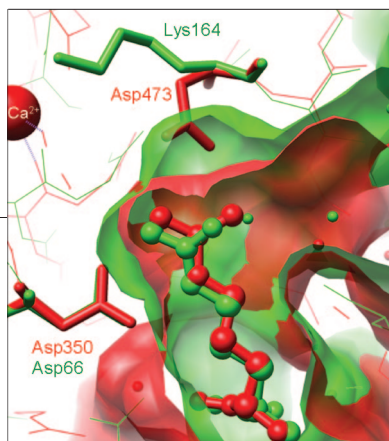


Figure 4. Structural overlay of an arginine analogue complexed with human PAD4 (red) and citrulline complexed with *Pseudomonas aeruginosa* DDAH (green). Ligands are shown as ball-and-stick models, and the active-site surfaces are color-coded to match the parent protein. The positioning of Asp residues interacting with the N^{δ} and N^{ϵ} nitrogens of the ligands are highly conserved, but the terminal Asp473 in PAD4, which makes a bidentate complex to each of the substrate's N^{ϵ} guanidino nitrogens, is not conserved in DDAH and serves to narrow the binding pocket and exclude N^{ϵ} -methylated ligands. The figure was created from PDB structures 1WDA and 1H70.

not effective PAD4 inhibitors (32) supports the idea that PAD4's preference for unmethylated substrates is due to steric exclusion rather than intrinsic stability of methylated guanidiniums.

A structural comparison with the related DDAH enzymes is instructive. DDAH is very selective for N^{ϵ} -methylated arginine substrates, and against unmodified arginine residues. This selectivity is essential for the proper physiological function of DDAH because it ensures that DDAH only hydrolyzes N^{ϵ} -methylated arginines, which are endogenous inhibitors of nitric oxide synthase, and does not hydrolyze arginine, which is a nitric oxide synthase substrate. A structural overlay of the PAD4 and DDAH active sites (Figure 4) shows very similar positioning of the substrates' side chains, and a side-on bidentate ligation of N^{δ} and N^{ϵ} by Asp350. Also conserved is a visible water channel that extends to the surface of the protein, presumably allowing the first product, ammonia, to diffuse away, and providing solvent access for the hydrolytic step. The most notable divergence between these two structures is the binding pocket for the two terminal N^{ϵ} nitrogens. PAD4 (and ADI) interacts with these nitrogens through a bidentate interaction with Asp473, whereas DDAH has a lysine at this position. In DDAH, the hydrophobic portion of this lysine residue helps to form a pocket for the N^{ϵ} -methyl groups. DDAH does possess a residue that is the functional equivalent of Asp473 (Glu65 in DDAH), but this residue comes from a different loop of the enzyme, forms only a monodentate interaction with the unsubstituted N^{ϵ} -nitrogen, and helps to direct the substituted N^{ϵ} -methyl group into its binding pocket. Although these particular substitutions are the most obvious differences between these two proteins, more extensive remodeling of the active-site environment (33, 37), including the positioning of Asn588, Val469, and Glu474, likely help to occlude N^{ϵ} -methyl group binding in PAD4 and underlie the preference of this enzyme toward unmethylated arginine residues.

Future Perspectives. Protein citrullination is emerging as an important PTM for both human disease and for gene regulation. However, significant questions about

PAD activity, specificity, and regulation remain unanswered. These questions represent excellent opportunities for the application of novel tools in proteomics and chemical biology and will likely be areas of intense future research.

Of primary interest is reconciling the *in vitro* and *in vivo* data for the "demethylation" activity of PAD4: are the *in vivo* results due to substrate depletion or epitope occlusion, or can PAD4 be transformed into an efficient demethylase *in vivo*? Several reasonable explanations for the latter possibility have been proposed, including the presence of accessory proteins or PTM of PAD4 that might increase its activity toward methylated substrates. The recent design and synthesis of covalent inactivators that either broadly target the AT superfamily (41) or specifically target the active form of the PADs (42) will be very useful in designing activity-based protein profiling reagents (43) that should be able to directly address these and other related issues.

The sequence specificities of PAD isozymes are also poorly defined, and recent crystal structures of PAD4 bound to short peptides (44) have not fully explained how the PADs, and in particular PAD4, regiospecifically modify particular arginine residues *in vitro* and *in vivo* (20–22). Although interactions between substrates and either the immunoglobulin-like domains or PAD binding proteins likely contribute to the substrate selectivity of these enzymes, it should be possible to determine a consensus sequence for this modification by analysis of a citrullinated proteome. Tools for detecting protein citrullination in the proteome include immunodetection of chemically modified citrulline residues (45) and MS analysis of either $^{16}\text{O}/^{18}\text{O}$ PAD-labeled products (46) or chemically modified citrulline residues. These residues can be derivatized with 2,3-butanedione alone or in combination with antipyrine to generate characteristic mass shifts in citrulline-containing peptides (47).

The regulation of PAD isoforms also raises important unanswered questions. PAD4 is a calcium-dependent enzyme that requires millimolar amounts of calcium to deiminate its protein substrates *in vitro* (15), yet the *in vivo* concentrations of calcium typically do not rise

above the low micromolar levels. Additionally, the presence of an active-site cysteine nucleophile and the sensitivity of the PADs to oxidation raise the possibility of redox regulation through reaction of this group with reactive oxygen or nitrogen species. An understanding of the physiologically relevant regulation mechanisms is essential for building a complete picture of how protein citrullination is controlled.

From a perspective of systems biology, the interplay of other enzymes and reactive metabolic compounds with residues targeted by PAD4 is also of significant interest. For example, if PAD4 cannot hydrolyze methylated substrates, then the possibility exists that other enzymes could serve as demethylases to truly reverse this modification. Interestingly, recent reports indicate that lysine-specific demethylase 1 (LSD1), a flavin-adenine dinucleotide-dependent amino oxidase, and members of the JmjC domain-containing histone demethylase (JHDM) family of enzymes, which are non-heme Fe(II), O₂, and α -ketoglutarate-dependent dioxygenases, can reverse histone lysine methylation (48–50). The presence of large numbers of LSD1 and JHDM homologues (~10 and 28, respectively) in an individual species (49) makes it reasonable to expect that a true arginine demethylase might be found among these coding sequences and would help to determine whether these enzymes act alone or possibly in combination with a modified form of PAD4 to control the levels of methylarginine. Also, the reaction of peptidylarginine and peptidyllysine with metabolically derived dicarbonyl reagents, such as methylglyoxal, results in PTMs collectively called advanced glycation end (AGE) products, which are well-studied hallmarks of diabetes and other

diseases (51–53). Dimethylation or citrullination would be expected to substantially change the susceptibility of peptidylarginine to these modifications.

Finally, the demonstration that citrulline levels in the N-terminal tails of histones H3 and H4 rise and then fall suggests that citrullination itself may be reversible. These results could be explained by epitope occlusion, histone tail clipping (54), or nucleosome displacement by chromatin remodeling enzymes; however, a more interesting proposal would be that the transitory nature of this PTM is due to the actions of an enzyme or enzymes that convert citrulline back into arginine. The existence of such a “decitrullinase”, as suggested by Bannister *et al.* (6), is a distinct possibility because precedents exist for the conversion of nonpeptidyl citrulline to arginine in the urea cycle by argininosuccinate synthetase and argininosuccinase.

In summary, a complete understanding of the interplay of citrullination with other PTMs of arginine (*e.g.*, methylation, AGE) will greatly inform our ideas about RA, gene regulation, and other aspects of human physiology. The characterization of these activities promises to be a fulfilling avenue of future research. Reconciling the results obtained from *in vitro* and *in vivo* studies of the PADs and PRMTs can best be achieved through the development and application of novel tools for studying the proteomics and chemical biology of arginine modification.

Acknowledgment: We thank the University of South Carolina Research Foundation (P.R.T.), the American Heart Association (0565409U; P.R.T.), the Arthritis National Research Foundation (W.F.), the American Cancer Society (RSG-05-061-01-GMC; W.F.), and the Robert A. Welch Foundation (F-1572; W.F.) for support.

REFERENCES

1. Rogers, G. E. (1962) Occurrence of citrulline in proteins, *Nature* **194**, 1149–1151.
2. Rogers, G. E., and Simmonds, D. H. (1958) Content of citrulline and other amino-acids in a protein of hair follicles, *Nature* **182**, 186–187.
3. Schellekens, G. A., de Jong, B. A., van den Hoogen, F. H., van de Putte, L. B., van Venrooij, W. J. and (1998) Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies, *J. Clin. Invest.* **101**, 273–281.
4. Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T., and Yamada, M. (2002) Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes, *Biochem. Biophys. Res. Commun.* **290**, 979–983.
5. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code, *Science* **293**, 1074–1080.
6. Bannister, A. J., Schneider, R., and Kouzarides, T. (2002) Histone methylation: dynamic or static? *Cell* **109**, 801–806.
7. Moscarello, M. A., Pritzker, L., Mastronardi, F. G., and Wood, D. D. (2002) Peptidylarginine deiminase: a candidate factor in demyelinating disease, *J. Neurochem.* **81**, 335–343.
8. Ishida-Yamamoto, A., Senshu, T., Takahashi, H., Akiyama, K., Nomura, K., and Iizuka, H. (2000) Decreased deiminated keratin K1 in psoriatic hyperproliferative epidermis, *J. Invest. Dermatol.* **114**, 701–705.
9. Bhattacharya, S. K., Crabb, J. S., Bonilha, V. L., Gu, X., Takahara, H., and Crabb, J. W. (2006) Proteomics implicates peptidyl arginine deiminase 2 and optic nerve citrullination in glaucoma pathogenesis, *Invest. Ophthalmol. Visual Sci.* **47**, 2508–2514.
10. Chang, X., and Han, J. (2006) Expression of peptidylarginine deiminase type 4 (PAD4) in various tumors, *Mol. Carcinog.* **45**, 183–196.

11. McGraw, W. T., Potempa, J., Farley, D., and Travis, J. (1999) Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase, *Infect. Immun.* **67**, 3248–3256.
12. Vossenaar, E. R., Zendman, A. J., van Venrooij, W. J., and Pruijn, G. J. (2003) PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease, *BioEssays* **25**, 1106–1118.
13. Vossenaar, E. R., and van Venrooij, W. J. (2004) Citrullinated proteins: sparks that may ignite the fire in rheumatoid arthritis, *Arthritis Res. Ther.* **6**, 107–111.
14. Shirai, H., Blundell, T. L., and Mizuguchi, K. (2001) A novel superfamily of enzymes that catalyze the modification of guanidino groups, *Trends Biochem. Sci.* **26**, 465–468.
15. Kearney, P. L., Bhatia, M., Jones, N. G., Luo, Y., Glascock, M. C., Catchings, K. L., Yamada, M., and Thompson, P. R. (2005) Kinetic characterization of protein arginine deiminase 4: a transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis, *Biochemistry* **44**, 10570–10582.
16. Arita, K., Hashimoto, H., Shimizu, T., Nakashima, K., Yamada, M., and Sato, M. (2004) Structural basis for Ca²⁺-induced activation of human PAD4, *Nat. Struct. Mol. Biol.* **11**, 777–783.
17. Nakashima, K., Hagiwara, T., and Yamada, M. (2002) Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J. Biol. Chem.* **277**, 49562–49568.
18. Taunton, J., Hassig, C. A., and Schreiber, S. L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p, *Science* **272**, 408–411.
19. Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines, *Nature* **383**, 269–272.
20. Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004) Histone deimination antagonizes arginine methylation. *Cell* **118**, 545–553.
21. Hagiwara, T., Hidaka, Y., and Yamada, M. (2005) Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes, *Biochemistry* **44**, 5827–5834.
22. Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D., and Coonrod, S. A. (2004) Human PAD4 regulates histone arginine methylation levels via demethyliminium, *Science* **306**, 279–283.
23. Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk, *Curr. Opin. Cell Biol.* **15**, 172–183.
24. Fischle, W., Wang, Y., and Allis, C. D. (2003) Binary switches and modification cassettes in histone biology and beyond, *Nature* **425**, 475–479.
25. Lee, Y. H., Coonrod, S. A., Kraus, W. L., Jelinek, M. A., and Stallcup, M. R. (2005) Regulation of coactivator complex assembly and function by protein arginine methylation and demethyliminium, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3611–3616.
26. Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K., and Kouzarides, T. (2002) Methylation at arginine 17 of histone H3 is linked to gene activation, *EMBO Rep.* **3**, 39–44.
27. Ma, H., Baumann, C. T., Li, H., Strahl, B. D., Rice, R., Jelinek, M. A., Aswad, D. W., Allis, C. D., Hager, G. L., and Stallcup, M. R. (2001) Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter, *Curr. Biol.* **11**, 1981–1985.
28. Schurter, B. T., Koh, S. S., Chen, D., Bunick, G. J., Harp, J. M., Hanson, B. L., Henschen-Edman, A., Mackay, D. R., Stallcup, M. R., and Aswad, D. W. (2001) Methylation of histone H3 by coactivator-associated arginine methyltransferase 1, *Biochemistry* **40**, 5747–5756.
29. Wang, H., Huang, Z.-Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P., and Zhang, Y. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor, *Science* **293**, 853–857.
30. Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001) Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1, *Curr. Biol.* **11**, 996–1000.
31. Chen, D., Ma, H., Hong, H., Koh, S. S., Huang, S. M., Schurter, B. T., Aswad, D. W., and Stallcup, M. R. (1999) Regulation of transcription by a protein methyltransferase, *Science* **284**, 2174–2177.
32. Hidaka, Y., Hagiwara, T., and Yamada, M. (2005) Methylation of the guanidino group of arginine residues prevents citrullination by peptidylarginine deiminase IV, *FEBS Lett.* **579**, 4088–4092.
33. Galkin, A., Lu, X., Dunaway-Mariano, D., and Herzberg, O. (2005) Crystal structures representing the Michaelis complex and the thioonium reaction intermediate of *Pseudomonas aeruginosa* arginine deiminase, *J. Biol. Chem.* **280**, 34080–34087.
34. Das, K., Butler, G. H., Kwiatkowski, V., Clark, A. D., Jr., Yadav, P., and Arnold, E. (2004) Crystal structures of arginine deiminase with covalent reaction intermediates; implications for catalytic mechanism, *Structure* **12**, 657–667.
35. Stone, E. M., Person, M. D., Costello, N. J., and Fast, W. (2005) Characterization of a transient covalent adduct formed during dimethylarginine dimethylaminohydrolase catalysis, *Biochemistry* **44**, 7069–7078.
36. Stone, E. M., Costello, A. L., Tiemey, D. L., and Fast, W. (2006) Substrate-assisted cysteine deprotonation in the mechanism of dimethylargininase (DDAH) from *Pseudomonas aeruginosa*, *Biochemistry* **45**, 5618–5630.
37. Lu, X., Li, L., Wu, R., Feng, X., Li, Z., Yang, H., Wang, C., Guo, H., Galkin, A., Herzberg, O., Mariano, P. S., Martin, B. M., and Dunaway-Mariano, D. (2006) Kinetic analysis of *Pseudomonas aeruginosa* arginine deiminase mutants and alternate substrates provides insight into structural determinants of function, *Biochemistry* **45**, 1162–1172.
38. Lu, X., Galkin, A., Herzberg, O., and Dunaway-Mariano, D. (2004) Arginine deiminase uses an active-site cysteine in nucleophilic catalysis of L-arginine hydrolysis, *J. Am. Chem. Soc.* **126**, 5374–5375.
39. Murray-Rust, J., Leiper, J., McAlister, M., Phelan, J., Tilley, S., Santa Maria, J., Vallance, P., and McDonald, N. (2001) Structural insights into the hydrolysis of cellular nitric oxide synthase inhibitors by dimethylarginine dimethylaminohydrolase, *Nat. Struct. Mol. Biol.* **8**, 679–683.
40. Jencks, W. P., and Regenstein, J. (1968) Ionization constants of acids and bases, in *Handbook of Biochemistry* (Sober, H. A., Ed.) pp 1-148–1-189, CRC Press, Cleveland, OH.
41. Stone, E. M., Schaller, T. H., Bianchi, H., Person, M. D., and Fast, W. (2005) Inactivation of two diverse enzymes in the amidinotransferase superfamily by 2-chloroacetamide: dimethylargininase and peptidylarginine deiminase, *Biochemistry* **44**, pp 13744–13752.
42. Luo, Y., Knuckley, B., Lee, Y. H., Stallcup, M. R., and Thompson, P. R. (2006) A fluoro-acetamide based inactivator of protein arginine deiminase 4 (PAD4): design, synthesis, and in vitro and in vivo evaluation, *J. Am. Chem. Soc.* **128**, 1092–1093.
43. Jessani, N., and Cravatt, B. F. (2004) The development and application of methods for activity-based protein profiling, *Curr. Opin. Chem. Biol.* **8**, 54–59.
44. Arita, K., Shimizu, T., Hashimoto, H., Hidaka, Y., Yamada, M., and Sato, M. (2006) Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5291–5296.
45. Senshu, T., Sato, T., Inoue, T., Akiyama, K., and Asaga, H. (1992) Detection of citrulline residues in deiminated proteins on polyvinylidene difluoride membrane, *Anal. Biochem.* **203**, 94–100.

46. Kubota, K., Yoneyama-Takazawa, T., and Ichikawa, K. (2005) Determination of sites citrullinated by peptidylarginine deiminase using ^{18}O stable isotope labeling and mass spectrometry, *Rapid Commun. Mass Spectrom.* **19**, 683–688.
47. Holm, A., Rise, F., Sessler, N., Sollid, L. M., Undheim, K., and Fleckenstein, B. (2006) Specific modification of peptide-bound citrulline residues, *Anal. Biochem.* **352**, 68–76.
48. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, J. R., Cole, P. A., and Casero, R. A. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1, *Cell* **119**, 941–953.
49. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins, *Nature* **439**, 811–816.
50. Whetstone, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases, *Cell* **125**, 467–481.
51. Thomas, M. C., Baynes, J. W., Thorpe, S. R., and Cooper, M. E. (2005) The role of AGEs and AGE inhibitors in diabetic cardiovascular disease, *Curr. Drug Targets* **6**, 453–474.
52. Thorpe, S. R., and Baynes, J. W. (2003) Maillard reaction products in tissue proteins: new products and new perspectives, *Amino Acids* **25**, 275–281.
53. Fackelmayer, F. O. (2005) Protein arginine methyltransferases: guardians of the Arg? *Trends Biochem. Sci.* **30**, 666–671.
54. Allis, C. D., Bowen, J. K., Abraham, G. N., Glover, C. V., and Gorovskiy, M. A. (1980) Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in *Tetrahymena micronuclei*, *Cell* **20**, 55–64.

An *ent*-Kaurene That Inhibits Mitotic Chromosome Movement and Binds the Kinetochores Protein Ran-Binding Protein 2

Natalie T. Rundle[†], Jim Nelson[‡], Mark R. Flory^{§,**}, Jomon Joseph^{¶,††}, John Th'ng^{||}, Ruedi Aebersold[§], Mary Dasso[¶], Raymond J. Andersen[‡], and Michel Roberge^{†,*}

[†]Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada, [‡]Departments of Earth-Ocean Sciences and Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada, [§]Institute for Systems Biology, Seattle, Washington 98103-8904, [¶]Laboratory of Gene Regulation and Development, National Institute of Child Health and Development, Bethesda, Maryland 20892-5431, ^{||}Cancer Care Ontario, Northwestern Ontario Regional Cancer Centre, Thunder Bay, Ontario P7B 6V4, Canada, ^{**}Present address: Department of Molecular Biology and Biochemistry, Wesleyan University, Hall-Atwater Laboratories, Middletown, Connecticut 06459. ^{††}Present address: National Centre for Cell Science, Ganeshkhind, India.

Mitosis is a dynamic and highly regulated process whereby a cell precisely divides its chromosomes into two daughter cells. Chemicals that perturb mitotic entry and progression can be valuable tools to uncover the mechanisms underlying this complex biological process (1). A large number of structurally distinct chemicals, generally called antimitotic agents, are known to block mitosis. Most have been found to disrupt the function of the mitotic spindle by binding directly to tubulins, the proteins that form microtubules (2), but in recent years, a number have been discovered that inhibit mitotic progression by acting on other targets. For example, monastrol and SB715992 inhibit kinesin motor proteins and cause cells to arrest with monopolar spindles (3, 4), the Plk1 inhibitor ONO1910 causes mitotic arrest with multipolar spindles (5, 6), hesperadin and ZM447439 inhibit Aurora kinases and cause defects in chromosome condensation and alignment (7–9), diminutol can inhibit NADP-binding proteins and affect microtubule morphogenesis (10), and protein serine/threonine phosphatase inhibitors can also block cells in mitosis (11).

Many distinct biochemical activities required for mitosis do not yet have chemical inhibitors (12). Novel inhibitors of those activities would be useful for unraveling the mechanisms of mitosis and may also find application in cancer therapy. Paclitaxel and the vinca alkaloids target microtubules and are used to treat various

ABSTRACT Using a chemical genetics screen, we have identified *ent*-15-oxokaurenoic acid (EKA) as a chemical that causes prolonged mitotic arrest at a stage resembling prometaphase. EKA inhibits the association of the mitotic motor protein centromeric protein E with kinetochores and inhibits chromosome movement. Unlike most antimitotic agents, EKA does not inhibit the polymerization or depolymerization of tubulin. To identify EKA-interacting proteins, we used a cell-permeable biotinylated form that retains biological activity to isolate binding proteins from living cells. Mass spectrometric analysis identified six EKA-binding proteins, including Ran-binding protein 2, a kinetochores protein whose depletion by small interfering RNA causes a similar mitotic arrest phenotype.

*Corresponding author,
michelr@interchange.ubc.ca.

Received for review May 9, 2006
and accepted July 25, 2006.

Published online August 11, 2006
10.1021/cb600196w CCC: \$33.50

© 2006 by American Chemical Society

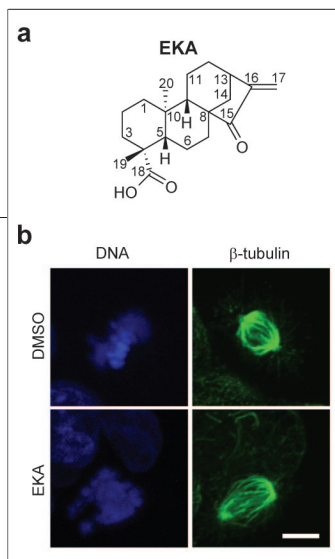


Figure 1. Structure and mitotic arrest phenotype of EKA. a) Structural formula of EKA. b) Mitotic spindles and chromosomes were examined by indirect immunofluorescence microscopy after treatment of HCT116 p53^{-/-} cells with DMSO (top row) or EKA (20 μ M, bottom row) for 8 h. The DNA (left) and β -tubulin (right) signals were visualized by confocal microscopy. Results are typical of several ($n > 6$) independent experiments. Bar, 5 μ m.

tumors (13, 14). However, tubulin-targeting agents also affect the function of microtubules in neuronal and other non-dividing normal cells, leading to toxicity that limits the clinical utility of these agents. Chemicals that selectively target mitotic proteins may be more effective therapeutic agents.

As part of a program to discover chemicals that modulate mitotic entry and progression (15, 16), we have identified a plant metabolite that blocks chromosome movement prior to metaphase, interferes with the proper localization of a motor protein required for chromosome movement, and binds to Ran-binding protein 2 (RanBP2), a protein required for chromosome segregation.

RESULTS AND DISCUSSION

An *ent*-Kaurene That Causes an Unusual Mitotic Arrest Phenotype. We have isolated the diterpenoid *ent*-15-oxokaurenoic acid (EKA) from a plant extract (Figure 1, panel a) as a compound eliciting an unusual mitotic arrest phenotype. Cells treated with EKA for 8 h and immunostained with a β -tubulin antibody and the DNA dye TOTO-3 showed 20–30% mitotic arrest with half-maximal activity at $8.3 \pm 1.5 \mu$ M. All arrested cells had a similar unusual phenotype resembling prometaphase in which condensed chromosomes were scattered across mitotic spindles that were clearly bipolar but showed an abnormal morphology (Figure 1, panel b). Most chromosomes were distributed across the mitotic spindle, apparently bound to microtubules, but some showed no evidence of attachment to spindle microtubules (Figure 2). This mitotic arrest phenotype appeared irreversible because cells did not exit mitosis upon drug removal. Similar effects were observed in four different cell lines.

During prometaphase, chromosomes are captured by microtubules and move in an oscillatory fashion before becoming aligned at the center of the mitotic spindle (17). To determine whether EKA blocks congression by inhibiting these oscillatory movements, we monitored chromosomes in live MCF-7 cells expressing a GFP-histone H1 fusion protein to fluorescently label the chromosomes. Movies of cells undergoing mitosis in the presence of dimethyl sulfoxide (DMSO) or EKA were

recorded with time-lapse 3D confocal microscopy. In the presence of DMSO, cells underwent prometaphase in ~ 10 min and completed mitosis in little over 1 h (Supplementary Video 1). In the presence of EKA (20 μ M), very little chromosome movement was detected and cells remained in a prometaphase-like state over the entire observation period, which varied from 30 to 60 min (Figure 3, panel a; Supplementary Video 2), an indication that EKA blocks chromosome movement.

Most agents that block mitotic progression target microtubules directly and interfere with their polymerization or depolymerization. To examine stimulation of microtubule polymerization, purified brain tubulin ($>95\%$ pure, 1 mg mL⁻¹) was preincubated with DMSO, EKA, or paclitaxel on ice, and the temperature was raised to 37 $^{\circ}$ C. At this low tubulin concentration, microtubules do not polymerize spontaneously, but microtubule-stabilizing agents such as paclitaxel induce polymerization. EKA did not significantly stabilize microtubules (Figure 3, panel b). To determine whether EKA can depolymerize microtubules, tubulin (3 mg mL⁻¹) was preincubated with DMSO, EKA, or nocodazole on ice, and the temperature was raised to 37 $^{\circ}$ C. At this higher tubulin concentration, microtubules polymerize spontaneously, but polymerization is prevented by microtubule-destabilizing agents such as nocodazole (Figure 3, panel c). EKA did not prevent microtubule polymerization (Figure 3, panel c). These results indicated that EKA interferes with chromosome movement by acting on targets other than tubulin.

The spindles of EKA-treated cells were clearly bipolar (Figure 1 and Figure 2), unlike those resulting from inhibition of kinesins by monastrol and SB715992 or of Plk1 by ON01910. Further, EKA caused sustained mitotic arrest and strong spindle assembly checkpoint activation (not shown) as measured by 3F3/2 immunostaining (18, 19), unlike the Aurora kinase inhibitors hesperadin and ZM447439, which cause defects in chromosome condensation and alignment but do not activate the spindle assembly checkpoint (7, 8). Further, EKA did not inhibit the protein serine/threonine phosphatase PP1 or PP2A *in vitro* (not shown). These data suggest that the target of EKA is not one of the above proteins.

EKA prevents the association of centromeric protein E (CENP-E) with kinetochores. Chromosome movement is controlled by mitotic motor proteins (12, 20, 21). CENP-E is a kinesin-like motor protein that plays an

essential role in chromosome congression (22, 23). Inhibition of CENP-E by immunodepletion, antibody microinjection, or overexpression of dominant interfering mutants results in cell cycle arrest in prometaphase with condensed chromosomes and bipolar spindles (24, 25). Because this phenotype bears resemblance to that resulting from EKA treatment, we next determined the effect of EKA on the intracellular localization of CENP-E by immunofluorescence microscopy. Cells were treated with DMSO or EKA (20 μM) for 8 h before immunolabeling CENP-E and β -tubulin and counterstaining chromosomes. In control cells undergoing prometaphase, CENP-E was found on kinetochores (Figure 2, top row), in agreement with published results (22, 23, 26, 27). Strikingly, after EKA treatment, kinetochore-associated CENP-E was dramatically reduced in $\sim 80\%$ of prometaphase-like cells, and its signal became very bright at the spindle poles (Figure 2, middle and bottom rows). This effect was observed in all four cell lines examined. That CENP-E was truly localized to spindle poles in EKA-treated cells was confirmed by co-immunolabeling CENP-E and the centrosomal protein γ -tubulin (not shown). Treatment with nocodazole (300 ng mL^{-1} , 8 h) did not cause CENP-E to localize to spindle poles, indicating that pole localization of CENP-E is not a common consequence of mitotic arrest by chemicals.

hBubR1 is a kinetochore protein that plays an early activating role in the checkpoint that monitors the

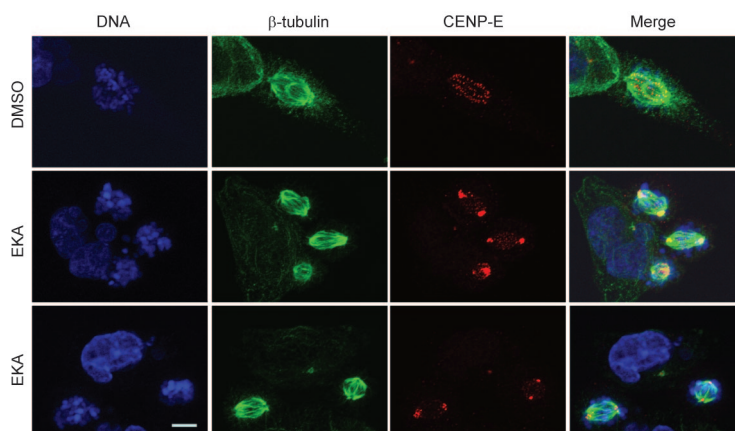


Figure 2. Effect of EKA on the intracellular localization of CENP-E. After incubation with DMSO or EKA (20 μM) for 8 h, HCT116 p53 $^{-/-}$ cells were immunostained with β -tubulin and CENP-E antibodies, and DNA was labeled with Hoechst 33258. Images were collected by confocal microscopy. Results are typical of six independent experiments. Bar, 5 μm .

mitotic spindle (28–30). We tested whether EKA also interferes with the kinetochore localization of this protein. DMSO-treated prometaphase cells exhibited bright punctate hBubR1 staining indicative of kinetochore localization (Figure 4, top row). After EKA treatment, a punctate pattern was also observed, indicating that hBubR1 was associated with kinetochores, but many hBubR1 dots were closer to the centrosomes than in control cells (Figure 4, bottom row). This distribution was observed in $>70\%$ of prometaphase-arrested cells

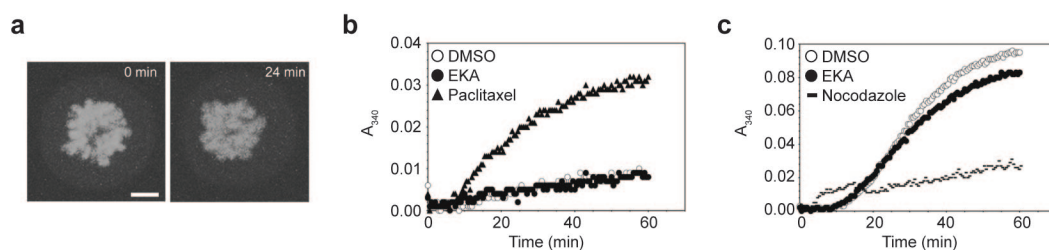


Figure 3. EKA inhibits mitotic chromosome movement but does not target microtubules. **a)** Chromosome movement is greatly reduced during incubation with EKA. 3D time-lapse recordings were taken of live cells that entered mitosis during EKA treatment. MCF-7 cells expressing a GFP–histone H1 fusion protein were incubated with EKA (5 μM) for 4 h before confocal microscopy. Projections from the 0 and 24 min time points are shown. Full-length movies of DMSO- (Supplementary Video 1) and EKA-treated cells (Supplementary Video 2) are available. Bar, 5 μm . **b)** and **c)** EKA neither promotes nor inhibits microtubule polymerization. **b)** Purified bovine brain tubulin (1 mg mL^{-1}) was incubated with DMSO (0.3%, \circ), EKA (20 μM , \bullet), or paclitaxel (10 μM , \blacktriangle) to test for microtubule-polymerizing activity. At 0 min, additions were made to prechilled tubulin aliquots, and samples were rapidly brought to 37 $^{\circ}\text{C}$. An increase in turbidity (A_{340}) indicates microtubule assembly. **c)** The tubulin concentration was increased to 3 mg mL^{-1} to promote spontaneous microtubule assembly. Incubations were carried out as in panel b, with nocodazole (30 μM , $-$) instead of paclitaxel.

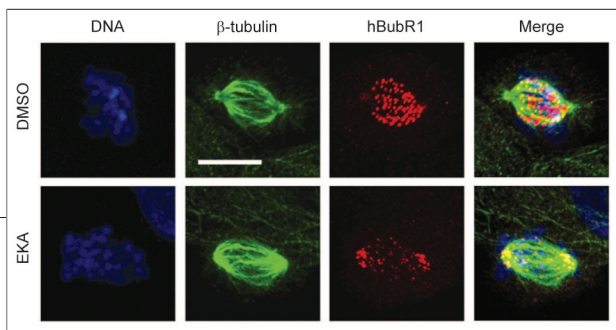


Figure 4. Effect of EKA on the intracellular localization of hBubR1. Cells treated as in Figure 2 were immunostained with β -tubulin and hBubR1 antibodies, and DNA was labeled with Hoechst 33258. Images were collected by confocal microscopy. Results are typical of six independent experiments. Bar, 5 μ m.

and in all four cell lines examined. This distribution is distinct from that of CENP-E and may be a reflection of a large number of chromosomes near centrosomes in EKA-treated cells rather than an effect of EKA on the association of hBubR1 with kinetochores.

Interference with the function of CENP-E is therefore likely to be a significant factor contributing to the prometaphase arrest elicited by EKA. We observed that in prophase, CENP-E localized to the poles of the assembling spindle before localizing to kinetochores at prometaphase. The localization of CENP-E at spindle poles in EKA-treated cells raises the possibility that EKA either interferes with the translocation of CENP-E to kinetochores or prevents its binding to kinetochores.

A Cell-Permeable Biotinylated EKA Derivative. EKA possesses an unsaturated α,β -carbonyl group that makes it reactive towards cysteine residues and able to form adducts with target proteins. In order to characterize further the mechanism of action of EKA, we wished to create a derivative suitable for isolating EKA-binding proteins and for studying the intracellular localization of EKA binding sites.

EKA was derivatized by reaction in basic refluxing acetone with *N*-iodoacetyl-*N*-biotinylhexylenediamine, a cell membrane-permeable biotinylating reagent with an extended linker group. Attachment occurred through an ester linkage at C18.

The structure of the reaction product, called biotinylated EKA (B-EKA; Figure 5, panel a), was confirmed by NMR and mass spectrometry. 1-Methyl-1-cyclohexane carboxylic acid, identical in structure to the C1–C5 ring and attached groups in EKA, was also reacted with *N*-iodoacetyl-*N*-biotinylhexylenediamine (Figure 5, panel b). This biotinylated control compound, named B-Ctrl, was included as a negative control to help identify non-specific interactions between proteins and the beads, linker group, or C1–C5 ring of EKA.

We used immunofluorescence microscopy to determine whether B-EKA retained the antimetabolic activity of the parent compound. Labeling of β -tubulin and DNA revealed that cells treated with 10 μ M B-EKA arrested in prometaphase with condensed chromosomes and

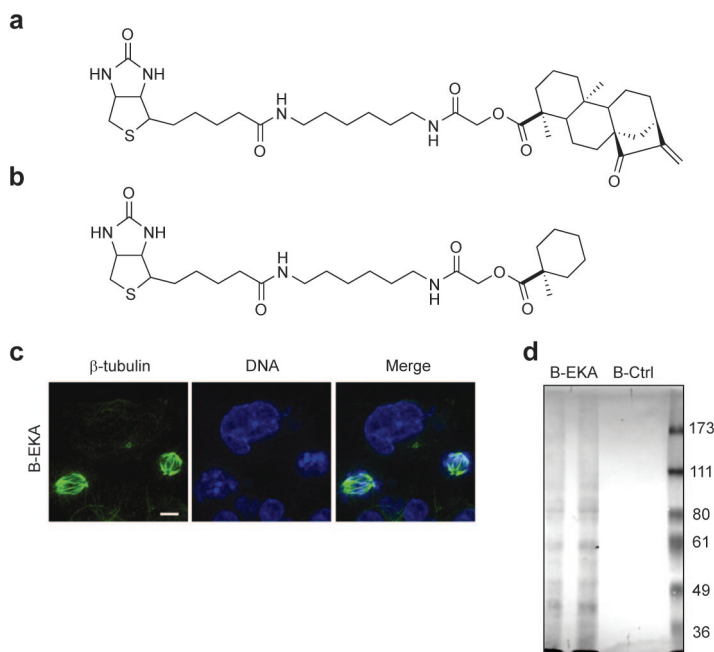


Figure 5. Biotinylated EKA (B-EKA) retains biological activity and binds covalently to cellular proteins. a) Structural formula of B-EKA. b) Structural formula of the biotinylated control compound B-Ctrl. c) HCT116 p53^{-/-} cells were treated with 5 μ M B-EKA for 8 h and immunostained with β -tubulin antibody, and DNA was labeled with TOTO-3. Images were collected by confocal microscopy. Bar, 5 μ m. d) HCT116 p53^{-/-} cells were treated with 5 μ M B-EKA or B-Ctrl for 8 h, lysates were prepared, and interacting proteins were isolated on streptavidin–agarose beads. The biotinylated proteins were then separated by SDS–polyacrylamide gel electrophoresis (PAGE), electrotransferred to polyvinylidene difluoride (PVDF) membrane, bound to horseradish peroxidase (HRP)-conjugated streptavidin, and detected with a chromogenic substrate. The molecular mass of protein markers in kilodaltons is indicated to the right.



Figure 6. Map of RanBP2 peptides identified by mass spectrometry. We incubated HCT116 p53^{-/-} cells with B-EKA (5 μ M for 8 h), biotinylated proteins were purified by streptavidin affinity precipitation and identified by mass spectrometry. The position and sequence of recovered peptides is shown in relation to the domain structure of RanBP2 (adapted from ref. 47).

bipolar mitotic spindles (Figure 5, panel c) and that CENP-E localized to centrosomes (not shown). We concluded that B-EKA was indeed able to penetrate cells and that its activity and potency were similar to EKA.

Identification of B-EKA-Binding Proteins. B-EKA was used as an affinity ligand to purify covalently interacting proteins. We first used a batch purification approach in which cells were incubated with B-Ctrl or B-EKA (8 h, both at 5 μ M) and lysed and biotinylated proteins were purified by affinity precipitation with streptavidin–agarose beads under stringent conditions. The proteins were eluted from the beads and digested with trypsin. Tryptic peptides were separated by reversed-phase HPLC, mass spectra and sequence data were collected by tandem mass spectrometry, and proteins were identified by comparison of the spectra and sequence data to the SEQUEST databank. In a second approach, proteins eluted from streptavidin–agarose beads were separated by SDS-PAGE and transferred to a membrane. Biotinylated proteins were then detected by incubation with HRP-conjugated streptavidin and a chromogenic substrate, and individual bands were excised, trypsinized, and identified by reversed-phase tandem mass spectrometry as above. Several biotinylated protein bands were detected on membranes when B-EKA was used as an affinity ligand but none with the B-Ctrl control compound (Figure 5, panel d), indicating that B-EKA bound covalently to cellular proteins. Six B-EKA-specific proteins were identified by both batch purification and band excision in five independent experiments: FRAP1 (FK506 binding protein 12-rapamycin associated protein, NP_004949), phosphate carrier precursor isoform 1a (NP_005879), Parc (p53-associated Parkin-like cytoplasmic protein, XP_166408.2), RanBP2 (NP_006258), and two unnamed proteins (XP_169369 and XP_172512).

RanBP2 was reproducibly identified by several peptides of between 7 and 24 amino acid residues in length distributed in the amino terminal half of the protein (Figure 6). We chose to study this protein further because of its documented involvement in mitotic progression (31). We examined the intracellular localization of

B-EKA and compared it to that of RanBP2. HeLa cells were treated with B-EKA, permeabilized with digitonin, fixed with formaldehyde, and reacted with fluorescent streptavidin. The distribution of RanBP2 was determined in the same cells using RanBP2 antibodies.

In interphase cells, RanBP2 localized mainly to the nuclear envelope (Figure 7, top row), where it is a component of nuclear pore complexes (32–35). B-EKA staining was strong at the nuclear envelope but also present in a discrete perinuclear region overlapping with the Golgi (Figure 7, top row). By contrast, B-Ctrl staining was very weak and diffuse throughout the cells (not shown). Merging the B-EKA and RanBP2 fluorescence signals showed co-staining at the nuclear envelope but not in the cytosol (Figure 7). The nuclear envelope localization of RanBP2 was not affected by either B-EKA or EKA treatment.

In mitotic cells, RanBP2 has previously been shown to localize to the base of spindles and at kinetochores

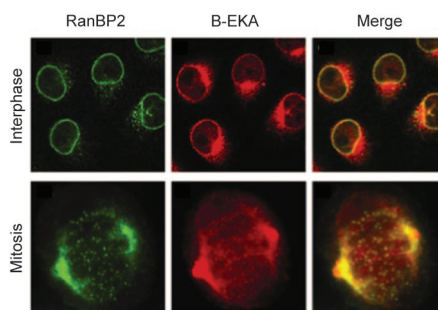


Figure 7. Colocalization of B-EKA with RanBP2 in interphase and mitotic cells. HeLa cells were treated with B-EKA (10 μ M, 8 h). RanBP2 was visualized using a RanBP2 antibody (green) and B-EKA using fluorescently labeled streptavidin (red). In the right column, the fluorescence signals were merged. Yellow color indicates colocalization.

(32). Treatment with EKA or B-EKA did not change the subcellular localization of RanBP2 in mitotic cells (Figure 7, bottom row). B-EKA also localized to the base of mitotic spindles and to numerous small dots between spindle poles (Figure 7, bottom row). Merging of the B-EKA and RanBP2 fluorescence signals showed colocalization at the base of spindles and at kinetochores, but a small fraction of B-EKA was not colocalized with RanBP2 (Figure 7, bottom row). Therefore, most if not all RanBP2 colocalizes with B-EKA in interphase and mitotic cells, corroborating the affinity chromatography result.

This study found no evidence of interaction between EKA and tubulin, protein serine/threonine phosphatases, kinesin motor proteins, Plk1, or Aurora kinases, and the mitotic arrest phenotype of EKA is distinct from that of chemical inhibitors of these proteins, suggesting a novel mechanism of action for EKA. B-EKA binds covalently to RanBP2 and significantly

colocalizes with this protein in interphase and mitotic cells at the light microscopy level. Salina *et al.* (36) have shown that depletion of RanBP2 by small interfering RNA (siRNA) caused cells to arrest at prometaphase. The phenotype of cells treated with RanBP2 siRNA described by these authors bears strong resemblance to that following EKA treatment. In both cases, the spindles display a somewhat perturbed morphology, with well-separated centrosomes and condensed chromosomes that fail to congress. After RanBP2 siRNA treatment, the association of CENP-E with kinetochores is also strongly reduced while its association with spindle poles is increased. The notably similar inhibition of mitotic progression by EKA treatment and by RanBP2 depletion using siRNA provides strong evidence that EKA causes mitotic arrest at least in part by inhibiting RanBP2 function. EKA should be a useful addition to an expanding repertoire of cell cycle-specific chemical biology tools.

METHODS

Cell Culture. Experiments were carried out using HeLa cells and MCF-7 or HCT116 cells and their derivatives lacking p53 function (MCF-7 mp53 and HCT116 p53^{-/-}) (37, 38). Entry into mitosis was evaluated using an enzyme-linked cytochemical assay with the mitosis-specific monoclonal antibody TG-3 (39), a generous gift of Dr. P. Davies, Albert Einstein College of Medicine, Bronx, NY.

Isolation and Identification of EKA. A methanol/dichloromethane extract of the South American fruiting shrub *Coccoloba acuminata* obtained from the National Cancer Institute Natural Products Repository was fractionated by size exclusion chromatography on a Sephadex LH 20 column using 100% methanol as the eluent, and the active fraction was further fractionated by chromatography on LH 20 using EtOAc/MeOH/H₂O (20:5:1) as the eluent. Final fractionation of the active LH 20 fraction using gradient silica gel flash chromatography (eluents, ethyl acetate/hexane, 5:95 → ethyl acetate/hexane, 15:85) yielded a pure active compound. The structure of the active compound was determined by analysis of ¹H and ¹³C NMR and high-resolution mass spectrometry data, and it was found to be identical to the known compound EKA (40).

3D Confocal Microscopy of Live Cells. MCF-7 cells that stably express a histone H1-GFP fusion protein (41) were cultured on glass-bottomed culture wells (Tekware) and treated with DMSO or EKA (20 μM) for 4 h before microscopy. Temperature was maintained at 37 °C during microscopy with a thermal insulating enclosure. During each 2 min recording cycle, a stack of 20 sections (total thickness 4 μm) was collected at the approximate center of the cell. Stacks for individual time points were projected and animated using ImageJ software (National Institutes of Health).

Purification and Detection of B-EKA Binding Proteins. Proteins that bind to B-EKA were purified by streptavidin affinity precipitation according to the method of Zimmer *et al.* (42) with modification. HCT116 p53^{-/-} cells (5 × 10⁶) were seeded onto six 10 cm² dishes and treated with B-Ctrl or B-EKA (both at

5 μM) for 8 h. Cells were collected by trypsinization, washed once in phosphate buffered saline (PBS) containing 0.2 mM phenylmethylsulfonyl fluoride (PMSF), suspended in 20 mM Tris-HCl, pH 8.0, and lysed by addition of an equal volume of 2× radio immune precipitation assay (RIPA) lysis buffer (20 mM Tris-HCl, pH 8.0, 0.2% SDS, 2% NP-40, 2% sodium deoxycholate, 300 mM NaCl, 0.04% sodium azide, 0.2 mM PMSF, and 10 μg mL⁻¹ each of leupeptin, aprotinin, antipain, and pepstatin). Lysates were passed several times through a syringe fitted with a 21-gauge needle and incubated on ice for 1 h. Supernatants were collected after centrifugation at 177,000g for 1 h at 4 °C. The protein concentration of each sample was measured by bicinchoninic acid protein assay and equalized between samples by adding 1× RIPA buffer. Lysates (~20 mg of protein) were loaded onto 300 μL of a 50% slurry of streptavidin-agarose beads (Molecular Probes) that had been prewashed three times in RIPA buffer and agitated overnight at 4 °C. The beads were washed three times in RIPA buffer, followed by one wash in 20 mM Tris-HCl, pH 6.8. Proteins were eluted by boiling in 50 mM Tris-HCl, pH 6.8, 2% SDS for 10 min with agitation. In some experiments, proteins purified by streptavidin affinity precipitation were separated by SDS-PAGE in 8% acrylamide gels. Gels were electrotransferred to polyvinylidene difluoride membrane (Millipore) and blocked overnight at 4 °C in PBS containing 5% bovine serum albumin (BSA) and 0.02% sodium azide. Biotinylated proteins were labeled by incubation with HRP-conjugated streptavidin (Molecular Probes), diluted 1:50,000 in PBS containing 0.1% Triton X-100 and 1% BSA for 1 h at RT. After washing the blots three times in PBS containing 0.1% Triton X-100 and 1% BSA, labeled bands were detected by incubation with 3,3',5,5'-tetramethylbenzidine (Pierce) chromogenic substrate for several minutes.

Analysis of B-EKA-Binding Proteins by Mass Spectrometry. Proteins isolated by B-EKA affinity precipitation and bands cut from PVDF membranes were incubated in 20 mM Tris-HCl, pH 8.3, 5 mM EDTA and digested with 20 μg mL⁻¹ trypsin overnight at 37 °C. The resulting peptides were purified *via* electro-

static charge by strong cation exchange liquid chromatography using ICAT cation exchange syringe cartridges (Applied Biosystems) according to the manufacturer's recommendations. Peptides were desalted and further purified using C18 ZipTip pipette tips (Millipore) according to the manufacturer's recommendations. Peptides were subjected to 1D reversed-phase chromatography with on-line mass spectrometry on an ion trap mass spectrometer (model LCQ, ThermoFinnigan, San Jose, CA). This was performed as described (43) employing a 2-h binary gradient from 5–80% solvent B during which each mass spectrum (MS) scan was followed by three MS/MS scans. Experimental mass spectra were compared with theoretical spectra generated from sequences from the human genome using the SEQUEST algorithm (44). Data were displayed and filtered using INTERACT software (45).

Immunofluorescence Labeling of CENP-E and hBubR1. Labeling of CENP-E and hBubR1 was carried out according to the method of Jablonski *et al.* (46). HCT116 p53^{-/-} cells were seeded onto coverslips (150,000 cells) and were treated 1 d later with DMSO or EKA (20 μ M) for 8 h. Cells were extracted for 1 min in 4 M glycerol, 0.1 M piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES), pH 6.9, 1 mM ethylene glycol bis(2-aminoethyl ether)tetraacetic acid (EGTA), containing 0.5% Triton X-100, washed 2 min in the above buffer without Triton X-100, and fixed by incubation in PBS, pH 6.8, containing 3.7% formaldehyde for 7 min. After washing twice in KB (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.1% BSA), coverslips were incubated with rabbit CENP-E or hBubR1 antibodies (gifts from Dr. G. Chan, University of Alberta), diluted 1:1,000 in KB, for 1 h. After being washed twice in KB, coverslips were incubated in Cy3-conjugated goat anti-rabbit secondary antibody (Sigma), diluted 1:1,000 in KB, for 30 min. After two washes in KB, samples were incubated with 500 U mL⁻¹ RNase A (Roche Diagnostics) in TBS (10 mM Tris-HCl, pH 7.4, 150 mM NaCl) for 30 min at 37 °C, followed by addition of 2 mM TOTO-3 (Molecular Probes) for 20 min in darkness. After being washed twice in TBS, the coverslips were mounted in 90% glycerol in PBS containing 0.2 M *N*-propyl gallate. Images were collected by confocal microscopy.

Intracellular Localization of B-EKA and RanBP2. HeLa cells were seeded onto coverslips and were treated with DMSO or 10 μ M B-EKA for 8 h. The cells were then permeabilized with digitonin and fixed with formaldehyde as described (32). RanBP2 was detected using rabbit anti-human RanBP2 antibodies and Alexa-488-conjugated goat anti-rabbit antibody (Molecular Probes), and B-EKA was detected using Alexa-594-conjugated streptavidin (Molecular Probes). Cells were mounted in Vectashield medium (Vector Laboratories) and visualized by fluorescence microscopy.

Acknowledgments: We thank B. Vogelstein for HCT116 cell lines, P. Davies and G. Chan for antibodies, G. Cragg and D. Newman for access to the NCI Natural Products Repository, and H. Anderson for helpful discussions. This work was supported by a studentship from the National Cancer Institute of Canada (NTR) and grants from the National Cancer Institute of Canada (RIA) and the Canadian Institutes of Health Research (MR).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Lampson, M. A., and Kapoor, T. M. (2006) Unraveling cell division mechanisms with small-molecule inhibitors, *Nat. Chem. Biol.* 2, 19–27.
- Hadfield, J. A., Ducki, S., Hirst, N., and McGown, A. T. (2003) Tubulin and microtubules as targets for anticancer drugs, *Prog. Cell Cycle Res.* 5, 309–325.
- Mayer, T. U., Kapoor, T. M., Haggarty, S. J., King, R. W., Schreiber, S. L., and Mitchison, T. J. (1999) Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen, *Science* 286, 971–974.
- Bergnes, G., Brejc, K., and Belmont, L. (2005) Mitotic kinesins: prospects for antimetastatic drug discovery, *Curr. Top. Med. Chem.* 5, 127–145.
- Gumireddy, K., Reddy, M. V., Cosenza, S. C., Boominathan, R., Baker, S. J., Papatih, N., Jiang, J., Holland, J., and Reddy, E. P. (2005) ONO1910, a non-ATP-competitive small molecule inhibitor of Plk1, is a potent anticancer agent, *Cancer Cell* 7, 275–286.
- Strebhardt, K., and Ullrich, A. (2006) Targeting polo-like kinase 1 for cancer therapy, *Nat. Rev. Cancer* 6, 321–330.
- van Meel, J. Hauf, S., Cole, R. W., LaTerra, S., Zimmer, C., Schnapp, G., Walter, R., Heckel, A., Rieder, C. L., and Peters, J. M. (2003) The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint, *J. Cell Biol.* 161, 281–294.
- Ditchfield, C., Johnson, V. L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N., and Taylor, S. S. (2003) Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores, *J. Cell Biol.* 161, 267–280.
- Andrews, P. D. (2005) Aurora kinases: shining lights on the therapeutic horizon? *Oncogene* 24, 5005–5015.
- Wignall, S. M., Gray, N. S., Chang, Y. T., Juarez, L., Jacob, R., Burlingame, A., Schultz, P. G., and Heald, R. (2004) Identification of a novel protein regulating microtubule stability through a chemical approach, *Chem. Biol.* 11, 135–146.
- Vandre, D. D., and Wills, V. L. (1992) Inhibition of mitosis by okadaic acid: possible involvement of a protein phosphatase 2A in the transition from metaphase to anaphase, *J. Cell Sci.* 101, 79–91.
- Miyamoto, D. T., Perlman, Z. E., Mitchison, T. J., and Shirasu-Hiza, M., Dynamics of the mitotic spindle—potential therapeutic targets, *Prog. Cell Cycle Res.* 5, 349–360.
- Kingston, D. G. I. (2005) Taxol and its analogs, in *Anticancer agents from natural products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds.) pp 89–122, Taylor and Francis, Boca Raton, FL.
- Gueritte, F., and Fahy, J. (2005) The vinca alkaloids, in *Anticancer agents from natural products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds.) pp 123–135, Taylor and Francis, Boca Raton, FL.
- Roberge, M., Berlink, R. G., Xu, L., Anderson, H. J., Lim, L. Y., Cuman, D., Stringer, C. M., Friend, S. H., Davies, P., Vincent, I., Haggarty, S. J., Kelly, M. T., Britton, R., Piers, E., and Andersen, R. J. (1998) High-throughput assay for G2 checkpoint inhibitors and identification of the structurally novel compound isogranulatimide, *Cancer Res.* 58, 5701–5706.
- Roberge, M., Cinel, B., Anderson, H. J., Lim, L., Jiang, X., Xu, L., Bigg, C. M., Kelly, M. T., and Andersen, R. J. (2000) Cell-based screen for antimetastatic agents and identification of analogues of thioxin, eleutherobin, and paclitaxel in natural extracts, *Cancer Res.* 60, 5052–5058.
- McEwen, B. F., Chan, G. K., Zubrowski, B., Savoian, M. S., Sauer, M. T., and Yen, T. J. (2001) CENP-E is essential for reliable bi-oriented spindle attachment, but chromosome alignment can be achieved via redundant mechanisms in mammalian cells, *Mol. Biol. Cell* 12, 2776–2789.
- Nicklas, R. B., Ward, S. C., and Gorbsky, G. J. (1995) Kinetochore chemistry is sensitive to tension and may link mitotic forces to a cell cycle checkpoint, *J. Cell Biol.* 130, 929–939.
- Daum, J. R., Tugendreich, S., Topper, L. M., Jorgensen, P. M., Hoog, C., Hieter, P., and Gorbsky, G. J. (2000) The 3F3/2 anti-phosphoepitope antibody binds the mitotically phosphorylated anaphase-promoting complex/cyclosome, *Curr. Biol.* 10, R850–R852.

20. Kapoor, T. M., and Compton, D. A. (2002) Searching for the middle ground: mechanisms of chromosome alignment during mitosis, *J. Cell Biol.* **157**, 551–556.
21. Wittmann, T., Hyman, A., and Desai, A. (2001) The spindle: a dynamic assembly of microtubules and motors, *Nat. Cell Biol.* **3**, E28–E34.
22. Yen, T. J., Compton, D. A., Wise, D., Zinkowski, R. P., Brinkley, B. R., Earnshaw, W. C., and Cleveland, D. W. (1991) CENP-E, a novel human centromere-associated protein required for progression from metaphase to anaphase, *EMBO J.* **10**, 1245–1254.
23. Yen, T. J., Li, G., Schaar, B. T., Szilak, I., and Cleveland, D. W. (1992) CENP-E is a putative kinetochore motor that accumulates just before mitosis, *Nature* **359**, 536–539.
24. Wood, K. W., Sakowicz, R., Goldstein, L. S., and Cleveland, D. W. (1997) CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome alignment, *Cell* **91**, 357–366.
25. Schaar, B. T., Chan, G. K., Maddox, P., Salmon, E. D., and Yen, T. J. (1997) CENP-E function at kinetochores is essential for chromosome alignment, *J. Cell Biol.* **139**, 1373–1382.
26. Brown, K. D., Wood, K. W., and Cleveland, D. W. (1996) The kinesin-like protein CENP-E is kinetochore-associated throughout poleward chromosome segregation during anaphase-A, *J. Cell Sci.* **109**, 961–969.
27. Cooke, C. A., Schaar, B., Yen, T. J., and Earnshaw, W. C. (1997) Localization of CENP-E in the fibrous corona and outer plate of mammalian kinetochores from prometaphase through anaphase, *Chromosoma* **106**, 446–455.
28. Chan, G. K., Jablonski, S. A., Sudakin, V., Hittle, J. C., and Yen, T. J. (1999) Human BUBR1 is a mitotic checkpoint kinase that monitors CENP-E functions at kinetochores and binds the cyclosome/APC, *J. Cell Biol.* **146**, 941–954.
29. Sudakin, V., Chan, G. K., and Yen, T. J. (2001) Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2, *J. Cell Biol.* **154**, 925–936.
30. Chan, G. K., Schaar, B. T., and Yen, T. J. (1998) Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUBR1, *J. Cell Biol.* **143**, 49–63.
31. Joseph, J., Liu, S. T., Jablonski, S. A., Yen, T. J., and Dasso, M. (2004) The RanGAP1-RanBP2 complex is essential for microtubule-kinetochore interactions *in vivo*, *Curr. Biol.* **14**, 611–617.
32. Joseph, J., Tan, S. H., Karpova, T. S., McNally, J. G., and Dasso, M. (2002) SUMO-1 targets RanGAP1 to kinetochores and mitotic spindles, *J. Cell Biol.* **156**, 595–602.
33. Wilken, N., Senecal, J. L., Scheer, U., and Dabauvalle, M. C. (1995) Localization of the Ran-GTP binding protein RanBP2 at the cytoplasmic side of the nuclear pore complex, *Eur. J. Cell Biol.* **68**, 211–219.
34. Wu, J., Matunis, M. J., Kraemer, D., Blobel, G., and Coutavas, E. (1995) Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region, *J. Biol. Chem.* **270**, 14209–14213.
35. Yokoyama, N., Hayashi, N., Seki, T., Pante, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, T., Aebi, U., Fukui, M., and Nishimoto, T. (1995) A giant nucleopore protein that binds Ran/TC4, *Nature* **376**, 184–188.
36. Salina, D., Enarson, P., Rattner, J. B., and Burke, B. (2003) Nup358 integrates nuclear envelope breakdown with kinetochore assembly, *J. Cell Biol.* **162**, 991–1001.
37. Fan, S., Smith, M. L., Rivet, D. J., 2nd, Duba, D., Zhan, Q., Kohn, K. W., Fornace, A. J. Jr., and O'Connor, P. M. (1995) Disruption of p53 function sensitizes breast cancer MCF-7 cells to cisplatin and pentoxifylline, *Cancer Res.* **55**, 1649–1654.
38. Bunz, F., Dutriaux, A., Lengauer, C., Waldman, T., Zhou, S., Brown, J. P., Sedivy, J. M., Kinzler, K. W., and Vogelstein, B. (1998) Requirement for p53 and p21 to sustain G2 arrest after DNA damage, *Science* **282**, 1497–1501.
39. Dranovsky, A., Vincent, I., Gregori, L., Schwarzman, A., Colflesh, D., Enghild, J., Strittmatter, W., Davies, P., and Goldgaber, D. (2001) Cdc2 phosphorylation of nucleolin demarcates mitotic stages and Alzheimer's disease pathology, *Neurobiol. Aging* **22**, 517–528.
40. Hutchison, M., Lewer, P., and MacMillan, J. (1984) Carbon-13 nuclear magnetic resonance spectra of eighteen derivatives of *ent*-kaur-16-en-19-oic acid, *J. Chem. Soc., Perkin Trans. 1*, 2363–2366.
41. Lever, M. A., Th'ng, J. P., Sun, X., and Hendzel, M. J. (2000) Rapid exchange of histone H1.1 on chromatin in living human cells, *Nature* **408**, 873–876.
42. Zimmer, G., Klenk, H. D., and Herrler, G. (1995) Identification of a 40-kDa cell surface sialoglycoprotein with the characteristics of a major influenza C virus receptor in a Madin-Darby canine kidney cell line, *J. Biol. Chem.* **270**, 17815–17822.
43. Gygi, S. P., Rist, B., Griffin, T. J., Eng, J., and Aebersold, R. (2002) Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags, *J. Proteome Res.* **1**, 47–54.
44. Yates, J. R., III, Eng, J., McCormack, A. L., and (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, *J. Am. Soc. Mass Spectrom.* **5**, 976–989.
45. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, *Anal. Chem.* **74**, 5383–5392.
46. Jablonski, S. A., Chan, G. K., Cooke, C. A., Earnshaw, W. C., and Yen, T. J. (1998) The hBUB1 and hBUBR1 kinases sequentially assemble onto kinetochores during prophase with hBUBR1 concentrating at the kinetochore plates in mitosis, *Chromosoma* **107**, 386–396.
47. Azuma, Y., and Dasso, M. (2002) A new clue at the nuclear pore: RanBP2 is an E3 enzyme for SUMO1, *Dev. Cell* **2**, 130–131.

Structural Insight into the Self-Sacrifice Mechanism of Eneidyne Resistance

Shanteri Singh^{†,‡}, Martin H. Hager[‡], Changsheng Zhang[‡], Byron R. Griffith[‡], Min S. Lee^{†,||}, Klaas Hallenga[§], John L. Markley^{†,§} and Jon S. Thorson^{‡,||,*}

[†]Center for Eukaryotic Structural Genomics, Department of Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706-1544, [‡]Laboratory for Biosynthetic Chemistry, Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin–Madison, 777 Highland Avenue, Madison, Wisconsin 53705, [§]National Magnetic Resonance Facility at Madison, Biochemistry Department, University of Wisconsin–Madison, 433 Babcock Drive, Madison, Wisconsin 53706, and ^{||}University of Wisconsin National Cooperative Drug Discovery Group, School of Pharmacy, University of Wisconsin–Madison, 777 Highland Avenue, Madison, Wisconsin 53705, ^{||}Present address: MinHak Pharmaceuticals, 6356 Paseo Cerro, Carlsbad, California 92009.

The enediynes antitumor antibiotics (Figure 1), are characterized structurally by an enediynes core unit consisting of two acetylenic groups conjugated to a double bond or incipient double bond within a 9- or 10-membered ring (1–4). To date, six naturally occurring 10-membered- (Figure 1, 1–6), and five 9-membered- (Figure 1, 7–11), or “chromoprotein” enediynes, have been elucidated structurally (5–7). In general, these enediynes contain three distinct structural elements: a DNA-recognition unit (*e.g.*, the aryltetrasaccharide of 1), which serves to deliver the metabolite to its target DNA; an activating component (*e.g.*, the methyl trisulfide of 1), which sets the stage for cycloaromatization; and the enediynes “warhead”, which cycloaromatizes to a highly reactive diradical species (*e.g.*, Figure 1, species II) and, in the presence of DNA, results in oxidative strand scission of the targeted sequence (8–11). *In vitro* and *in vivo* studies are consistent with the role of enediynes as DNA-damaging agents and suggest that they may even favor cleavage at certain chromosomal sites and/or tertiary structures (12, 13). Although this extraordinary reactivity invokes incredible potency (some enediynes are >8000-fold more potent than adriamycin), the enediynes are similar to most cytotoxics in their general lack of specificity. However, the clinical success of enediynes has been derived *via* targeting with tumor-specific monoclonal antibodies (mAb) (as in the 1-based MyloTarg to treat acute myelogenous leukemia) (14–17) or through the application of polymer-assisted delivery devices (such as 1-poly[styrene-maleic acid]-conjugated 8).

ABSTRACT The recent discovery of the first “self-sacrifice” mechanism for bacterial resistance to the enediynes antitumor antibiotics, where enediynes-induced proteolysis of the resistance protein CalC inactivates both the highly reactive metabolite and the resistance protein, revealed yet another ingenious bacterial mechanism for controlling reactive metabolites. As reported herein, the first 3D structures of CalC and CalC in complex with calicheamicin (CLM) divulge CalC to be a member of the steroidogenic acute regulatory protein (StAR)-related transfer (START) domain superfamily. In contrast to previous studies of proteins known to bind DNA-damaging natural products (*e.g.*, bleomycins, mitomycins, and nine-membered chromoprotein enediynes), this is the first demonstrated involvement of a START domain fold. Consistent with the CalC self-sacrifice mechanism, CLM in complex with CalC is positioned for direct hydrogen abstraction from Gly113 to initiate the oxidative proteolysis-based resistance mechanism. These structural studies also illuminate, for the first time, a small DNA-binding region within CalC that may serve to localize CalC to the enediynes target (DNA). Given the role of START domains in nuclear/cytosolic transport and translocation, this structural study also may implicate START domains as post-endocytotic intracellular chaperones for enediynes-based therapeutics such as MyloTarg.

*Corresponding author,
jsthorson@pharmacy.wisc.edu.

Received for review July 8, 2006
and accepted July 28, 2006.

Published online August 18, 2006
10.1021/cb6002898 CCC: \$33.50

© 2006 by American Chemical Society

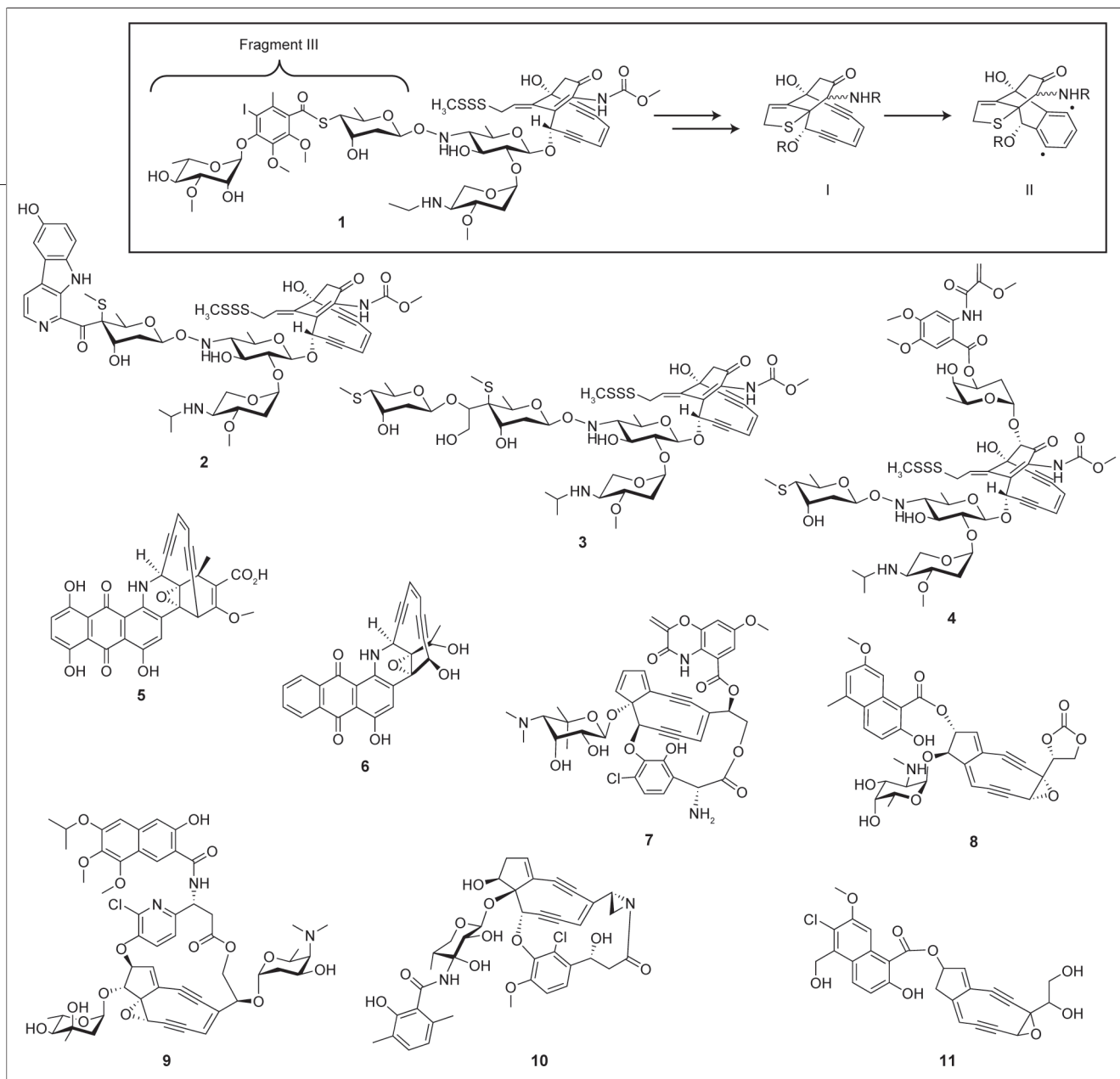


Figure 1. Naturally occurring enediynes: 1, CLM γ_1^1 , (fragment III is highlighted by the bracket, and cycloaromatized products I, dihydrothiophene, and II, CLM ϵ , are also illustrated); 2, shishijimicin A; 3, namenamicin; 4, esperamicin A₁; 5, dynemicin A; 6, uncialamycin; 7, C-1027; 8, neocarzinostatin; 9, kedaracidin; 10, maduropeptin; 11, N1999A₂.

Recently, the locus encoding for calicheamicin (CLM) biosynthesis in *Micromonospora echinospora* was discovered by screening a *M. echinospora* genomic library for cosmids that conferred CLM resistance (6, 18, 19). Preliminary analysis of these cosmids revealed a single gene (*calC*) responsible for CLM resistance. Subsequent *in vitro* studies of CLM inactivation by CalC revealed the mechanism of inactivation to proceed *via* abstraction of a CalC Gly113 C α -hydrogen by the transient enediyne diradical species, thereby quenching the reactive enediyne moiety while generating a CalC Gly113 C α radical (Figure 2) (19). Reminiscent of the mechanism of enediyne-based DNA scission (20), this CalC Gly113 C α

radical subsequently reacts with oxygen to ultimately provide oxidative site-specific proteolysis of CalC. This cumulative reaction, wherein the sacrifice of the CalC protein accompanies CLM inactivation, was noted as the first “self-sacrifice” mechanism of antibiotic resistance. CalC was also shown to inactivate two other members of the 10-membered enediyne family, shishijimicin and namenamicin (Figure 1, 2 and 3, respectively). In an effort to further our understanding of this distinctive CalC enediyne inactivation mechanism, we pursued the structural analyses of CalC and the CalC–CLM complex. The studies reported herein reveal CalC to be a member of the steroidogenic acute regulatory protein (StAR)-

related transfer (START) (21, 22) domain structural superfamily and to contain a small DNA-binding region. The potential implications of these findings to the CalC self-sacrifice mechanism and possibly even the post-endocytotic intracellular transport of enediyne-based therapeutics are also discussed.

RESULTS AND DISCUSSION

Overview of the CalC Structure.

The 15N-HQSC spectrum of full-length CalC1–181 indicated the first 30 residues of CalC to be unstructured as predicted by the PSI-PRED protein structure prediction server (23). N-terminal truncation of CalC provided a protein (CalC27–181) that was functionally equivalent and that gave a 15N-HQSC signature that could be superimposed with that of the full-length protein (Supplementary Figure 1). Given the minimal influence of the 26 N-terminal residues upon activity or structural integrity, CalC27–181 was employed for all studies described herein and for simplicity is referred to as “CalC”. The CalC structure was determined at 30 °C (0.8 mM CalC, 10 mM NaPO₄, pH 7.3, and 150 mM NaCl) by multidimensional heteronuclear NMR spectroscopy of uniformly ¹⁵N- and ¹⁵N/¹³C-labeled protein samples. On the basis of gel filtration and NMR relaxation measurements, CalC was found to exist as a monomer in dilute solution with a tendency to aggregate above ~1 mM CalC. Nearly complete ¹H, ¹⁵N, and ¹³C assignments were obtained from analysis of data from a standard set of NMR experiments. The global fold was established unambiguously from assigned chemical shifts and long-range nuclear Overhauser effects (NOEs).

Distance constraints were derived as described in Methods. The program CNS (24) with ARIA (25) was used to calculate an ensemble of 100 conformers of CalC by simulated annealing on the basis of 2518 unambiguous and 248 ambiguous distance constraints. Additional constraints used in the final calculation of 100 conformers consisted of 160 dihedral angle con-

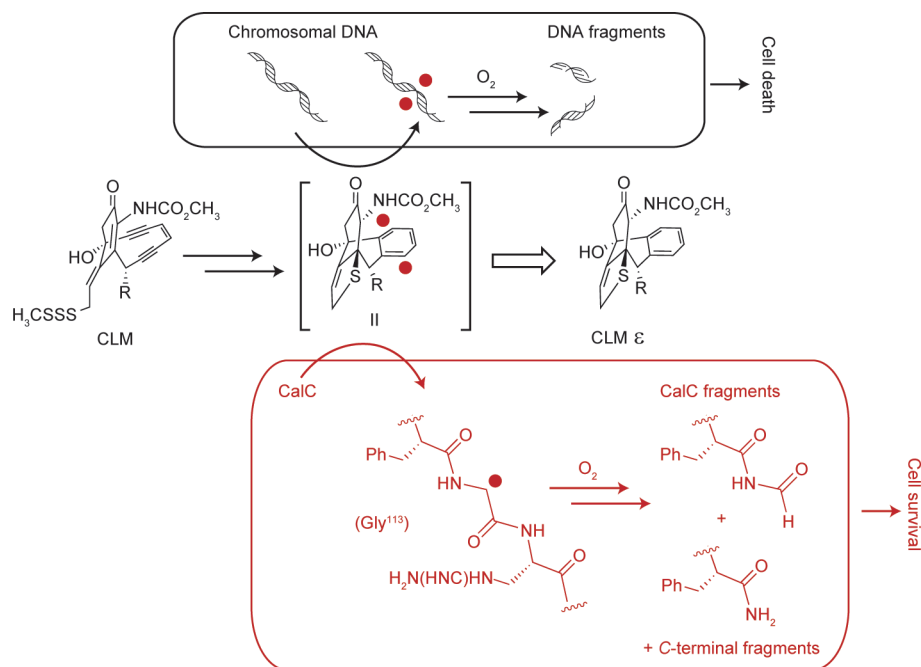


Figure 2. The competing pathways of the enediyne-induced DNA strand scission that leads to cell death (top) and the CalC self-sacrifice enediyne resistance mechanism (bottom).

straints obtained from TALOS (26) on the basis of backbone chemical shift values and 100 hydrogen bond constraints deduced from characteristic NOEs of the secondary structures (see Methods). The 20 conformers with the lowest total energy were used for final analysis. All conformers exhibited good geometry, no violations of distance restraint >0.5 Å, and no dihedral angle violations >5° (Supplementary Table 1). The final superimposed ensemble of 20 conformers (Figure 3, panel a) had root mean square deviation (rmsd) values (over residues 27–181) of 0.65 ± 0.11 Å for backbone heavy atoms and 1.16 ± 0.12 Å for all heavy atoms.

The CalC secondary structural elements (Figure 3) include three α-helices, one ₃₁₀-helix, and seven anti-parallel β-strands. The seven β-strands form a concave open barrel that is closed by helix α1 from the N-terminus, two capping helices α2 and α3 from the C-terminus, and L1 loop. This fold provides a large hydrophobic cavity that spans nearly the entire length of CalC. The hydrophobic core contains aromatic side chains stacked on the aliphatic side chains of other hydrophobic residues. These interactions are mani-

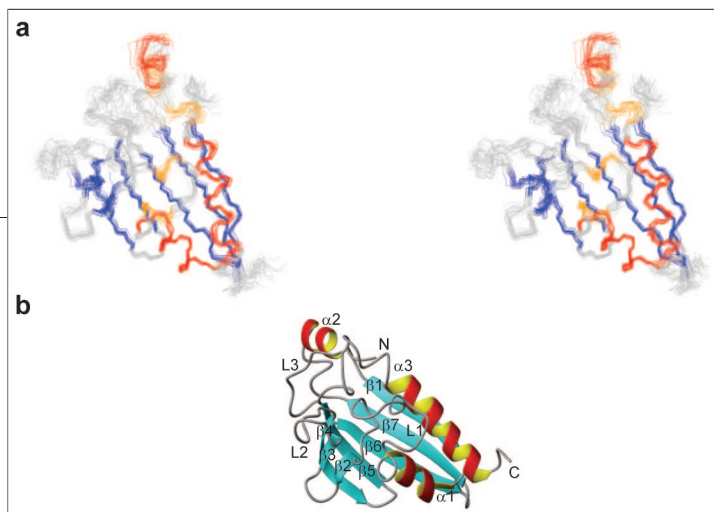


Figure 3. NMR solution structure of CalC. **a)** Backbone stereoview of the ensemble of 20 conformers that represent the structure (residues 27–181). Coloring scheme: blue, β -strand residues (residues 32–38, 71–73, 78–81, 89–98 with bulges at Arg95, 102–106, 123–130, and 134–141); red, α -helical residues (residues 41–50, 149–156, and 163–176); yellow, 3_{10} -helix (residues 53–55); gray, other residues. **b)** Ribbon representation of the representative structure of CalC (the conformer closest to the average). Numbers represent the order of β -strands, α -helices, and loops (overall order $\beta 1$ - $\alpha 1$ - 3_{10} - $\beta 2$ - $\beta 3$ - $\beta 4$ - $\beta 5$ - $\beta 6$ - $\beta 7$ - $\alpha 2$ - $\alpha 3$), and the N- and C-termini of the protein are indicated. The molecular graphics program MOLMOL was used in generating these views of the structure.

fested by low-frequency chemical shifts of protons from several methyl groups. For example, ring current effects from Phe127, Phe123, and Phe53, respectively, account for abnormal chemical shifts observed for the methyl protons of Val137, Val139, and Ile166.

Similarity to the START Domain Superfamily. There are no CalC sequence homologues, and all *de novo* sequence-based models failed to provide any CalC structural insights. The minimum energy structure of CalC was submitted to the fold recognition programs DALI (27) and VAST (28). From this analysis, the closest structural homologues identified displayed very low sequence identity to CalC (4–13%) and include a prokaryotic protein with unknown function, plant allergen

TABLE 1. Structural homologues of CalC

PDB	Z-score	rmsd (Å)	% Identity	Length ^a	Protein
1XUV	12	3.1	16	160	Unknown
1XFS	11	3.3	13	149	Unknown
1XN5	8.2	3.1	17	138	Unknown
1BV1	8.4	2.9	9	126	Plant allergen
1VJH	6.7	3.4	13	102	Plant allergen
1JSS	6.7	3.6	4	110	Lipid binding (STARD4)
1EM2	6	3.3	5	111	Lipid binding (STARD3)
1T17	5.6	4.1	9	148	Unknown
1ND0	4.3	3.3	10	109	Non-heme iron dioxygenase

^aSequence length over which percentage sequence identity is calculated.

proteins (29), and human/murine lipid binding proteins (Table 1). These structural homologues are recognized as part of the START superfamily of proteins, the members of which characteristically bind hydrophobic ligands and frequently appear in multidomain proteins where they regulate the activity of associated domains (21, 22). The START superfamily encompasses a wide range of bacterial, archaeal, and eukaryotic proteins that have been implicated in a variety of cellular functions, including lipid transport and metabolism, signal transduction, and transcriptional regulation (21, 22, 30, 31). A unique structural feature conserved within START domains is a hydrophobic cavity in which subtle modifications dictate ligand/substrate specificity. Apart from this, the START superfamily is characterized by notable structural variations, including the absence or distortion of the first α -helix, differences in the number of strands forming the β -sheet, and the size of the hydrophobic tunnel (30). Consistent with this familial variation, CalC contains two α -helices at the C-terminus ($\alpha 2$ and $\alpha 3$), whereas the closest CalC structural homologues each possess a single continuous C-terminal helix. As discussed below, the unique C-terminal structural feature of CalC appears to be involved in DNA binding.

The CalC–CLM Complex. CLM (Figure 1 and Figure 2), which contains a core bicyclo (7.3.1) tridecadiyne moiety (enediyne warhead) appended by an aryltetrasaccharide chain (6), is highly hydrophobic and is sparingly soluble in aqueous buffer. We used the NMR chemical shift perturbation method (32) to determine the interaction surface of the CalC protein with CLM. A series of 2D-¹H–¹⁵N heteronuclear single quantum correlation (HSQC) spectra of [¹⁵N]-CalC in 15% dimethyl sulfoxide (DMSO) were recorded to follow the effects of titration with ligand/substrate from a CLM-DMSO stock solution. Chemical shift changes upon titration identified residues that interact directly with CLM along with those affected indirectly by substrate/ligand-binding (Supplementary Figure 2). Upon CLM titration, chemical shift changes were observed mainly in the loop regions, indicating their involvement in the interaction with CLM (Supplementary Figure 2). However, although the chemical shift perturbations provided a general indication of the CalC–CLM interaction surface, the limited solubility of CLM, even in 15% DMSO, complicated more precise refinement of the CalC–CLM complex structure.

To obtain a more accurate and quantitative measurement of the protein–ligand interaction, a more soluble CLM fragment (Figure 1, fragment **III**) was employed. Titration of [U-¹⁵N]-CalC with fragment **III** (Figure 4) followed by NMR 2D-1H–15N HSQC, identified a set of residue perturbations similar to those previously found with CLM binding. As with CLM, large chemical shift changes (>0.13 ppm in ¹H and 1.1 ppm in ¹⁵N) were observed mainly in the loop regions, indicating their involvement in the interaction with fragment **III**. These changes increased upon continued addition of fragment **III**. Signals corresponding to Thr63, Arg108, Gly162, and the side chain of Trp163 broadened and disappeared upon the addition of 1 molar equivalent of fragment **III**. On the basis of the largest observed chemical shift change (100 Hz for the amide proton of Leu109), we determined a lower limit of ~ 600 s⁻¹ for the dissociation rate constant, which corresponds to an estimated K_d of ~ 60 μ M (assuming k_{on} to be under diffusion control with a value of $\sim 10^7$ M⁻¹ s⁻¹). The titration results suggest that the protein is in intermediate to fast exchange on the NMR time scale in its complex with fragment **III**, as judged by ¹H–¹⁵N HSQC cross-peak broadening or repositioning. It is not unreasonable to assume that the natural substrate, CLM, binds more tightly to CalC than the highly truncated fragment **III**.

The ¹H and ¹⁵N chemical shift changes for backbone amides (Figure 4) define the CalC-binding surface for fragment **III**/CLM with CalC. Both CLM and fragment **III** induce similar conformational changes in CalC upon binding as indicated by similar chemical shift changes. Chemical shift perturbations upon binding of CLM/fragment **III** were found to be concentrated within loop regions L1 (residues 53–70), L2 (residues 79–93), L3 (residues 107–121), and helices $\alpha 2$ (residues 146–156) and $\alpha 3$ (residues 159–170). Most of the hydrophobic interactions within the determined CalC–CLM complex involve residues in the loop L1 (Phe53, Pro54, Pro58), helix $\alpha 3$ (Met159, Trp163, Pro164, Ile166), and loop L3 (Phe112, Leu109) (Figure 4, panel d). In addition, charged residues such as Arg108, Arg114 from loop L3, Asn60 from loop L1, and Asp160 and Thr165 from $\alpha 3$ -helix contribute to the electrostatic interactions. The observed NMR shift perturbations upon fragment **III**/CLM-binding were used as input to HADDOCK (33) to derive a model for the 1:1 CalC–CLM complex. It should also be noted that including the NMR-elucidated interacting residues and the biochemically established CLM

enediayne–Gly113 interaction (19, 20) as an ambiguous interaction restraint (AIRs) in HADDOCK did not significantly alter the final structural model. In the resulting model (Figure 4), the enediayne warhead is oriented toward the neutral loop (L3) and a $26 \text{ \AA} \times 14 \text{ \AA} \times 14 \text{ \AA}$ hydrophobic tunnel engulfs the CLM aryltetrasaccharide (Figure 4, panel c). A number of intermolecular hydrophobic and hydrogen-bonding interactions account for the specific recognition within this complex (Figure 4, panel d).

DNA binding to CalC. There exists no prior evidence for CalC–DNA interaction. However, given that several START domain proteins are either associated with, or directly involved in, transcriptional regulation (31), we used the NMR chemical shift perturbation method described above to test the ability of CalC to bind DNA. Given that there was no *a priori* knowledge regarding the putative CalC DNA “target”, a structurally characterized double-stranded DNA (dsDNA) 11-mer derived from the *Escherichia coli* RNA-polymerase promoter region (34) was employed for this initial study.

Titration of [U-¹⁵N]-CalC with dsDNA, followed by NMR 2D-1H–15N HSQC, identified a set of residue perturbations. These perturbations revealed that CalC binds to the dsDNA probe and localized the CalC–DNA interaction surface to the unique C-terminal helix, $\alpha 2$ within CalC (Figure 5). Large chemical shift perturbations were restricted primarily to helix $\alpha 2$, (Lys150–Asn157), while residues spatially close to the helix $\alpha 2$ (Phe112, Gly113, Arg114, Ile115 from Loop L3, and Tyr28 and Asp29 from the N-terminus) also showed some perturbation in their chemical shift values (Figure 5). Notably, the high-frequency shifts of several backbone nitrogens and amide protons (Asp29, Arg114, Lys150, Lys153, Arg154, and Asn157) are indicative of possible hydrogen bonding to the DNA phosphate backbone (35). From the maximum observed chemical shift change of 225 Hz for amide proton of the residue Lys153, we determined a lower limit of ~ 1410 s⁻¹ for the dissociation rate constant, which corresponds to an estimated K_d range of ~ 14 –141 μ M (assuming k_{on} to be under diffusion control with a value of $\sim 10^6$ – 10^7 M⁻¹ s⁻¹). The observed changes indicate that the CalC–DNA interaction is in fast exchange regime on NMR chemical shift time scale, indicating that the binding is weak. Many proteins that interact with specific DNA sequences bind to nonspecific DNA sequences with weak affinity, and such initial nonspecific interactions play an important

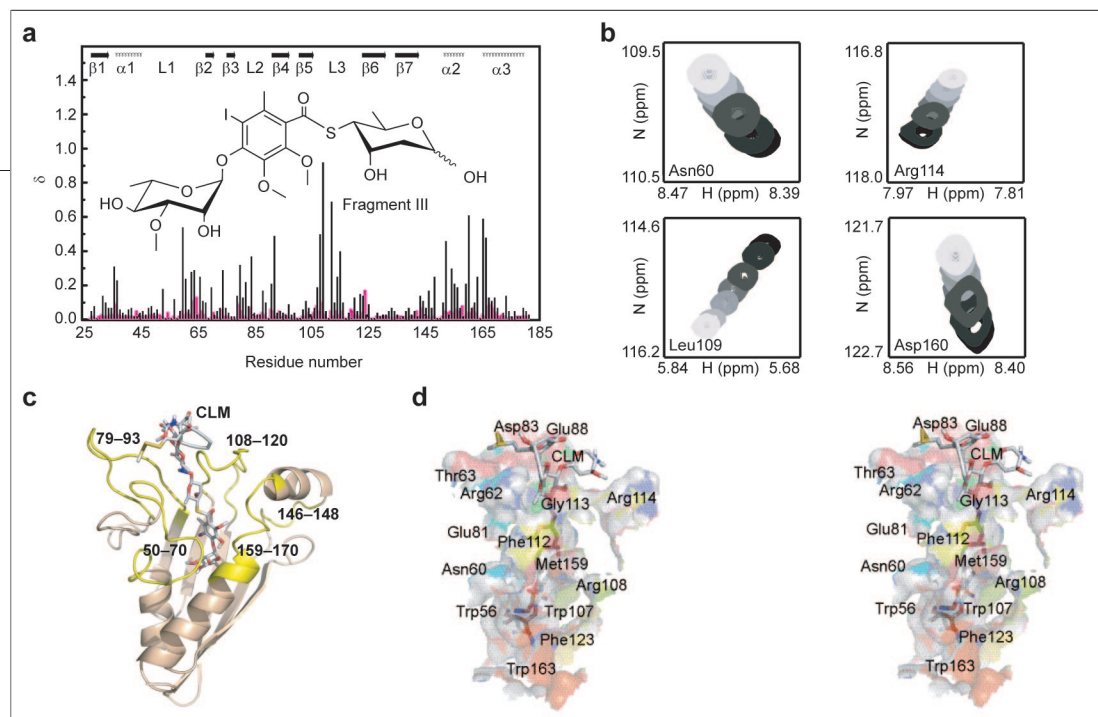


Figure 4. NMR chemical shift titration curves for the binding of fragment III to [U-¹⁵N]-CalC. **a)** Changes in chemical shifts upon titration of CalC with fragment III plotted by residue number: black, backbone amide nitrogens; magenta, amide protons. The absence of a bar in the plot indicates the presence of a proline residue or an unmeasured shift due to overlap. **b)** Regions of the HSQC spectrum displaying changes in cross-peak positions of various residues (Asn60, Leu109, Phe112, and Asp160) upon addition of fragment III at concentrations of 0, 0.1, 0.4, 0.7, 1.1, 1.5, 2.0, and 3.5 mM (decreasing darkness). **c)** CalC–CLM interacting surface: CalC is represented by the ribbon model, CLM as a stick model; protein residues with significant chemical shift changes in the presence of CLM are colored yellow. **d)** Stereoview of the CalC–CLM interaction surface.

role in translocation of DNA binding proteins to their target sequence *in vivo* (36). While the CalC–DNA binding may be tighter to a specific DNA sequence, such specific CalC binding sequences have yet to be determined. The restraints derived from NMR shift perturbations served as input to HADDOCK (33) to derive a model for the CalC–DNA complex. The interaction of positively charged residues from helix α 2 with the phosphate backbone of the DNA is shown (Figure 5).

Mechanistic Implications. To date, the structural elucidation of proteins that sequester highly reactive metabolites capable of DNA damage have revealed the involvement of two major structural folds (37). The first structural fold used for sequestering DNA-damaging metabolites is a dimer of monomers containing two $\beta\alpha\beta\beta$ -repeats found both within the bleomycin (BLM) and mitomycin (MTM) resistance proteins (BlmA and Mrd, respectively). In both BlmA and Mrd, this $\beta\alpha\beta\beta$ -rich structure provides a concave binding pocket (37). This pocket within BlmA predominately employs electrostatic interactions to bind BLM, while the analogous cavity in Mrd utilizes π -stacking interactions of aromatic residues to bind MTM. The second structural fold known to control DNA-damaging natural products is the immunoglobulin fold used by the 9-membered chromoprotein

enediynes (Figure 1, 7–11). Specifically, stabilizing 9-membered enediyne carrier proteins employ this β -sheet-rich fold to provide a hydrophobic pocket that protects the extremely labile enediyne chromophores (37). Distinct from BlmA, Mrd, and enediyne chromoproteins, CalC is a member of the START domain structural superfamily. Thus, this study illuminates, for the first time, the START domain as another fold used by nature to specifically sequester/control highly reactive metabolites capable of DNA damage.

The START domain is a conserved protein module that displays remarkable functional versatility and is found in eukaryotes, archaea, and bacteria (30). START domains are involved in diverse cellular functions that include ligand binding [e.g., the START domain of human MLN64, STARD3, which binds to and regulates the transport of cholesterol (38, 39)], catalysis [e.g., the cyclase/aromatase TcmN in tetracenomycin biosynthesis (40)], and even signal transduction [e.g., the potential tumor suppressor STARD12 (21, 41)]. The most prominent structural feature of the START domain is the presence of a helix-grip fold, which forms a binding pocket for hydrophobic ligands (30). Within CalC, this prominent structural fold serves to bind the CLM aryltetrasaccharide and thereby bury the DNA-recognition unit of CLM. In the

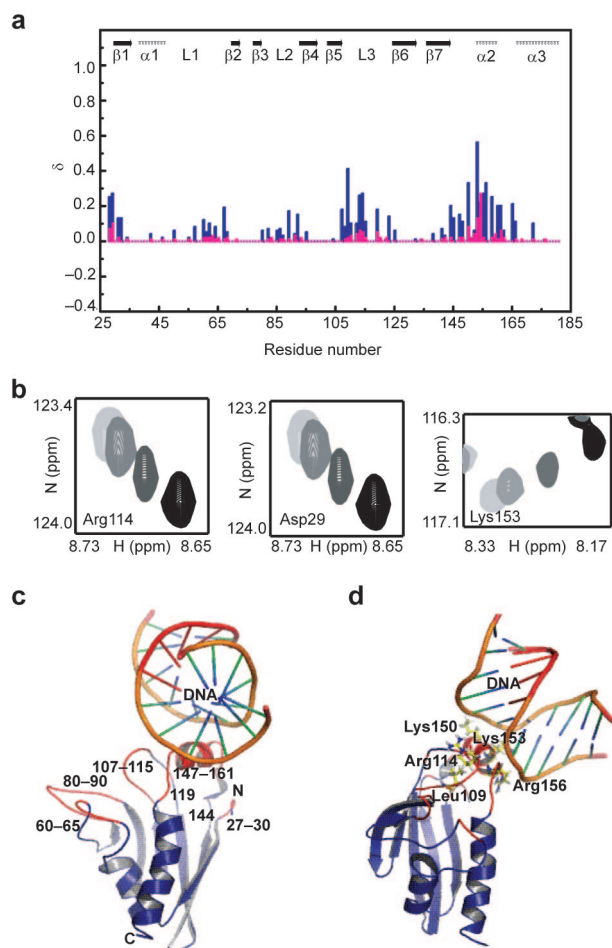


Figure 5. NMR chemical shift titration curves for the binding of DNA [U-¹⁵N]-CalC. a) Changes in chemical shifts upon titration of CalC with DNA, (blue) backbone amide nitrogen and (magenta) amide proton, plotted as a function of the residue number. The absence of a bar in the plot indicates the presence of a proline residue or an unmeasured shift due to overlap. b) Regions of HSQC spectra displaying changes in the positions of cross peaks corresponding to various residues (Asp29, Arg114, Lys153) at DNA concentrations of 0, 0.05, 0.08, and 0.1 mM (decreasing darkness). c) CalC–DNA interaction surface: CalC is represented by the blue ribbon; residues exhibiting significant chemical shift perturbations are shown in red. d) CalC residues exhibiting the largest chemical shift perturbations upon titration with DNA.

CalC–CLM complex, the trisulfide trigger of CLM is solvent-exposed to facilitate activation, while the enediyne warhead is positioned for Gly113 α -hydrogen abstraction to initiate the proteolytic CalC self-sacrifice event. Thus, the structure of the CalC–CLM complex is consistent with the previous in-depth mechanistic

studies of the inactivation of CLM through CalC self-sacrifice and the established minimal enediyne-disaccharide unit required for CalC recognition (19).

The demonstration that START domains are efficient scaffolds for CLM-binding may also have broader relevance. For example, in the context of CLM-based therapeutics such as MyloTarg, the current route of CLM intracellular transport from the lysosome to the nucleus (where the enediyne ultimately renders its lethal effects) remains unknown (16, 17). Interestingly, human START domains are broadly distributed within cells, including the nucleus (STARD4, D5, D6, and D10), cytosol (STARD2, D4, D5, D11, D13, and D15), mitochondria (STARD10), Golgi (STARD11), plasma membrane (STARD12), and even endosomes (STARD3) (21). Therefore, the present study, which reveals the START domain as an efficient enediyne-binding scaffold, presents a compelling basis from which to propose that START domain proteins may also serve to chaperone and/or mediate the requisite intracellular transport and translocation of CLM within human cells.

CLM-induced DNA strand scission is initiated by reductive activation. For example, in the presence of glutathione (GSH), CLM initially binds DNA reversibly, and the DNA-bound CLM reacts with GSH to form GSH–CLM disulfide as the major product (9). A number of long-lived, DNA-free intermediates exist en route to this key GSH–CLM disulfide, which can also dissociate from DNA. The GSH–CLM disulfide intermediate ultimately leads to the hetero-Michael adduct (Figure 1, dihydrothiophene I), which rapidly cycloaromatizes to give the reactive 1,4-diradical species required for hydrogen abstraction from the DNA backbone (10, 42, 43). The current structural study revealed a minimum CalC–fragment III association of 10^4 M^{-1} , and efforts to elucidate an association for CalC–CLM were prohibited by CLM solubility. Given the CLM–DNA association is estimated to be 10^6 – 10^8 M^{-1} (44, 45) and *in vitro* CLM inactivation via CalC self-sacrifice is favored even in the

presence of excess DNA (19, 20), we conclude the estimated fragment III–CalC weak association may not reflect the true CalC–CLM binding affinity.

The current NMR study also clearly divulged, for the first time, evidence for association of CalC with DNA. In the context of “self-sacrifice”, the apparent localization of CalC to the intended enediyne target (DNA) in conjunction with an apparent affinity of CalC for CLM is consistent with the demonstrated efficient capture and inactivation

of CLM by CalC to ultimately prevent DNA damage. Given that the interaction of many regulatory DNA-binding proteins with their target sites is usually preceded by binding to nonspecific DNA, the current study does not support or refute the existence of specific CalC-binding sequences. It will be of interest in future studies to determine whether CalC displays sequence specificity and whether the observed CalC–DNA interaction offers additional direct or indirect effects within the CLM producing bacterium.

METHODS

Cloning, Expression, and *in Vivo* Assays. The N-terminal truncation expression constructs of CalC were amplified from pJB2011 using the following primer pairs: CalC27-181, 5'-AAGCCGAAG-CATATGAAGTACGACCGTTC-3'/5'-ATATATAAGCTTTCACCTTCGCGCCCTTCC-3'; CalC¹⁻¹¹², 5'-ATATATCATATGCATCACCATCACCATCACTACTCAGGAGAAGACCGCA-3'/5'-CGGGTCAAGCTTTCAGAACCGTTGAGCCG-3'; CalC27-112, 5'-ATATATCATATGCATCACCATCACCATCACTACTCAGGACCGTTCGTC-3'/5'-CGGGTCAAGCTTTCAGAACCGTTGAGCCG-3'; and CalC114-181, 5'-AACGGCCATATGCGGATCGA-CCCGGAC-3'/5'-ATATATAAGCTTTCAGTATGGTATGGTATGATGCTTCTTCGCCCCCTTCC-3'. After restriction digestion with NdeI/HindIII (Promega), the PCR products were ligated into NdeI/HindIII-linearized pET21a (Novagen) to yield constructs pMH16 (CalC27-181), pMH20 (CalC1-112), pMH21 (CalC27-112), and pMH22 (CalC114-181), respectively. Each expression plasmid was confirmed by sequencing and subsequently transformed into *E. coli* strain BL21(DE3)Gold (Stratagene) for expression at 37 °C. *In vivo* CLM resistance assays for CalC, CalC27-181, CalC1-112, CalC27-112, and CalC114-181 were performed on LB-agar plates supplemented with 75 μg mL⁻¹ carbenicillin, 1 mM isopropyl-β-D-thiogalactopyranoside, and 30 ng mL⁻¹ CLM as previously described (19). In this study, cells containing the CalC expression construct, as well as control cells containing empty vector, were grown to an o.d. of ~2, and an equal cell density from each starter culture was plated. Plates were incubated for 12 h at 37 °C, and the clones were assessed for survival.

Labeling. Labeling of CalC27-181, CalC1-112, CalC27-112, and CalC114-181 with ¹⁵N and double-labeling of CalC27-181 with ¹³C and ¹⁵N was achieved by growth of BL21(DE3)Gold transformed with pMH16, pMH20, pMH21, and pMH22 in modified minimal medium according to Zhao *et al.* (46). The minimal salts medium for ¹⁵N/¹³C-labeling was supplemented with 4 g of ¹³C-glucose L⁻¹ and adjusted to pH 7.2. Protein expression was performed in 2.5 L Fernbach flasks, and cells were grown for 14 h at 37 °C to an o.d. of 4.8. Cells were harvested by centrifugation at 3000g for 20 min.

Protein Purification. Cells were resuspended in binding buffer (50 mM sodium phosphate, 25 mM NaCl, pH 6.8) and disrupted in a pressure cell (Thermo Electron) at 12,000 psi. Insoluble debris was removed by centrifugation at 10,000g for 1 h, and the supernatant was loaded onto a HiTrap SP FF column (Amersham Biosciences) for ion exchange chromatography using a NaCl-gradient (50 mM sodium phosphate, pH 6.8, 25 mM–1 M NaCl over 41 mL, A₂₈₀ detection, 0.5 mL min⁻¹). Under these conditions, the desired protein eluted at 175 mM NaCl, and CalC-containing fractions were collected and concentrated by ultrafiltration (10,000 MWCO-membrane, Centrprep and Centricon YM10, Millipore). Concentrated protein was loaded onto a HiPrep 16/60 Sephacryl S-200 HR column (Amersham Biosciences) for gel filtration in 50 mM BisTris, 200 mM

NaCl, pH 7.3 (A₂₈₀ detection, 0.5 mL min⁻¹). CalC-containing fractions (67–85 mL) were collected, concentrated, and stored in 5 mM BisTris, 100 mM NaCl, 10 mM MgCl₂, pH 7.3 (NMR buffer).

Generation of CLM Fragment III. Fragment III (Figure 1) was produced by refluxing 10 mg of CLM α₃ in 10 mL of wet acetone with pyridinium *p*-toluene-sulfonate (0.1 equiv) at 70 °C. The hydrolysis was monitored by HPLC using a reverse-phase column (Phenomenex Luna C18, 250 × 4.6 mm) with a gradient of 10%–90% acetonitrile in H₂O over 20 min at a flow rate of 1 mL min⁻¹. The products were detected under UV-280 nm by a PDA detector with a retention time for CLM α₃ of 15.5 min and for fragment III of 11.6 min. After 20 h, solvent was evaporated, and the fragment III was purified by preparative HPLC isolation using a reverse-phase column (DiscoveryBIO C18, 25 cm × 21.2 mm, 10 μm, Supelco) with a gradient of 10–90% acetonitrile in H₂O over 20 min at a flow rate of 10 mL min⁻¹. After lyophilization, 0.8 mg of fragment III was obtained (8% yield). The product was confirmed by ¹H NMR on Varian UNITY INOVA 500 MHz using a CapNMR probe.

Binding Studies. For CLM-binding studies, CLM was dissolved in 45–50 μL of DMSO to give a stock solution of 15–20 mM CLM. 2D ¹H–¹⁵N HSQC spectra of CalC (100 μM) were recorded in NMR buffer with the successive addition of increasing CLM at 30 °C. The 2D ¹H–¹⁵N-HSQC of 100 μM CalC in NMR buffer containing 15% DMSO was used as a control. For fragment III-binding studies, fragment III was dissolved in NMR buffer to give a stock solution of 15 mM. The 2D ¹H–¹⁵N HSQC spectra of 400 μM CalC with 0.14, 0.57, 1.0, 1.57, 2.14, 2.85, and 5.0 equiv of fragment III were recorded at 30 °C. For CalC–DNA studies, the DNA employed was dsDNA: 1, 5'-GCATATGATAG and 5'-CTATCATATGC and the 2D ¹H–¹⁵N HSQC spectra of CalC (100 μM) recorded in NMR buffer with 0.05, 0.21, 0.42, 0.84, and 1.26 equiv of DNA at 30 °C. All 1H–15N HSQC spectra were recorded using a time domain data size of 400 t₁, 1596 t₂ complex points, and 16 transients per t₁ increment.

NMR Spectra Used for the Structure Determination. NMR experiments were carried out at 30 °C on Varian UNITY INOVA 900 MHz, 800 MHz, and 600 MHz spectrometers equipped with a ¹H–¹³C–¹⁵N triple-resonance cold probe (800 MHz, 600 MHz) with a 5 mm z-shielded gradient or a RT probe (900 MHz) with a 5 mm x, y, z-shielded gradient. For the backbone resonance assignments, (3D) HNCO, HNCACB, CBCA(CO)NH, NOESY-(¹⁵N, ¹H)-HSQC spectra were recorded as previously described (47), and for the side chain resonance assignments, 3D HBHA(CO)NH, H(CO)NH, C(CO)NH, H(C)CH TOCSY, and (H)CCG TOCSY data sets were recorded. NOE distance constraints were obtained from 3D-NOESY(¹⁵N, ¹H)-HSQC, and 3D NOESY-(¹³C, ¹H)-HSQC spectra with 100 ms mixing times. Spectra were processed with NMRPipe (48) and analyzed using NMRView (49) software packages.

Chemical Shift-Derived and Hydrogen Bond Restraints. Assigned chemical shifts were determined for 97% of the nuclei (BioMagResBank accession code 6726). The ¹³C^α, ¹³C^β, ¹³C^γ, ¹H^α, and ¹⁵N chemical shifts of the assigned residues served as input for

the TALOS program (26). For better convergence, a number of hydrogen bond restraints were introduced for the backbone amide protons on the basis of amide ^1H - ^2H exchange results, $^{13}\text{C}^\alpha$ -/ $^{13}\text{C}^\beta$ secondary shifts, and NOE data. Hydrogen bonds were enforced by using the following restraints: 1.8–2.3 Å for $d_{(\text{N-H}, \text{O})}$ and 2.7–3.3 Å for $d_{(\text{N}, \text{O})}$. The hydrogen bonds between N–H_i and O–C_j in the β -sheet structures were included as restraints only if the β -sheet interstrand d_{NN} (i, j), and $d_{\alpha\text{N}}$ (i, j) NOE cross peaks were observed. Hydrogen bond constraints for α -helices were included when NOEs corresponding to the secondary structure $d_{\alpha\text{N}}$ (i, i+3) for α -helices were observed.

Structure Calculations. Structure calculations (50) were performed with CNS (24) using the ARIA setup and simulated annealing protocols (25). The best 20 structures out of 100 calculated were selected on the basis of lowest total energy. PROCHECK-NMR (51) was used to assess the quality of the final ensemble of conformers. Analysis of the final 20 structures showed that 93.3% were in the favored region of the Ramachandran plot. Structures were visualized with the programs MOLMOL (52) or PyMOL (www.pymol.org).

Modeling the CalC–CLM and CalC–DNA Complexes. Coordinates of CLM were taken from the NMR structure of the CLM–DNA complex (2PIK) (53). Topology and parameter files were generated by the PRODRG server (54). The experimentally determined distance and dihedral restraints for CalC (Supplementary Table 1) were applied in a simulated annealing protocol using CNS and HADDOCK (33). AIRs were generated based on the chemical shift mapping data of CalC with fragment III. “Active” and “passive” residues were distinguished on the basis of their amount of chemical shift perturbation; in the NMR titration data, active residues correspond to all residues showing a significant chemical shift perturbation (>0.03 ppm for ^1H and >0.3 ppm for ^{15}N) upon complex formation. From the list of active and passive residues, 24 AIRs were defined with an upper distance boundary of 3.5 Å. A total of 500 conformers of the CalC–CLM complex was generated using only the AIRs, van der Waals energy, and electrostatic terms in CNS. The 100 conformers with lowest molecular energies were subsequently subjected to semiflexible simulated annealing and refinement with explicit water and only backbone restraints for residues outside the interface. Residues 56–66, 79–92, 106–119, 130–144, and 150–168 of CalC, and full CLM were allowed to be flexible in all stages of the docking procedure. After the calculation, structures were ranked according to their intermolecular energy (the sum of electrostatic, van der Waals, and AIR energy terms). Ten structures with low restraint energy were accepted as a final structure.

For the CalC–DNA complex, the DNA coordinates were taken from the DNA structure 1SKP, d(gcatatgatag).d(ctatcatatgc), a consensus sequence for promoters recognized by sigma-k RNA polymerase (34). Fourteen AIRs were defined with an upper distance boundary of 3.5 Å. A total of 20 conformers of the CalC–DNA complex was generated using only the AIRs, van der Waals energy, and electrostatic terms in CNS. The 10 conformers with lowest molecular energies were subsequently subjected to semiflexible simulated annealing and refinement with explicit water and only backbone restraints for residues outside the interface. Residues 28–32, 62–65, 89–93, 109–116, and 148–155 of CalC, and the middle four base pairs in DNA were allowed to be flexible in all stages of the docking procedure. After the calculation, structures were ranked according to their intermolecular energy (the sum of electrostatic, van der Waals, and AIR energy terms). The structure with low restraint energy was accepted as a final structure.

Accession Codes. The chemical shifts of CalC at pH 7.3 and 30 °C have been deposited in the BioMagResBank under accession number 6726. The atomic coordinates of the ensemble of 20 structures that represent the solution structure of CalC, together with the complete list of constraints used for the final structure calculation, have been deposited in the Protein Data Bank (PDB), under accession number 1ZXF. Atomic coordinates of the ensemble of 10 struc-

tures that represent the model of the complex of CalC–CLM and a minimum restraint energy of complex of CalC–DNA, along with the complete list of AIR restraints derived from chemical shift perturbations, have been deposited in the PDB, under accession number 2GKC and 2GKD, respectively.

Acknowledgments: This work was supported in part by the National Institutes of Health grants CA84374, AI52218, and GM70637, and a NCDDG grant from the National Cancer Institute (U19 CA113297). M. H. Hager is a postdoctoral fellow of the German Academy of Scientists Leopoldina (BMBF-LPD 9901/8-82) and J. S. Thorson is a Romnes Fellow. NMR data were collected at the National Magnetic Resonance Facility at Madison, which is supported in part by grants P41 RR02301, P41 GM66326, the National Institutes of Health, and a Protein Structure Initiative through Grant P50 GM64598. We gratefully thank Wyeth Research for providing the parent enediyne and M. Tonelli from NMRFAM for providing dsDNA for binding studies.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Jones, G. B., and Fouad, F. S. (2002) Designed enediyne antitumor agents, *Curr. Pharm. Des.* 8, 2415–2440.
2. Danishefsky, S. J., and Shair, M. D. (1996) Observations in the chemistry and biology of cyclic enediyne antibiotics—total syntheses of calicheamicin and dynemicin, *J. Org. Chem.* 61, 16–44.
3. Smith, A. L., and Nicolaou, K. C. (1996) The enediyne antibiotics, *J. Med. Chem.* 39, 2103–2117.
4. Davies, J., Wang, H., Taylor, T., Warabi, K., Huang, X. H., and Andersen, R. J. (2005) Uncialamycin, a new enediyne antibiotic, *Org. Lett.* 7, 5233–5236.
5. Thorson, J. S., Shen, B., Whitwam, R. E., Liu, W., Li, Y., and Ahlert, J. (1999) Enediyne biosynthesis and self resistance. A Progress Report, *Bioorg. Chem.* 27, 172–88.
6. Thorson, J. S., Sievers, E. L., Ahlert, J., Shepard, E., Whitwam, R. E., Onwueme, K. C., and Ruppen, M. (2000) Understanding and exploiting nature’s chemical arsenal: the past, present and future of calicheamicin research, *Curr. Pharm. Des.* 18, 1841–1879.
7. Oku, N., Matsunaga, S., and Fusetani, N. (2003) Shishijimicins A-C, novel enediyne antitumor antibiotics from the ascidian *Didemnum proliferum*(1), *J. Am. Chem. Soc.* 125, 2044–2045.
8. Zein, N., Sinha, A. M., McGahren, W. J., and Ellestad, G. A. (1988) Calicheamicin gamma 11: an antitumor antibiotic that cleaves double-stranded DNA site specifically, *Science* 240, 1198–1201.
9. Myers, A. G., Cohen, S. B., and Kwon, B. M. (1994) A study of the reaction of calicheamicin gamma(1) with glutathione in the presence of double-stranded DNA, *J. Am. Chem. Soc.* 116, 1255–1271.
10. DeVoss, J. J., Hangeland, J. J., and Townsend, C. A. (1990) Characterization of the in vitro cyclization chemistry of calicheamicin and its relation to DNA cleavage, *J. Am. Chem. Soc.* 112, 4554–4556.
11. Kappen, L. S., and Goldberg, I. H. (1993) DNA conformation-induced activation of an enediyne for site-specific cleavage, *Science* 261, 1319–1321.
12. Watanabe, C. M. H., Supekova, L., and Schultz, P. G. (2002) Transcriptional effects of the potent enediyne anti-cancer agent Calicheamicin gamma (I)(1), *Chem. Biol.* 9, 245–251.
13. Stassinopoulos, A., Ji, J., Gao, S., and Goldberg, I. H. (1996) Solution structure of a two-base DNA bulge complexed with an enediyne cleaving analog, *Science*. 272, 1943–1946.
14. Damle, N. K. (2004) Tumour-targeted chemotherapy with immun-conjugates of calicheamicin, *Exp. Opin. Biol. Ther.* 4, 1445–1452.
15. Abe, S., and Otsuki, M. (2002) Styrene maleic acid neocarzinostatin treatment for hepatocellular carcinoma, *Curr. Med. Chem.* 2, 715–726.

16. Jedema, I., Barge, R. M., van der Velden, V. H., Nijmeijer, B. A., van Dongen, J. J., Willemze, R., and Falkenburg, J. H. (2004) Internalization and cell cycle-dependent killing of leukemic cells by Gemtuzumab Ozogamicin: rationale for efficacy in CD33-negative malignancies with endocytic capacity, *Leukemia* **18**, 316–325.
17. Wu, A. M., and Senter, P. D. (2005) Arming antibodies: prospects and challenges for immunoconjugates, *Nat. Biotechnol.* **23**, 1137–1146.
18. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M., and Thorson, J. S. (2002) The calicheamicin gene cluster and its iterative type I enediyne PKS, *Science* **297**, 1173–1176.
19. Biggins, J. B., Onwueme, K. C., and Thorson, J. S. (2003) Resistance to enediyne antitumor antibiotics by CalC self-sacrifice, *Science* **301**, 1537–1541.
20. Biggins, J. B., Prudent, J. R., Marshall, D. J., Ruppen, M., and Thorson, J. S. (2000) A continuous assay for DNA cleavage: the application of “break lights” to enediynes, iron-dependent agents, and nucleases, *Proc. Natl. Acad. Sci. U.S.A.* **25**, 13537–13542.
21. Alpy, F., and Tomasetto, C. (2005) Give lipids a START: the StAR-related lipid transfer (START) domain in mammals, *J. Cell. Sci.* **118**, 2791–2801.
22. Soccio, R. E., and Breslow, J. L. (2003) StAR-related lipid transfer (START) proteins: mediators of intracellular lipid metabolism, *J. Biol. Chem.* **278**, 22183–22186.
23. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server, *Bioinformatics* **16**, 404–405.
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr., D: Biol. Crystallogr.* **54**, 905–921.
25. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation, *Bioinformatics* **19**, 315–316.
26. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR* **13**, 289–302.
27. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol.* **233**, 123–138.
28. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) Surprising similarities in structure comparison, *Curr. Opin. Struct. Biol.* **6**, 377–385.
29. Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lowenstein, H., and Spangfort, M. D. (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy, *Nat. Struct. Biol.* **3**, 1040–1045.
30. Iyer, L., Koonin, E., and Aravind, L. (2001) Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily, *Proteins: Struct., Funct., Genet.* **43**, 134–144.
31. Ponting, C. P., and Aravind, L. (1999) START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins, *Trends Biochem. Sci.* **24**, 130–132.
32. Dekker, N., Cox, M., Boelens, R., Verrijzer, C. P., van der Vliet, P. C., and Kaptein, R. (1993) Solution structure of the POU-specific DNA-binding domain of Oct-1, *Nature* **362**, 852–855.
33. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information, *J. Am. Chem. Soc.* **125**, 1731–1737.
34. Tonelli, M., Ragg, E., Bianucci, A. M., Lesiak, K., and James, T. L. (1998) Nuclear magnetic resonance structure of d(GCATATGATAG). d(CTATCATATGC): a consensus sequence for promoters recognized by σ^k RNA polymerase, *Biochemistry* **37**, 11745–11761.
35. Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. (2000) Crystal structure of the holliday junction DNA in complex with a single RuvA tetramer, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8257–8262.
36. Jen-Jacobson, L. (1997) Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state, *Biopolymers* **44**, 153–180.
37. Galm, U., Hager, M. H., Van Lanen, S. G., Ju, J., Thorson, J. S., and Shen, B. (2005) Antitumor antibiotics: bleomycin, enediynes, and mitomycin, *Chem. Rev.* **105**, 739–758.
38. Tsujishita, Y., and Hurley, J. H. (2000) Structure and lipid transport mechanism of a StAR-related domain, *Nat. Struct. Biol.* **7**, 408–414.
39. Romanowski, M. J., Soccio, R. E., Breslow, J. L., and Burley, S. K. (2002) Crystal structure of the *Mus musculus* cholesterol-regulated START protein 4 (StarD4) containing a StAR-related lipid transfer domain, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6949–6954.
40. Thompson, T. B., Katayama, K., Watanabe, K., Hutchinson, C. R., and Rayment, I. (2004) Structural and functional analysis of tetracycline F2 cyclase from *Streptomyces glaucescens*. A type II polyketide cyclase, *J. Biol. Chem.* **279**, 37956–37963.
41. Homma, Y., and Emori, Y. (1995) A dual functional signal mediator showing RhoGAP and phospholipase C-delta stimulating activities, *EMBO J.* **14**, 286–291.
42. Chatterjee, M., Cramer, K. D., and Townsend, C. A. (1993) Kinetic nature of thiol activation in DNA cleavage by calicheamicin, *J. Am. Chem. Soc.* **115**, 3374–3375.
43. Kuduvali, P. N., Townsend, C. A., and Tullius, T. D. (1995) Cleavage by calicheamicin γ_1 of DNA in a nucleosome formed on the 5S RNA gene of *Xenopus borealis*, *Biochemistry* **34**, 3899–3906.
44. Drak, J., Iwasawa, N., Danishefsky, S., and Crothers, D. M. (1991) The carbohydrate domain of calicheamicin gamma I1 determines its sequence specificity for DNA cleavage, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7464–7468.
45. Ding, W. D., and Ellestad, G. A. (1991) Evidence for hydrophobic interaction between calicheamicin and DNA, *J. Am. Chem. Soc.* **113**, 6617–6620.
46. Zhao, Q., Frederick, R., Seder, K., Thao, S., Sreenath, H., Peterson, F., Volkman, B. F., Markley, J. L., and Fox, B. G. (2004) Production in two-liter beverage bottles of proteins for NMR structure determination labeled with either 15N- or 13C-15N, *J. Struct. Funct. Genomics* **5**, 87–93.
47. Palmer, A. G., III, Fairbrother, W. J., Cavanagh, J., Wright, P. E., and Rance, M. (1992) Improved resolution in three-dimensional constant-time triple resonance NMR spectroscopy of proteins, *J. Biomol. NMR* **2**, 103–108.
48. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* **6**, 277–293.
49. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods. Mol. Biol.* **278**, 313–352.
50. Bonvin, A. M., Boelens, R., and Kaptein, R. (2005) NMR analysis of protein interactions, *Curr. Opin. Chem. Biol.* **9**, 501–508.
51. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures, *J. Mol. Biol.* **231**, 1049–1067.
52. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics* **14**, 51–55.
53. Kumar, R. A., Ikemoto, N., and Patel, D. J. (1997) Solution structure of the calicheamicin gamma 1-DNA complex, *J. Mol. Biol.* **265**, 187–201.
54. Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr., D: Biol. Crystallogr.* **60**, 1355–1363.

A High-Throughput Screen for Compounds That Inhibit Aggregation of the Alzheimer's Peptide

Woojin Kim[†], Yunkyoung Kim[‡], Jaeki Min[‡], Dong Jin Kim[§], Young-Tae Chang^{‡,§,*}, and Michael H. Hecht^{†,*}

[†]Department of Chemistry, Princeton University, Princeton, New Jersey 08544, [‡]Department of Chemistry, New York University, New York, New York 10003, and [§]Biomedical Research Center, Korea Institute of Science and Technology, 39-1 Hawolgok-dong, Seongbuk-gu, Seoul 136-791, South Korea

Alzheimer's disease (AD) is estimated to affect nearly 10% of people over the age of 65 (www.alz.org/AboutAD/statistics.asp). As the "baby boom" generation ages and medical advances enable people to live longer, the number of people afflicted by AD is expected to increase dramatically. Given these trends, there is a tremendous need to develop therapeutics that block or reverse this debilitating neurodegenerative disease.

Although a number of drugs are in clinical use for the treatment of AD, most of these compounds target the symptoms of the disease, rather than its underlying molecular cause. Reducing the incidence of AD and slowing its progression will require new drugs that disrupt the underlying molecular etiology of AD. Although the molecular underpinnings of AD are not fully understood, a range of genetic and biochemical studies suggest that aggregation of the Alzheimer's peptide, A β , plays a causative role in the development of AD (1–7). Therefore, compounds that inhibit production and/or aggregation of A β are attractive candidates as therapeutics for the prevention and/or treatment of AD.

A β peptides are produced *in vivo* by proteolytic cleavage of the amyloid precursor protein (APP) by β and γ secretases (1). Because γ secretase can cleave at several alternative sites, the resulting A β peptides vary in length. The most abundant forms found in amyloid plaque are the 40-mer and the 42-mer. A β 40 is produced in greater abundance; however, A β 42 aggregates more readily and comprises the major component of amyloid plaque in diseased brains (8–10).

ABSTRACT Aggregation of the Alzheimer's peptide A β produces toxic multimeric species that play a key role in the development of Alzheimer's disease. Compounds that inhibit this aggregation may prove useful as therapeutic agents for the prevention or treatment of Alzheimer's disease. Although aggregation inhibitors may already exist in combinatorial libraries, finding these compounds in a cost-effective high-throughput manner poses an enormous challenge. To meet this challenge, we have developed a novel high-throughput screen capable of isolating inhibitors of A β aggregation from large libraries of inactive candidates. The screen uses a fusion of A β 42 to GFP. In the absence of inhibition, the rapid misfolding and aggregation of A β 42 causes the entire fusion protein to misfold, thereby preventing fluorescence. Compounds that inhibit A β 42 aggregation enable GFP to fold into its native structure and be identified by the resulting fluorescent signal. By implementing the screen on a pilot library of triazine derivatives, we have identified several putative inhibitors. One of the selected compounds was studied in detail by a series of biochemical and biophysical methods. These studies confirmed that the selected compound inhibits aggregation of synthetic A β 42 peptide. The fluorescence-based method described here is rapid and inexpensive and can be used to screen large libraries for inhibitors of A β 42 aggregation and/or amyloidogenesis.

*Corresponding authors,
hecht@princeton.edu,
yt.chang@nyu.edu.

Received for review March 27, 2006
and accepted July 2, 2006.

Published online August 4, 2006
10.1021/cb600135w CCC: \$33.50

© 2006 by American Chemical Society

Each step in the production and aggregation of A β can be considered as a target for intervention. These steps include (i) expression of the APP, (ii) proteolytic cleavage of APP into A β peptides, (iii) clearance of A β peptides from the system, and (iv) aggregation of A β into oligomers and insoluble amyloid. Numerous studies, both in academic labs and in the pharmaceutical industry, have targeted the first three steps of this pathway. Although initial results were promising (11–15), attempts to block expression, cleavage, or clearance have not produced an effective pharmaceutical.

In addition to these upstream processes, efforts have also focused on blocking the aggregation step itself. Although a few aggregation inhibitors have been reported (16–23), no clinically useful therapeutics have emerged. The discovery of compounds that block A β aggregation has been stymied by two major hindrances: First, structure-based rational drug design is precluded by the unavailability of a high-resolution structure. Although structural studies of A β have advanced significantly in recent years and models of A β amyloid have been built (24, 25), these structures are not of sufficiently high resolution to enable atom-by-atom drug design. Moreover, these structural models are for fibrils of A β , rather than the oligomers now assumed to be the toxic species. Second, large-scale screening of combinatorial libraries has been hindered by the unavailability of a cost-effective high-throughput screen for inhibitors of the early steps of A β aggregation.

Although several methods to screen for inhibitors of A β aggregation have been reported (26, 27), these methods are hampered by several shortcomings: Published methods typically require synthetic A β peptide. Because A β 42 aggregates, synthesis of this 42-residue peptide is laborious and time-consuming. Consequently, synthetic A β 42 is too expensive to use in screens aiming to analyze large libraries of compounds. In addition to its prohibitive cost, the aggregation of synthetic A β 42 can also interfere with the efficacy of a screen: Synthetic A β 42 often contains oligomeric “seeds” that can nucleate further aggregation. Since current models of AD pathogenesis implicate small oligomers on the pathway toward amyloid as the most toxic species (5, 6, 28–31), a screen relying on samples that contain pre-existing seeds may actually miss the most important inhibitors, including those that block the initial formation of soluble A β oligomers.

To overcome these challenges, we have developed a novel screen for inhibitors of A β 42 aggregation. Our screen does not require synthetic A β 42. Moreover, the new screen uses fluorescence for rapid and high-throughput detection. Here we describe the development of this screen, and its implementation to isolate inhibitors of A β 42 aggregation from a novel combinatorial library of triazine derivatives.

RESULTS AND DISCUSSION

A High-Throughput Screen for Inhibitors of A β

Aggregation. A high-throughput screen for inhibitors of A β aggregation requires the solubility/aggregation behavior of A β to be coupled to a property that can be assayed on many samples in parallel. Such coupling can be achieved by fusing the A β sequence to a reporter protein with an observable function that is blocked by A β aggregation but enabled by agents that inhibit A β aggregation.

Our screen achieves this goal by fusing the sequence of A β 42 to GFP. Because folding of GFP into its native fluorescent structure occurs slowly (32), the fluorescence of the A β 42–GFP fusion depends on the folding and solubility of the fused A β 42. Misfolding and aggregation of the A β 42 sequence cause the entire A β 42–GFP fusion to misfold prior to formation of the correct fluorescent structure. Inhibitors that retard (or block) A β 42 aggregation enable GFP to fold into its native structure and can be distinguished by the resulting fluorescent signal.

We chose to work with A β 42 rather than A β 40 because the longer peptide is the major component of senile plaque and the ratio of A β 42/A β 40 is increased in diseased brains (8, 9). The 42-residue peptide also forms aggregates more rapidly *in vitro* (10). In our fusion construct, A β 42 is separated from the N-terminus of GFP by a linker encoding the sequence GSAGSAAGSGEF (33). This sequence was shown previously to be effective in coupling the aggregation state of N-terminal fusions with the fluorescence of GFP (33–35). (Longer or more disordered linkers are not suitable because they would uncouple the properties of the N-terminal peptide from those of GFP.)

In previous work, we demonstrated that fusions of A β 42 to GFP do not fluoresce, and expression of the A β 42–GFP fusion in *Escherichia coli* yields nonfluorescent colonies (34, 35). We used these fusions as an artificial genetic system in *E. coli* to screen for mutations

in A β 42 that inhibit aggregation (34). Nonaggregating mutants were isolated by screening random mutations in A β 42 for those that produce green fluorescent colonies. The ability of the selected amino acid substitutions to diminish the aggregation of A β 42 was confirmed by biophysical studies of mutant versions of the synthetic 42-residue peptide.

Just as mutations in the sequence of A β 42 can retard aggregation by intramolecular effects (34), one might expect exogenous compounds to inhibit A β 42 aggregation by intermolecular effects. To test this hypothesis, we now demonstrate the use of the A β 42–GFP fusion as a high-throughput screen to search a library of small molecules for inhibitors of A β 42 aggregation.

In this initial test of the screen (Figure 1), we focus on a library of \sim 1000 compounds based on the triazine scaffold (36, 37). This library was prepared by varying the substituents at positions X, Y, and Z on the scaffold (Figure 1). The tested compounds were prepared previously for other assays (37) or freshly synthesized using a slightly modified method developed in our lab (36). The substituents at positions X, Y, and Z are described in our earlier work (36, 37).

To implement the screen, we added *E. coli* cells expressing the A β 42–GFP fusion protein to 96-well plates. Candidate molecules from the library of triazine derivatives were added to each well, and expression of the A β 42–GFP fusion protein was induced by addition of isopropyl- β -D-thiogalactopyranoside. Following 3 h of induction, the fluorescence of each well was measured on an automated plate reader. Several wells containing specific triazine derivatives fluoresced at levels significantly above background (Figure 2). Compounds were tested in quadruplicate, and the identification of “hits” was consistent across the four repetitions. The fluorescent hits are considered as putative inhibitors of A β 42 aggregation.

In Vitro Studies Confirm the Activity of a Selected Inhibitor. The ability of the A β 42–GFP fluorescence screen to identify potential inhibitors of aggregation is encouraging; however, because the screen relies on several artificial features, one might question whether compounds isolated by this screen actually inhibit aggregation of the A β 42 peptide in a well-defined biochemical system. The artificial features of the A β 42–GFP screen include (i) a fusion in which the relevant 42-residue A β sequence is only a small fraction of the 292-residue fusion protein and (ii) expression in *E. coli*,

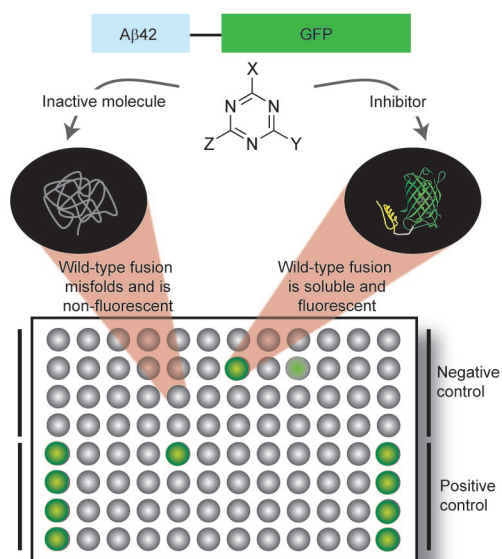


Figure 1. Fluorescence-based screen using the A β 42–GFP fusion. In the absence of inhibition, the A β 42 portion of the fusion aggregates rapidly and causes the entire A β 42–GFP fusion to misfold and aggregate (left). Therefore, no fluorescence is observed. However, inhibition of A β 42 aggregation enables GFP to form its native green fluorescent structure (right). (The green part of the ribbon diagram shows the structure of GFP; the yellow part is merely a schematic representation of a nonaggregated form of A β 42.) The triazine scaffold is shown at the center of the figure. Combinatorial diversity was introduced at sites marked X, Y, and Z. A 96-well plate is shown at the bottom of the figure. Compounds were added to each well, followed by *E. coli* cells expressing the A β 42–GFP fusion. Negative (colorless) and positive (green) controls are shown in the columns on the edges of the plate. For negative controls, no test compounds were added to the wells. For positive controls, the wild-type A β 42–GFP fusion was replaced with a fusion in which the A β 42 sequence contained mutations F19S and L34P. This double mutant was shown previously to inhibit aggregation and enable fluorescence of the A β 42–GFP fusion (34).

which is clearly not the natural system for AD. Consequently, it is essential to verify that fluorescence observed for the A β 42–GFP fusion expressed in *E. coli* indeed correlates with diminished aggregation of the A β 42 peptide.

In earlier work, we demonstrated that mutations in A β 42 that yield green fluorescence in the context of the A β 42–GFP fusion expressed in *E. coli* indeed diminish the aggregation of synthetic A β 42 peptide studied *in*

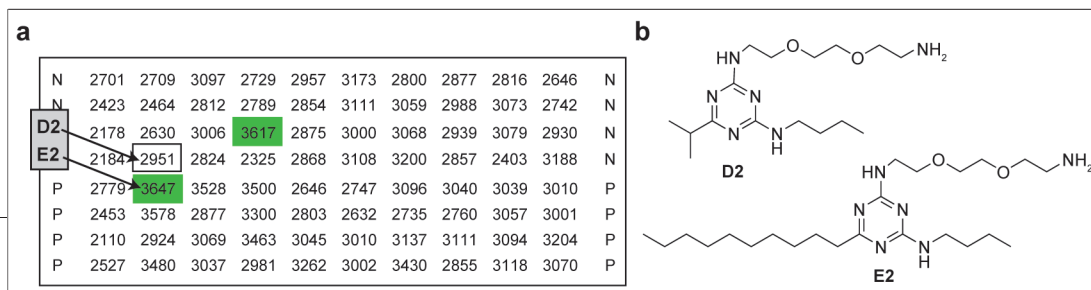


Figure 2. Screening results for the triazine library. a) Digital readout of the fluorescence of *E. coli* cells expressing the A β 42–GFP fusion in the presence of compounds from a combinatorial library of triazine derivatives. N denotes negative control wells without compound (2706 average, 238 standard deviation). P denotes positive control wells expressing a GFP fusion to the soluble F19S/L34P mutant of A β 42 (4610 average, 155 standard deviation). Compounds that reproducibly yielded fluorescence signals 3 standard deviations above the average of the negative control are highlighted in green. Compounds E2 (green) and D2 (control) were chosen for further studies. b) Structures of the aggregation inhibitor, E2, and the inactive control compound, D2. (The triazine scaffold of the combinatorial library is shown in Figure 1.)

vitro (34). To confirm that this correlation holds not only for mutations in A β 42 but also for the small molecule inhibitors isolated by our screen, we probed the effects of the selected triazine derivative E2 (Figure 2) and the unselected control D2 (Figure 2) on the aggregation behavior of synthetic A β 42 peptide.

Soluble monomeric A β 42 peptide can be prepared using organic solvents, sonication, and filtration (38). When such samples are diluted into aqueous buffer, the peptide aggregates into fibrillar amyloid structures, which can be assayed by the binding and resulting fluorescence of thioflavin T (39). The rate of A β 42 aggregation depends on the conditions of the incubation: Under “quiescent” conditions, aggregation is slow, whereas agitation causes A β 42 to aggregate more rapidly.

We studied the effects of compounds D2 and E2 on the aggregation of A β 42 under both quiescent and agitated conditions. For the quiescent conditions, synthetic A β 42 at a concentration of 20 μ M was incubated for 2 h in the presence of various concentrations of either D2 or E2. Fibril formation was assayed by the shifted fluorescence of thioflavin T that accompanies binding to fibrils (39). Compound E2 inhibits aggregation in a concentration-dependent manner, with an IC₅₀ of \sim 30 μ M. At 80 μ M, E2 produces nearly complete inhibition of A β 42 aggregation (Figure 3). In contrast, the control compound D2 shows no inhibitory effect.

Compounds D2 and E2 were also tested for their inhibitory effects under agitated incubation conditions. Here the effect was even more dramatic: While the control compound D2 was inactive, the selected compound E2 caused a 90% reduction in thioflavin T fluorescence at a concentration of only 50 μ M (Figure 4). The inhibitory effect of E2 was compared to dopamine and tannic acid, which were shown previously to inhibit A β 42 aggregation (40, 41). At concentrations of 25 and 50 μ M, the inhibitory effect of E2 was similar to, or slightly better than, dopamine or tannic acid (Figure 4).

Despite its inhibitory activity at 25 and 50 μ M, E2 at lower concentrations seems to cause a slight increase in amyloid formation (Figure 4). This surprising behavior is

reminiscent of the effect of trifluoroethanol (TFE), which inhibits fibrillogenesis at high concentrations but increases the rate of fibrillogenesis when added at low concentrations (42). It is not clear whether E2 acts by a mechanism similar to that exerted by TFE on peptide structure.

Finally, the ability of E2 to inhibit the assembly of A β 42 into amyloid fibrils was assessed by electron microscopy. A β 42 peptide was incubated for 5 d, either alone or in the presence of compounds D2 or E2. Five days is a relatively long incubation time; in the absence of inhibitors, A β 42 readily forms visible fibrils after 1 or 2 d (data not shown). Following the 5 d incubation, samples were stained with uranyl acetate and imaged by electron microscopy. The control compound D2 was inactive at all concentrations (Figure 5). Compound E2, however, inhibited fibrillogenesis in a dose-dependent manner. At 50 μ M, E2 had no effect, at 100 μ M only short fibrils (perhaps “protofibrils”) were observed, while at a concentration of 200 μ M, compound E2 completely inhibited fibril formation.

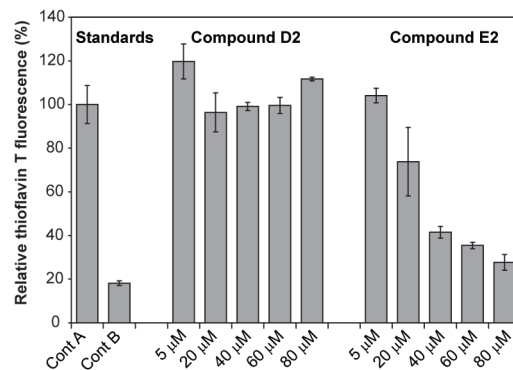


Figure 3. Aggregation of synthetic A β 42 peptide under quiescent conditions. Amyloid formation was assayed by binding and fluorescence of the diagnostic dye, thioflavin T. Control A is synthetic A β 42 alone. Control B is buffer alone. Compound E2 inhibits amyloid formation in a dose-dependent manner. The related compound, D2, has no activity. Additional controls showed that in the absence of peptide, compounds D2 and E2 had no effect on ThT fluorescence (data not shown). Fluorescence is shown as a percentage of the control (synthetic A β 42 alone).

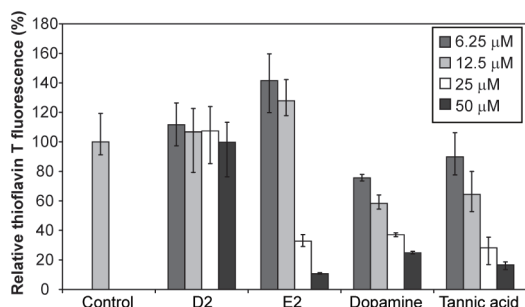


Figure 4. Amyloid formation assayed by thioflavin T fluorescence after incubation under agitated conditions. Compound E2 inhibits amyloid formation in a dose-dependent manner. The control compound, D2, has no activity. The right side of the figure shows results for dopamine and tannic acid, which were shown previously to inhibit A β 42 aggregation (40, 41). At concentrations of 25 and 50 μ M, the inhibitory effect of E2 was similar to, or slightly better than, dopamine or tannic acid. Fluorescence is shown as a percentage of the control (synthetic A β 42 alone).

These results confirm that the novel fluorescence-based assay using an A β 42–GFP fusion expressed in *E. coli* can detect compounds that indeed inhibit aggregation and/or amyloidogenesis of the A β 42 peptide.

Screening for Inhibitors of the Early Steps of Aggregation. An extensive range of genetic and biochemical studies (1–6) support the “amyloid cascade” hypothesis (43), which posits that accumulation of aggregated A β initiates a multistep cascade that ultimately leads to AD. While insoluble amyloid plaque has long been thought to play a causative role in AD, recent work suggests that smaller aggregates (A β oligomers) on the pathway toward amyloid may in fact be more toxic than insoluble plaque (5, 6, 28–31, 44, 45). Although the relative importance of the various stages of aggregation remains a topic of investigation, it is clear that aggregation of A β into some form of multimeric complex (ranging from small oligomers to large fibrils) produces toxic species that lead to AD.

Because the exact structure and oligomeric state of the toxic aggregate of A β 42 are not known, it is important to consider what stages of aggregation might be blocked by compounds scored as “hits” in high-throughput screens. In traditional screens relying on turbidity or binding of thioflavin T, a compound is scored as a hit if it prevents assembly into amyloid fibrils. Since fibrils occur late in the aggregation pathway, a potential

disadvantage of these older screens is the likelihood that some compounds isolated by these screens will inhibit the later steps of amyloidogenesis but fail to inhibit the upstream formation of toxic soluble oligomers. A more effective method would screen for compounds that block early misfolding and aggregation without requiring the formation of amyloid fibrils. Our A β 42–GFP screen for misfolding and aggregation may satisfy this requirement. Although we do not know the exact level of A β 42 aggregation (dimers? tetramers? hexamers?) that prevents fluorescence of the A β 42–GFP fusion, it seems likely that the nonfluorescent phenotype of the misfolded aggregate would be apparent at or before the dodecameric stage, which has been proposed to be the toxic species responsible for memory impairment in AD (31). Once active inhibitors are isolated, the exact oligomerization stage at which they function and the precise mechanism of their action can be assessed by biophysical studies.

For a screen to find inhibitors of the earliest stages of aggregation, it is important that the compounds being tested are present prior to the initial steps of the aggregation pathway. For screens that relied on synthetic A β 42 peptide, this posed a serious challenge: Because A β 42 aggregates so readily, it is difficult to prepare aqueous samples that are entirely free of partially aggregated seeds. The presence of these seeds (which presumably contain oligomers) meant that the species that must be inhibited would have already been present prior to addition of putative inhibitors. Consequently, screens using synthetic A β 42 peptide could miss the

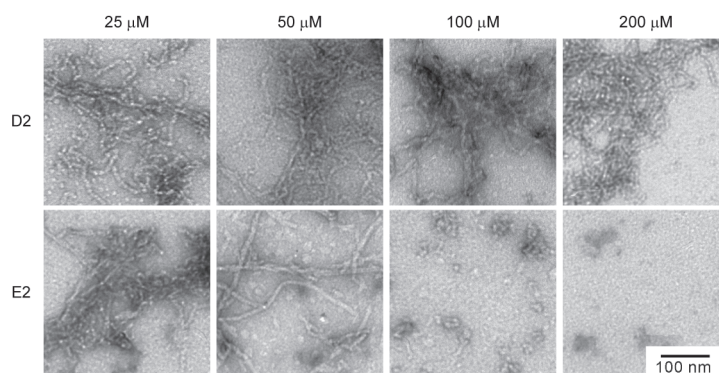


Figure 5. Electron microscopy of fibrils of A β 42 after incubation with D2 or E2. Synthetic A β 42 peptide was incubated for 5 d with various concentrations of either D2 or E2. At elevated concentrations, E2 inhibited fibrillogenesis. In contrast, the control compound D2, was inactive at all concentrations.

very compounds that ultimately will provide leads for the development of anti-AD therapeutics. The A β 42–GFP fusion system overcomes these problems: In the new screen, A β 42 is *not* present prior to addition of the test compounds; expression of the A β 42–GFP fusion is induced only *after* the test compounds have been added.

The triazine collection described above was used as a pilot library to demonstrate that the A β 42–GFP screen can indeed distinguish hits from inactive compounds. Although we are not suggesting that triazine is the optimal scaffold for drug discovery, we note that compound RS-0466, which was shown by Selkoe and co-workers (46) to block A β oligomerization and rescue long-term potentiation, is also a triazine derivative.

Sensitivity of the Screen. An effective screen must be sensitive enough to detect compounds with relatively low levels of inhibitory activity. This is important for two reasons: First, initial implementation of a screen typically searches for lead compounds, rather than final drugs. Therefore, a screen should be sensitive enough to detect first-generation compounds with only moderate effects on aggregation. (Such leads can be optimized at later stages.) Second, detection of compounds with low activity is important because drugs with modest effects on aggregation may in fact be sufficient to treat AD: In early onset AD caused by familial mutations in APP or in the presenilins, levels of A β 42 are increased by as little as 30% (3). This small increase in A β 42 can advance the onset of AD by 30–40 years. Therefore, compounds with only moderate inhibitory activity may suffice to delay the onset of AD to the point where it is no longer a major health problem. The A β 42–GFP fusion system described here has the required level of sensitivity. This was demonstrated explicitly by earlier work using the fusion to screen for mutations in A β 42 that diminish aggregation (35). In that work, we showed that mutations that alter the aggregation rate only moderately are readily detected by changes in the fluorescence of the A β 42–GFP fusion. Thus, the A β 42–GFP fusion system is well-suited for the detection of compounds having a range of inhibitory activity.

The A β 42–GFP fusion system is sensitive to inhibitory effects at sites throughout the length of 42-residue A β sequence (34). One might be concerned that the presence of a linker following residue 42 would interfere with inhibitory effects on the C-terminal residues of A β 42, which are known to be important for aggregation (8–10,

47). Our earlier studies, however, demonstrated that the A β 42–GFP fusion can discriminate small differences in aggregation rates caused by mutations throughout A β 42, including those at residues 41 and 42 (34, 35).

When screening for compounds that inhibit aggregation, it is important to ensure that the screen does not inadvertently identify generic inhibitors of protein folding. This possibility must be considered because aggregation into β -sheet fibrils and folding into native globular structures are similar processes: Both involve self-assembly of a polypeptide into an ordered structure. Although A β 42 aggregation is intermolecular and protein folding is intramolecular, the two processes are governed by the same types of interactions (hydrogen bonding, the hydrophobic effect, propensities for secondary structure, side chain packing, *etc.*). Therefore, it is important to establish that a screen for inhibitors of aggregation does not inadvertently identify inhibitors of protein folding, particularly the folding of β -sheet proteins. The A β 42–GFP fusion system is internally controlled for this possibility. A positive signal (fluorescence) is required to identify hits, and this signal is observed if and only if GFP folds into its native structure. Therefore, generic inhibitors of protein folding will not be isolated by this screen. Moreover, since GFP is a β -sheet protein, generic inhibitors of β -sheet structure will not be isolated. These undesirable effects are “weeded out” by the requirement for correct folding of GFP.

Applications of the Screen. Because the A β 42–GFP fusion is sensitive enough to detect both low and high levels of inhibitory activity, the screen can be used to determine structure/activity relationships. For example, compounds D2 and E2 are identical at positions X and Y but differ at position Z (Figure 2). By screening collections of molecules that differ at only one position, we have begun to establish the relationship between the substituents at each position and the resulting level of inhibition (Kim *et al.*, unpublished).

AD is one of several diseases involving protein misfolding and aggregation. Others include prion encephalopathies and Huntington’s disease (48, 49). The GFP fusion system described here may also be suitable to screen for inhibitors of the aggregation processes that underlie these other diseases.

Certain classes of compounds that successfully inhibit A β 42 aggregation may nonetheless be missed by our screen. Two examples include compounds that fluoresce at the same wavelength as GFP and compounds

that are toxic to cells. To enable screening of libraries containing green fluorescent compounds, it may be necessary to modify our system for future use with variants of GFP that fluoresce in other parts of the spectrum (e.g., yellow fluorescent protein (50, 51). Cytotoxic compounds will also be missed by our screen; however, this may be advantageous since such compounds are unlikely to be suitable as drugs.

Finally, we note that the initial version of the A β 42–GFP fluorescent screen described in this work relies on the fusion protein expressed in *E. coli*. Screening in *E. coli* has several advantages: It is fast, inexpensive, and highly reproducible. Moreover, it favors compounds that (i) are nontoxic and (ii) readily penetrate biological barriers. Nonetheless, expression in *E. coli* may also introduce a limitation: To be scored as a hit in this cell-

based screen, a compound must enter the bacterial cell. Inhibitors of A β aggregation that fail to enter cells will not produce fluorescent signals and will escape detection in this initial version of our screen. The significance of this limitation will depend on the type of library being screened. Some chemical moieties are inherently more likely than others to enter cells (52). To effectively screen compounds that do not enter bacterial cells, we recently developed a modified version of the A β 42–GFP fluorescent screen in which the fusion protein is expressed *in vitro* using a cell-free transcription and translation system (Kim, Wurth, and Hecht, unpublished). This cell-free system readily distinguished between aggregating and nonaggregating mutants of A β 42. Future work will adapt this cell-free system to screen for small-molecule inhibitors of A β 42 aggregation.

METHODS

Fluorescent Screen for Inhibitors of A β Aggregation. The vector for expressing the A β 42–GFP fusion was described previously (34, 35). Strain BL21(DE3) *E. coli* cells (53) harboring the A β 42–GFP fusion vector were grown in LB media supplemented with 35 $\mu\text{g mL}^{-1}$ kanamycin. When cultures reached an OD₆₀₀ = 0.8, 100 μL of culture was transferred to the wells of 96-well plates. Candidate compounds from the triazine library were added to each well, and protein expression was induced by adding isopropyl- β -D-thiogalactopyranoside to a final concentration of 1 mM. Samples were incubated with gentle agitation at 37 °C. Following 3 h of incubation, the fluorescence of each well was measured at 512 nm (excitation 490 nm) using an automated plate reader (SpectraMAX Gemini XS, Molecular Devices). To verify that cell densities were consistent across all samples, the OD₆₀₀ was also measured. Compounds were tested in quadruplicate: twice at a final concentration of 30 μM and twice at a final concentration of 100 μM . The identification of hits was consistent across several repetitions. Occasionally, we observed compounds that yielded fluorescent signals below that of the negative control; these compounds may be cytotoxic.

Overall, screening a library of ~1000 compounds required several hours. Scale-up procedures using robotic sample handling will enable screening of much larger libraries on a high-throughput scale.

Synthesis of the Library of Triazine Derivatives. The general design and synthesis of a triazine-based library was reported previously (36, 37). In this solid-phase synthetic approach, three types of building blocks were prepared separately and assembled by chemically orthogonal reactions. The putative inhibitor E2 and the structurally related but inactive compound D2 were resynthesized for further characterization. Data for E2: ¹H NMR (400 MHz, MeOH-*d*₄) δ 3.72–3.63 (8H, m), 3.59 (2H, m), 3.37 (2H, m), 3.12 (2H, t, *J* = 5.0 Hz), 2.46–2.33 (2H, m), 1.67 (2H, m), 1.56 (2H, m), 1.43–1.24 (16H, m), 0.95 (3H, t, *J* = 7.5 Hz), 0.89 (3H, t, *J* = 7.0 Hz). LC–MS (*m/z*): Calculated for C₂₃H₄₆N₆O₂: 438.4. Found: 439.4 [M + H]⁺. Data for D2: ¹H NMR (MeOH-*d*₄) δ 3.71–3.64 (8H, m), 3.60 (2H, m), 3.38 (2H, m), 3.12 (2H, t, *J* = 5.0 Hz), 2.65 (1H, m), 1.57 (2H, m), 1.39 (2H, m), 1.21 (6H, d, *J* = 6.8 Hz), 0.95 (3H, t, *J* = 7.3 Hz); LC–MS (*m/z*): Calculated for C₁₆H₃₂N₆O₂: 340.3. Found: 341.3 [M + H]⁺.

Preparation of Synthetic A β 42 Peptide. A β 42 peptide (unpurified) was purchased from the Keck Institute at Yale University and

purified on a C4 reverse phase column (Vydac). After purification, the peptide was snap-frozen at –75 °C and lyophilized. Monomeric samples were prepared by 15 min of sonication after addition of trifluoroacetic acid (TFA, 1 mg mL^{–1} peptide concentration) (54). Residual TFA was removed by addition of hexafluoroisopropanol (Sigma Aldrich) and argon blow.

Thioflavin T Assays. Synthetic A β 42 peptide was incubated at 20 μM in phosphate buffered saline (PBS, 50 mM NaH₂PO₄, 100 mM NaCl, 0.02% NaN₃) in the presence or absence of candidate inhibitors at various concentrations. Following incubation with or without agitation, thioflavin T was added to a final concentration of 7 μM , and fluorescence was measured at 490 nm (excitation 450 nm).

Electron Microscopy. A β 42 peptide at a concentration of 20 μM in PBS buffer was incubated in the presence or absence of the test compounds at various concentrations. Following 5 d of incubation at 37 °C under quiescent conditions, Formvar carbon-coated grids were floated on a drop of the sample for 2 min. The grids were blotted using filter paper and then stained for 2 min with freshly made 1% uranyl acetate. Samples were imaged using a Zeiss 912ab electron microscope.

Acknowledgments: We thank M. Bishop for assistance with the electron microscopy, J. Crawford for peptide synthesis, and C. Wurth for initial experiments using the A β 42–GFP fusion.

REFERENCES

1. Selkoe, D. J. (2001) Alzheimer's disease: genes, proteins, and therapy, *Physiol. Rev.* 81, 741–766.
2. Koo, E., Lansbury, P. J., and Kelly, J. (1999) Amyloid diseases: abnormal protein aggregation in neurodegeneration, *Proc. Natl. Acad. Sci. U.S.A.* 96, 9989–9990.
3. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T. D., Hardy, J., Hutton, M., Kukull, W., Larson, E., Levy-Lahad, E., Viitanen, M., Peskind, E., Poorkaj, P., Schellenberg, G., Tanzi, R., Wasco, W., Lannfelt, L., Selkoe, D., and Younkin, S. (1996) Secreted amyloid- β protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease, *Nat. Med.* 2, 864–870.

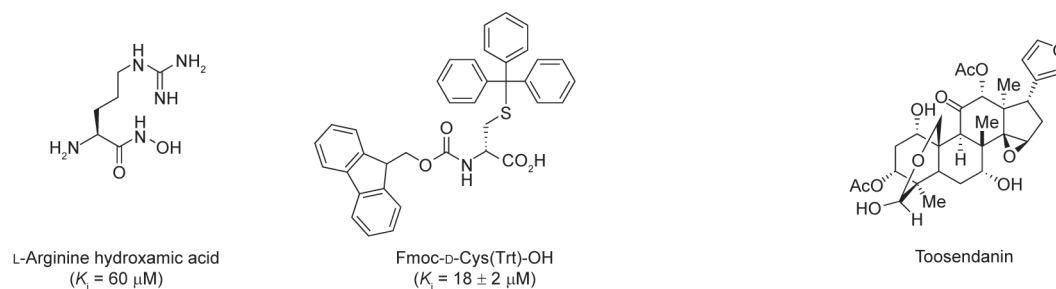
4. Geula, C., Wu, C., Saroff, D., Lorenzo, A., Yuan, M., and Yankner, B. (1998) Aging renders the brain vulnerable to amyloid- β protein neurotoxicity, *Nat. Med.* **4**, 827–831.
5. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002) Naturally secreted oligomers of amyloid- β protein potently inhibit hippocampal long-term potentiation *in vivo*, *Nature* **416**, 535–539.
6. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Rowan, M. J., and Selkoe, D. (2002) Amyloid- β oligomers: their production, toxicity and therapeutic inhibition, *Biochem. Soc. Trans.* **30**, 552–557.
7. Irie, K., Murakami, K., Masuda, Y., Morimoto, A., Ohgashi, H., Ohashi, R., Takegoshi, K., Nagao, M., Shimizu, T., and Shirasawa, T. (2005) Structure of β -amyloid fibrils and its relevance to their neurotoxicity: implications for the pathogenesis of Alzheimer's disease, *J. Biosci. Bioeng.* **99**, 437–447.
8. Gravina, S. A., Ho, L., Eckman, C. B., Long, K. E., Otvos, L., Jr., Younkin, L. H., Suzuki, N., and Younkin, S. G. (1995) Amyloid- β protein (A β) in Alzheimer's disease brain, *J. Biol. Chem.* **270**, 7013–7016.
9. Roher, A. E., Lowenson, J. D., Clarke, S., Woods, A. S., Cotter, R. J., Gowing, E., and Ball, M. J. (1993) β -Amyloid-(1-42) is a major component of cerebrovascular amyloid deposits: implications for the pathology of Alzheimer disease, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10836–10840.
10. Jarrett, J., Berger, E., and Lansbury, P., Jr. (1993) The carboxy terminus of the β -amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease, *Biochemistry* **32**, 4693–4697.
11. Dash, P., Moore, A., and Orsi, S. (2005) Blockade of γ -secretase activity within the hippocampus enhances long term memory, *Biochem. Biophys. Res. Commun.* **338**, 777–782.
12. Asai, M., Hattori, C., Iwata, N., Saido, T. C., Sasagawa, N., Hashimoto, Y., Maruyama, K., Tanuma, S., Kiso, Y., and Ishiura, S. (2006) The novel β -secretase inhibitor KMI-429 reduces amyloid- β peptide production in amyloid precursor protein transgenic and wild type mice, *J. Neurochem.* **96**, 533–540.
13. Bacskai, B. J., Kajdasz, S. T., Christie, R. H., Carter, C., Games, D., Seubert, P., Schenk, D., and Hyman, B. T. (2001) Imaging of amyloid- β deposits in brains of living mice permits direct observation of clearance of plaques with immunotherapy, *Nat. Med.* **7**, 369–372.
14. Schenk, D., Barbour, R., Dunn, W., Gordon, G., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Liao, Z., Lieberburg, I., Motter, R., Mutter, L., Soriano, F., Shopp, G., Vasquez, N., Vandevent, C., Walker, S., Wogulis, M., Yednock, T., Games, D., and Seubert, P. (1999) Immunization with amyloid- β attenuates Alzheimer-disease-like pathology in the PDAPP mouse, *Nature* **400**, 173–177.
15. Bard, F., Cannon, C., Barbour, R., Burke, R., Games, D., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Lieberburg, I., Motter, R., Nguyen, M., Soriano, F., Vasquez, N., Weiss, K., Welch, B., Seubert, P., Schenk, D., and Yednock, T. (2000) Peripherally administered antibodies against amyloid- β peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer's disease, *Nat. Med.* **6**, 916–919.
16. Ritchie, C. W., Bush, A. I., Mackinnon, A., Macfarlane, S., Mastwyk, M., Macgregor, L., Kiers, L., Cherny, R., Li, Q., Tammer, A., Carrington, D., Mavros, C., Volitakis, I., Xilinas, M., Ames, D., Davis, S., Beyreuther, K., Tanzi, R. E., and Masters, C. L. (2003) Metal-protein attenuation with iodochloroquinol (cloquinol) targeting A β amyloid deposition and toxicity in Alzheimer disease, *Arch. Neurol.* **60**, 1685–1691.
17. Levine, H., III (2002) The challenge of inhibiting A β polymerization, *Curr. Med. Chem.* **9**, 1121–1133.
18. Wood, S. J., MacKenzie, L., Maleeff, B., Hurle, M. R., and Wetzel, R. (1996) Selective inhibition of A β fibril formation, *J. Biol. Chem.* **271**, 4086–4092.
19. Lashuel, H. A., Hartley, D. M., Balakhaneh, D., Aggarwal, A., Teichberg, S., and Callaway, D. J. (2002) New class of inhibitors of amyloid- β fibril formation, *J. Biol. Chem.* **277**, 42881–42890.
20. Talaga, P. (2001) β -Amyloid aggregation inhibitors for the treatment of Alzheimer's disease: dream or reality? *Mini-Rev. Med. Chem.* **1**, 175–186.
21. Findeis, M. A. (2000) Approaches to discovery and characterization of inhibitors of amyloid β -peptide polymerization, *Biochim. Biophys. Acta* **1502**, 76–84.
22. Cohen, T., Frydman-Marom, A., Rechter, M., and Gazit, E. (2006) Inhibition of amyloid fibril formation and cytotoxicity by hydroxyindole derivatives, *Biochemistry* **45**, 4727–4735.
23. Porat, Y., Abramowitz, A., and Gazit, E. (2006) Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism, *Chem. Biol. Drug. Des.* **67**, 27–37.
24. Lührs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Döbeli, H., Schubert, D., and Riek, R. (2005) 3D structure of Alzheimer's amyloid- β (1–42) fibrils, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17342–17347.
25. Petkova, A., Yau, W., and Tycko, R. (2006) Experimental constraints on quaternary structure in Alzheimer's β -amyloid fibrils, *Biochemistry* **45**, 498–512.
26. Esler, W. P., Stimson, E. R., Ghilardi, J. R., Felix, A. M., Lu, Y., Vinters, H., Mantyh, P., and Maggio, J. (1997) A β deposition inhibitor screen using synthetic amyloid, *Nat. Biotechnol.* **15**, 258–263.
27. Blanchard, B. J., Chen, A., Rozeboom, L. M., Stafford, K. A., Weigele, P., and Ingram, V. M. (2004) Efficient reversal of Alzheimer's disease fibril formation and elimination of neurotoxicity by a small molecule, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14326–14332.
28. Kaye, R., Head, E., Thompson, J., McIntire, T., Milton, S., Cotman, C., and Glabe, C. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis, *Science* **300**, 486–489.
29. Lambert, M., Barlow, A., Chromy, B., Edwards, C., Freed, R., Morgan, T., Rozovsky, I., Trommer, B., Viola, K., Wals, P., Zhang, C., Finch, C., Krafft, G. A., and Klein, W. L. (1998) Diffusible, nonfibrillar ligands derived from A β 1–42 are potent central nervous system neurotoxins, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6448–6453.
30. Whalen, B., Selkoe, D. J., and Hartley, D. (2005) Small non-fibrillar assemblies of amyloid- β protein bearing the Arctic mutation induce rapid neuritic degeneration, *Neurobiol. Dis.* **20**, 254–266.
31. Lesne, S., Koh, M. T., Kotilinek, L., Kaye, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006) A specific amyloid- β protein assembly in the brain impairs memory, *Nature* **440**, 352–357.
32. Cubitt, A., Heim, R., Adams, S., Boyd, A., Gross, L., and Tsien, R. (1995) Understanding, improving and using green fluorescent proteins, *Trends Biochem. Sci.* **20**, 448–455.
33. Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein, *Nat. Biotechnol.* **17**, 691–695.
34. Wurth, C., Guimard, N. K., and Hecht, M. H. (2002) Mutations that reduce aggregation of the Alzheimer's A β 42 peptide: an unbiased search for the sequence determinants of A β amyloidogenesis, *J. Mol. Biol.* **319**, 1279–1290.
35. Kim, W., and Hecht, M. H. (2005) Sequence determinants of enhanced amyloidogenicity of Alzheimer A β 42 peptide relative to A β 40, *J. Biol. Chem.* **280**, 35069–35076.
36. Bork, J. T., Lee, J., and Chang, Y. (2003) Palladium-catalyzed cross-coupling reaction of resin bound chlorotriazines, *Tetrahedron Lett.* **44**, 6141–6144.

37. Khersonsky, S. M., Jung, D., Kang, T., Walsh, D. P., Moon, H., Jo, H., Jacobson, E., Shetty, V., Neubert, T. A., and Chang, T. (2003) Facilitated forward chemical genetics using a tagged triazine library and zebrafish embryo screening, *J. Am. Chem. Soc.* **125**, 11804–11805.
38. Bitan, G., and Teplow, D. (2005) Preparation of aggregate-free, low molecular weight amyloid- β for assembly and toxicity assays, in *Methods in Molecular Biology* (Sigurdsson, E. M., Ed.) pp 3–9, Humana Press, Totowa, NJ.
39. Levine, H., III (1993) Thioflavin T interaction with synthetic Alzheimer's disease β -amyloid peptides: detection of amyloid aggregation in solution, *Protein Sci.* **2**, 404–410.
40. Li, J., Zhu, M., Manning-Bog, A., Di Monte, D., and Fink, A. (2004) Dopamine and L-dopa disaggregate amyloid fibrils: implications for Parkinson's and Alzheimer's disease, *FASEB J.* **18**, 962–964.
41. Ono, K., Hasegawa, K., Naiki, H., and Yamada, M. (2004) Anti-amyloidogenic activity of tannic acid and its activity to destabilize Alzheimer's β -amyloid fibrils *in vitro*, *Biochim. Biophys. Acta.* **1690**, 193–202.
42. Fezoui, Y., and Teplow, D. B. (2002) Kinetic Studies of amyloid- β protein fibril assembly, *J. Biol. Chem.* **277**, 36948–36954.
43. Hardy, J. A., and Higgins, G. A. (1992) Alzheimer's disease: the amyloid cascade hypothesis, *Science* **256**, 184–185.
44. Lee, E., Leng, L., Zhang, B., Kwong, L., Trojanowski, J., Abel, T., and Lee, V. (2006) Targeting amyloid- β peptide (A β) oligomers by passive immunization with a conformation-selective monoclonal antibody improves learning and memory in A β precursor protein (APP) transgenic mice, *J. Biol. Chem.* **281**, 4292–4299.
45. Kokubo, H., Kaye, R., Glabe, C., and Yamaguchi, H. (2005) Soluble A β oligomer ultrastructurally localize to cell process and might be related to synaptic dysfunction in Alzheimer's disease brain, *Brain Res.* **1031**, 222–228.
46. Walsh, D. M., Townsend, M., Podlisny, M. B., Shankar, G. M., Fadeeva, J. V., Agnaf, O. E., Hartley, D. M., and Selkoe, D. J. (2005) Certain inhibitors of synthetic amyloid β -peptide (A β) fibrillogenesis block oligomerization of natural A β and thereby rescue long-term potentiation, *J. Neurosci.* **25**, 2455–2462.
47. Harper, J., and Lansbury, P., Jr. (1997) Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins, *Annu. Rev. Biochem.* **66**, 385–407.
48. Carrell, R., and Lomas, D. (1997) Conformational disease, *Lancet* **350**, 134–138.
49. Thomas, P., Qu, B., and Pederson, P. (1995) Defective protein folding as a basis of human disease, *Trends Biochem. Sci.* **20**, 456–459.
50. Tsien, R. Y. (1998) The green fluorescent protein, *Annu. Rev. Biochem.* **67**, 509–544.
51. Lippincott-Schwartz, J., and Patterson, G. H. (2003) Development and use of fluorescent protein markers in living cells, *Science* **300**, 87–91.
52. Pardridge, W. M. (1998) CNS drug design based on principles of blood-brain barrier transport, *J. Neurochem.* **70**, 1781–1792.
53. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, in *Methods in Enzymology* (Goeddel, D., Ed.) pp 60–89, Academic Press, Orlando, FL.
54. Jao, S., Ma, K., Talafous, J., Orlando, R., and Zagorski, M. G. (1997) Trifluoroacetic acid pretreatment reproducibly disaggregates the amyloid- β peptide, *Amyloid: Int. J. Exp. Clin. Invest.* **4**, 240–252.

ACS Chem. Biol., 2006, 1, 359–369**Correction: The Use of Small Molecules to Investigate Molecular Mechanisms and Therapeutic Targets for Treatment of Botulinum Neurotoxin A Intoxication**

Tobin J. Dickerson* and Kim D. Janda*

Incorrect structures were inadvertently introduced for L-arginine hydroxamic acid and Fmoc-D-Cys(Trt)-OH (Figure 3), and limonoid toosendanin (Figure 4) in this review article. The corrected structures are reproduced below. We regret this error and apologize for any inconvenience this may have caused.

**Figure 3. BoNT/A LC metalloprotease inhibitors.****Figure 4. Structure of the limonoid toosendanin.**

10.1021/cb600333c
Published online August 18, 2006

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

The Author's Choice

Those of you who have published with ACS know about *ACS Articles on Request*. This policy currently provides (free of charge) to all contributing authors a unique URL directed to the ACS web site that authors may e-mail to their colleagues or post on external web sites. By this link, authors may provide interested readers with up to 50 article downloads within the first year of publication; thereafter, these same author-directed URL links provide unlimited access to final published articles from the ACS web site. This open access policy provides a mechanism for authors to distribute their work more broadly. To complement this linking approach, ACS is pleased to offer a new publishing option, the *ACS AuthorChoice*, for authors who wish or need to make their published research articles available without restriction immediately upon publication.

The *ACS AuthorChoice* option (www.pubs.acs.org/pressrelease/author_choice) provides a mechanism for individual authors or their research funding agencies to sponsor immediate access to their article upon online publication on the ACS journals' web sites. This new policy will be implemented later in the fall and will enable paying authors to post electronic copies of published articles on their personal web sites. In addition, authors can also place electronic copies of their paper in institutional repositories for noncommercial scholarly purposes immediately upon publication. ACS retains copyright to the article.

ACS AuthorChoice is a fee-based option for contributing authors. Authors who are ACS members and/or affiliated with an ACS subscribing institution receive significant discounts as outlined below:

\$3,000: non-ACS member unaffiliated with a subscribing institution

\$2,000: non-ACS member affiliated with a subscribing institution

\$1,500: ACS member unaffiliated with a subscribing institution

\$1,000: ACS member affiliated with a subscribing institution

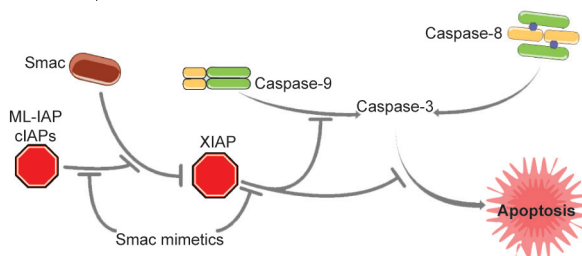
To ensure complete separation between the editorial decision process and economic consideration, the author will be able to select the *ACS AuthorChoice* option only after the paper has been accepted for publication. Once the author has chosen this publishing option, ACS will make the article freely available as soon as it is published on the journal's web site.

ACS is committed to experimenting with new publishing models to broaden access of our highly valued peer-reviewed research. We will monitor the usage of the *ACS AuthorChoice* option and collect web statistics on manuscripts made immediately available on the ACS web sites. These data inform our future decisions on access options.

The *ACS Chemical Biology* Staff

“Smac”ked to Death

Deregulation of apoptosis, the highly controlled process of cell suicide, can result in a variety of disease states, including cancer. Many cancer cells manage to elude the activity of anticancer agents by becoming resistant to apoptotic pathways. Specialized proteins termed inhibitors of apoptosis (IAPs) act generally to prevent cell death, and many IAP proteins have been implicated as accomplices in cancer cell resistance to apoptosis. The protein second mitochondrial activator of caspases (Smac) is an antagonist of IAPs that promotes apoptosis, making Smac-like molecules potential anticancer agents. Zobel *et al.* (p 525) report the design, synthe-



sis, and biological activity of small molecules designed to mimic the four N-terminal residues of Smac that interact with IAPs.

Molecular modeling aided the design and subsequent synthesis of [7,5]-bicyclic peptidomimetics that possessed submicromolar K_i values against IAP proteins, as determined by an *in vitro* fluorescence polarization assay. Notably, all of the key contacts observed in the structure of an IAP derivative with a Smac-derived peptide substrate were retained in the crystal structure of one of the peptidomimetics with the protein. Biological activity studies showed that the inhibitors prevented the interaction of an IAP protein and Smac in cell lysates. In addition, the proapoptotic properties of the peptidomimetics were evident when the compounds triggered apoptosis in two cancer cell lines and synergized with chemotherapeutic agents to promote cancer cell death.

Published online September 15, 2006

10.1021/cb600369w CCC: \$33.50

© 2006 by American Chemical Society

Sticking It to NMR

Natural products are a rich source of biologically active compounds and potential therapeutics. Unfortunately, natural product isolation and characterization are notoriously laborious because of the small quantities of material available and the limitations of analytical techniques. Recently, new methods utilizing microcoil and cryogenic technologies have enabled NMR analysis of samples of significantly reduced quantities. Now, Dossey *et al.* (p 511) cleverly combine the use of small-diameter samples and cryogenics to analyze the secreted material from individual walking stick insects by NMR.

An NMR probe made from high-temperature superconducting material was used to conduct an analysis of secretions from two species of walking sticks, *Anisomorpha buprestoides* and *Peruphasma schultei*. Aqueous and organic soluble material examined from individual *A. buprestoides* insects revealed that the secretions contained approximately equal amounts of



glucose, which was not previously known to be a component of walking stick secretions, and a stereoisomeric mixture of the monoterpene dialdehyde dolichodial. Surprisingly, it was discovered that the ratio of dolichodial-like stereoisomers present in the secretions varied among different *A. buprestoides* insects and within individual insects over time. In addition, examination of secretions from *P. schultei* led to the identification of a new dolichodial stereoisomer that the authors named peruphasmal. The ability to explore chemical biodiversity at the individual animal level provides intriguing information not previously accessible in a practical manner.

Adapting Aptamers

The explosion in genomics and proteomics has provided an extensive tool kit for dissecting the roles of genes and proteins in cellular function. Methods for examining the contributions of small molecules in regulating complex cellular processes, however, are less well established. Niles and Marletta (p 515) describe a method for using RNA aptamers to probe the role of the small-molecule heme in regulating heme biosynthesis in *Escherichia coli*.

RNA aptamers can be expressed intracellularly and can be evolved to bind essentially to any small molecule, and this makes them attractive tools for exploring small-molecule function in a cellular context. To this end, the authors generated several RNA aptamers specific for heme and used them to examine the effects of modulating heme levels in the *E. coli* heme auxotroph. Heme-limiting conditions in the heme auxotroph cause a growth defect and lead to increased levels of δ -aminolevulinic acid (δ -ALA), an intermediate in heme biosynthesis. The authors found that expression of heme-specific aptamers impaired growth and resulted in accumulation of δ -ALA, both of which could be reversed upon addition of heme. The ability of the aptamers to modulate the heme biosynthetic pathway in a predictable fashion validates this innovative use of RNA aptamer technology, and the generality of this method should facilitate exploration of the role that small molecules play in regulating biological processes.



Introducing our AUTHORS

ACS
chemical
biology



Kerry Zobel

Current position: Genentech, Inc., Department of Protein Engineering and Medicinal Chemistry, research associate with Dr. Kurt Deshayes

Education: University of Wisconsin-Madison, B.S. in chemistry, 1999

Nonscientific interests: Bicycling the streets of San Francisco, camping and backpacking in the redwood forests, and guitar playing

As an organic chemist, a major interest of mine has been designing and creating new organic compounds. My work at Genentech allows me to incorporate this passion into the study of biology that is crucial for understanding and developing therapeutics. For example, the regulation of apoptosis, which is the current focus of our lab. To me this paper is an exciting example of progressing from a protein target to specific tight binding small molecules through the use of rationally designed peptidomimetics. We demonstrate this through a class of [7,5] bicyclic lactame compounds that bind to the baculoviral inhibitor of apoptosis protein (IAP) repeat (BIR) and BIR3 domains of melanoma IAP and X-chromosome-linked IAP. (Read Zobel's article on p 525.)

This project was my first chance to incorporate my passion for studying insects with my formal training in biochemistry. One aspect that I found most fascinating was that we were able to analyze venom from a single insect and discover unreported components of that substance. Such a discovery opens new doors to understanding arthropod chemistry. Indeed, only a tiny fraction of the total chemical biodiversity that exists in insects alone has been determined. I hope to continue exploring the large potential for discovery that exists in these creatures. Using cutting-edge technologies such as the microsample NMR used in our study of phasmid insect venom, we can now begin to push the frontiers of natural products chemistry. I am currently looking for future work involving medicinal and natural product discovery from invertebrates. (Read Dossey's article on p 511.)

Current position: University of Florida, College of Medicine, Department of Biochemistry and Molecular Biology, postdoctoral research associate with Prof. Arthur Edison

Education: Oklahoma State University, B.S. in biochemistry and molecular biology, *cum laude*, 2001; University of Florida, Gainesville, Ph.D. in biochemistry and molecular biology, with Prof. Art Edison, 2006

Nonscientific interests: Entomology, keeping and breeding invertebrates, comedy, playing trumpet, nature photography, travel, gardening, camping, fishing, and hiking through the wilderness



Aaron Dossey



Jacquin Niles

Current position: University of California, Berkeley, Department of Chemistry, postdoctoral research associate with Prof. Michael Marletta

Education: Massachusetts Institute of Technology, S.B. in chemistry, 1994; Massachusetts Institute of Technology, Ph.D. in toxicology with Dr. Steven Tannenbaum, 2001; Harvard Medical School, M.D., 2002

Nonscientific interests: Most sports—definitely cricket!

One of my interests is the development of versatile and easily accessible tools for addressing fundamental questions in biology. Aptamer technology has many attributes compatible with this goal, especially because diverse cellular processes can be targeted with exquisite specificity. In this paper, we have used heme-binding aptamers to specifically target the well-studied heme biosynthetic pathway in the model organism *Escherichia coli*, demonstrating the applicability of this technology in modulating a small-molecule-regulated pathway. In the long run, I am interested in extending these approaches to nonmodel organisms that are less amenable to traditional genetic methods for studying essential cellular processes. (Read Niles's article on p 515.)

Published online September 15, 2006 • 10.1021/cb600387e CCC: \$33.50

© 2006 by American Chemical Society

www.acschemicalbiology.org

VOL.1 NO.8 • ACS CHEMICAL BIOLOGY

473

Spotlight

Stem Cells Go Global

The clinical promise of human embryonic stem cell research (hESCR) is tantalizing scientists throughout the world, but controversial ethical issues continually jeopardize progress in the field. Countries around the globe have distinct regulations guarding hESCR, but borders can become blurred when scientists from different countries attempt to work together. More than 50 scientists,



Image courtesy of Getty Images

ethicists, journal editors, lawyers, and policy makers from 14 countries recently convened in Hinxton, Cambridge, U.K., to create guiding principles for international collaborations in hESCR (Mathews *et al.*, *Science* 2006, 313, 921–922).

One hotly contested topic centered on extraterritorial jurisdiction over hESCR. Should countries that have banned hESCR have the power to prohibit their scientists from participating in hESCR collaborations in countries in which it is legal? Currently, at least one country appears to assert extraterritorial jurisdiction over their scientists, while others do not. It is not reasonable to expect that all countries will eventually adopt similar policies with respect to hESCR. However, the Hinxton Group urged lawmakers, research institutions, and journal editors to take appropriate measures to ensure that, provided the research is conducted in a legal and ethical manner and that participation in hESCR is not expressly prohibited, scientists feel comfortable pursuing international collaborations without fear of prosecution, restriction, or discrimination.

The Hinxton Group also discussed the responsibilities that researchers and journal editors bear to ensure scientific and ethical integrity in hESCR. For example, it was suggested that scientists should submit new stem cell lines to depositories that subscribe to internationally accepted standards of quality. In addition, editors and authors should work together to make all pertinent information regarding hESCR research readily available, including details about the cell lines used and ethical considerations taken.

In the fall of 2006, the Hinxton Group will make available a public database for the deposition of documents relevant to the policies and ethics of hESCR. The group stressed that the rapid evolution of hESCR will require continual development of ethical practices that consider academic, professional, and public opinion. **EG**

Published online September 15, 2006 • 10.1021/cb600392e CCC: \$33.50

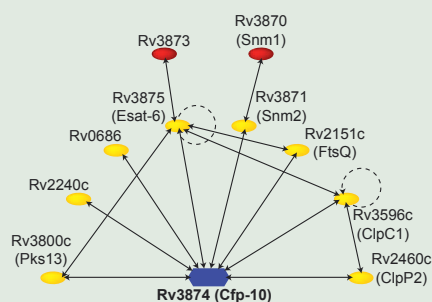
© 2006 by American Chemical Society

Networking Mycobacteria

Astonishingly, nearly one-third of the world's population is infected with *Mycobacterium tuberculosis* (*Mtb*), but the molecular details of its pathogenicity are not well understood. Deciphering the protein interaction networks utilized by *Mtb* would help unravel some of the mysteries of *Mtb* virulence and facilitate drug development against tuberculosis. Yeast two-hybrid (Y2H) technology has been an invaluable tool for unveiling protein interaction networks in many organisms, but the use of yeast as a host can pose various limitations. Singh *et al.* (*PNAS* 2006, 103, 11346–11351) now report the development of a mycobacteria-based cousin of Y2H, termed mycobacterial protein fragment complementation (M-PFC), as an effective method for exploring *Mtb* protein–protein interactions in mycobacteria.

The method is based on the functional reconstitution of murine dihydrofolate reductase (mDHFR) upon the interaction of two mycobacterial proteins, which are independently fused to two mDHFR domains. Active mDHFR confers mycobacterial resistance against trimethoprim (TRIM), and thus mycobacterial growth in the presence

of TRIM is indicative of a protein–protein interaction. Several identified protein pairs, including the



Reprinted with permission from *Proceedings of the National Academy of Sciences*

Mtb secreted immunogenic antigens Esat-6 and Cfp-10, were used to initially validate the system. In addition, a modified assay was developed to quantify the strength of specific protein–protein interactions, significantly

(continued on page 475)

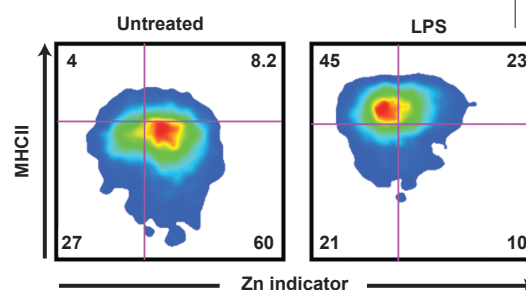
A Link to Zinc

Zinc has received substantial attention for its purported ability to prevent or alleviate symptoms of the common cold. Regardless of the efficacy of this metal as a cold remedy, zinc is an essential element that plays a role in a wide variety of cellular processes, including a well-established but not well-characterized effect on the immune response. Kitamura *et al.* (*Nature Immunology* 2006, 7, 971–977) explore the relationship between zinc homeostasis and immune cell function by examining the effects of manipulating intracellular free zinc levels on dendritic cell (DC) maturation.

An integral part of the immune response depends on the maturation of DCs, which is concomitant with expression of class II major histocompatibility complex (MHC class II) proteins through which antigens are presented to T cells. It is known that the endotoxin lipopolysaccharide (LPS) induces DC maturation through Toll-like receptor (TLR) stimulation. A

zinc ion sensitive fluorescent probe was used to observe that LPS-induced DC stimulation resulted in decreased intracellular zinc concentrations. In addition, treatment of DCs with the zinc chelating reagent *N,N,N',N'*-tetrakis(2-pyridylmethyl)ethylenediamine (TPEN) caused an increase in cell surface expression of MHC class II and induced CD4⁺ T cell activation. By contrast, increasing intracellular DC zinc levels led to inhibition of TPEN-mediated increases in surface expression of MHC class II and LPS-induced movement of MHC class II positive vesicles from the perinuclear area to the cell surface. Intracellular zinc levels are modulated by zinc importer and exporter proteins, and examination of transporter levels upon LPS treatment revealed a TLR-dependent net increase in zinc export. The connection between LPS exposure and free zinc levels was verified *in vivo* when mice injected with LPS exhibited decreased intracellular zinc concentrations and

altered zinc transporter expression. Taken together, these results provide a biochemical connection between intracellular free zinc concentrations, zinc transporter levels, and TLR signaling, illuminating one pathway by which zinc homeostasis modulates the immune response. Furthermore, the implications could go well beyond the immune system; these data indicate that the level of intracellular free



Reprinted with permission from *Nature Immunology*

zinc changes in response to extracellular stimuli, suggesting that zinc acts as a signaling molecule like calcium. If this process can be generalized to other cell types, this would be an exciting finding in the field of signaling pathways. **EG**

Networking Mycobacteria, *continued*

expanding the versatility of the method. To demonstrate the capacity of M-PFC to identify unknown protein interactions, a mycobacterial genomic library was screened for proteins that interact with Cfp-10. Six proteins were uncovered in the screen: one was Esat-6, and the other five were previously unknown to interact with Cfp-10. Intriguingly, examination of the

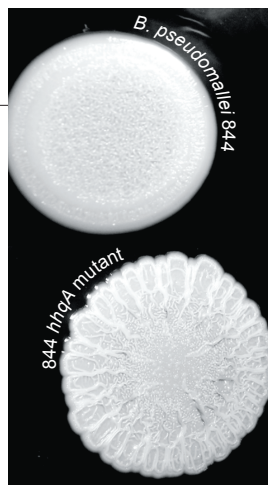
potential functions of these proteins revealed suggestive linkages to the secretory system, including specific involvement in membrane targeting and translocation. This powerful method enabled the elucidation of several components of the protein interaction network of Cfp-10, paving the way for delineating the elusive secretory mechanisms of *Mtb*. **EG**

Spotlight

Making Sense of Quorum Sensing

Quorum sensing refers to the ability of bacteria to use signaling molecules to communicate with one another. A variety of structurally diverse small molecules produced by bacteria regulate critical aspects of their function, including pathogenicity, secondary metabolism, and biofilm development. The opportunistic human pathogen *Pseudomonas aeruginosa* synthesizes dozens of 2-alkyl-4(1*H*)-quinolones (AHQs), including a molecule termed pseudomonas quinolone signal (PQS) that is known to regulate virulence gene expression. However, the roles that other AHQs play in cellular communication mechanisms in *P. aeruginosa* and other bacteria are not well characterized. Diggle *et al.* (*Chem. Biol.* 2006, 13, 701–710) now report that AHQs are synthesized by several species of bacteria and are likely to be an integral part of their quorum-sensing network.

A combination of bioinformatics, bacterial genetics, bioreporters, and analytical chemistry were cleverly combined to determine the existence and potential function of AHQs in several bacterial species. Genomic analysis revealed that in addition to *P. aeruginosa*, other strains of *Pseudomonas* and *Burkholderia* produced AHQs. Of special interest was the human pathogen *B. pseudomal-*



Reprinted with permission from *Chemistry & Biology*

lei, which is responsible for the life-threatening disease melioidosis and has potential uses as a bioweapon. It was initially demonstrated that genetic disruption of PQS synthesis in *P. aeruginosa* can be restored by complementation with the corresponding gene in *B. pseudomallei*, verifying that the genes have a similar function in both species. Use of an innovative combination of thin-layer chromatography and an AHQ bioreporter revealed that, of 20 bacterial strains tested, 9 had the ability to synthesize AHQs, although notably only *P. aeruginosa* strains were capable of producing PQS. A critical role for AHQ signaling was demonstrated in *B. pseudomallei*, when genetic disruption of AHQ synthesis resulted in a striking, wrinkled phenotype and an increase in elastase production. The authors propose that the ability of certain bacteria to generate distinct AHQs may be a critical component of the intricate mechanisms by which quorum sensing is regulated. **EG**

Cannabinoid Crossing

The molecular details behind the enticing therapeutic and psychological effects of cannabinoids like those found in marijuana have been the subject of investigation for decades. The active components of marijuana, such as Δ^9 -tetrahydrocannabinol, as well as endogenous cannabinoids (called endocannabinoids), such as anandamide, elicit their biological effects through interactions with cannabinoid receptors in the brain and select peripheral tissues. However, the uptake and catabolism of these compounds are less well understood. The serine hydrolase fatty acid amide hydrolase (FAAH) has been implicated in the

degradation of fatty acid amides, including anandamide, but the mechanism by which anandamide crosses into the cell for delivery to this enzyme remains elusive. Two recent studies (Dickason-Chesterfield *et al.*, *Cell. Mol. Neurobiol.*, published online May 31, 2006, DOI: 10.1007/s10571-006-9072-6, and Alexander and Cravatt, *JACS* 2006, 128, 9699–9704) employ inhibitors of endocannabinoid transport to provide insight into fatty acid amide metabolism.

Several hypotheses are circulating about how fatty acid amides are delivered from outside the cell to the intracellular membrane compartments that house FAAH. One

model promotes simple diffusion, aided by the lipophilic nature of fatty acid amides. A second hypothesis argues for the existence of a plasma-membrane-associated transporter that facilitates anandamide uptake. Still another paradigm invokes an endocytic process for uptake and transport of fatty acid amides to FAAH. Though structurally unrelated to anandamide, the potent, small-molecule inhibitor of anandamide uptake LY2183240 enabled researchers to refine existing models of endocannabinoid transport and define the utility of compounds of this class.

(continued on page 477)

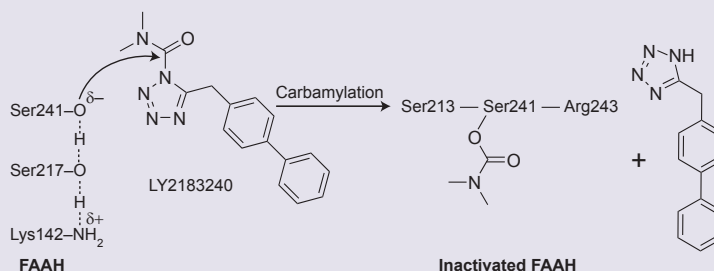
Cannabinoid Crossing, *continued*

Dickason-Chesterfield and colleagues compared several reported inhibitors of anandamide transport, including LY2183240, for their ability to prevent cellular uptake of anandamide and to block FAAH hydrolytic activity *in vitro*. Cellular uptake was assessed in rat basophilic leukemia cells, which actively express FAAH, and in HeLa cervical cancer cells, which do not express FAAH. All of the compounds tested prevented cellular uptake and inhibited FAAH activity, but the potency of each inhibitor was dramatically right-shifted in functional anandamide uptake in the HeLa cells. In addition, in cell membranes from both rat basophilic leukemia and HeLa cell lines, ^3H -LY2183240 identified a high-affinity plasma membrane associated binding site independent of FAAH. Notably, the rank order and K_i values for displacing ^3H -LY2183240 matched the functional anandamide uptake inhibitory constants, lending support for a specific reuptake anandamide transport protein. The authors propose that taken together, the data suggest the existence of a distinct transport protein for anandamide that can adopt high and low binding affinity states, depending on the presence or absence of FAAH.

In a separate study, Alexander and Cravatt scrutinize the inhibitory properties of LY2183240 using functional proteomics. Brain proteomes were treated with LY2183240, and use of the activity-based serine hydrolase probe fluorophosphonate-biotin led to the identification of several enzymes, including FAAH, that were inhibited by this compound. Tryptic digestion and mass spectrometry analysis of purified FAAH treated with LY2183240 revealed that the compound covalently inhibits FAAH. Furthermore, studies in mice confirmed that LY2183240 covalently inhibits FAAH and several other serine hydrolases *in vivo* at pharmacologically efficacious doses. An advanced functional proteomic platform termed ABPP-MudPIT

(activity-based protein profiling multidimensional protein identification technology) was used to identify the other serine hydrolases inhibited by LY2183240, which included α/β -hydrolase 6 and monoacylglycerol lipase. The authors suggest that the promiscuity of this inhibitor is likely due to the reactivity of its heterocyclic urea group, precluding its incorporation into potential pharmaceutical agents. However, use of LY2183240 as a tool to probe endocannabinoid transport suggested the intimate involvement of FAAH in anandamide metabolism. The authors propose that hydrolysis of anandamide by FAAH results in a concentration gradient that drives uptake of the molecule.

The question remains: does a specific anandamide transport protein exist? Evidence presented in these papers provides compelling circumstantial evidence for



Reprinted with permission from the *Journal of the American Chemical Society*

the presence of a transport protein, but it also suggests that inhibition of anandamide uptake by LY2183240 involves direct interaction of LY2183240 with FAAH. However, the inability of the radiolabeled structural analogue of LY2183240 to cross the plasma membrane complicates reconciliation of all the data. Clearly, more studies are needed to elucidate the puzzling process by which anandamide enters the cell. Whether anandamide uptake occurs through simple diffusion or is carried in by a transporter is uncertain, but if the putative transporter does exist, selective small-molecule probes would be welcome tools to help clarify the mechanism. **EG**

Flipping the Lid

Protein degradation in eukaryotic cells is a carefully regulated and essential process. The proteasome is the large ATP-dependent trash barrel for proteins that have been tagged for disposal. Proteins are marked for degradation by covalent attachment of a 76 amino acid polypeptide known as ubiquitin. Ubiquitin status is recognized by the top of the trash barrel, the 19S regulatory lid of the proteasome. After the lid checks for ubiquitin and removes it, the trash barrel portion, the 20S proteolytic particle, is unmasked, and the targeted protein is degraded. A high-resolution look at the proteasome has long proved difficult because of its large size, dozens of protein subunits, and a complex *in vivo* assembly pathway. A recent study from Sharon *et al.* (*PLoS Biol.* 2006, 4, 1314–1324), however, has taken a new approach to proteasome study by harnessing the power of mass spectrometry (MS) and chemical cross-linking. The authors developed a robust method for purifying the intact regulatory subunit, the lid, directly from yeast cell extracts. MS confirmed the presence of all known protein subunits as one major complex and also displayed some putative intermediates and subcomplexes. To gain further

insight, specific parent ion peaks were accelerated through an increased pressure chamber known as a collision cell. Under these conditions, subunits that are on the periphery or are less stably associated are thought to dissociate, and their identity can be assessed using mass spectrometry. Using different collision cell conditions, the authors showed that a number of subunits can be released from the lid. The topology of the lid was also investigated using a chemical cross-linking reagent coupled with MS to identify multimers of the various protein subunits. Several of the associations are in comforting agreement with genetic interaction assays. This study, along with other recent studies on large machines like the ribosome, highlights biochemical applications for MS not only for protein identification but also for spatial and temporal clues into complex assemblies. **JU**

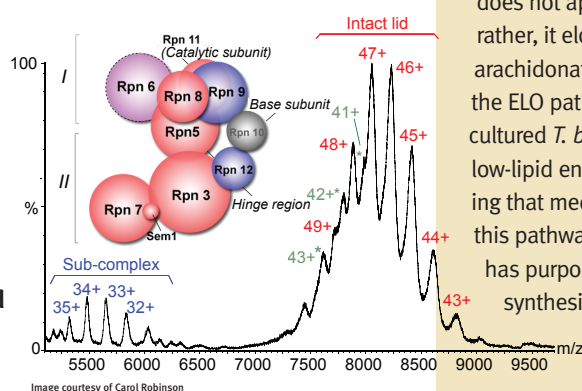
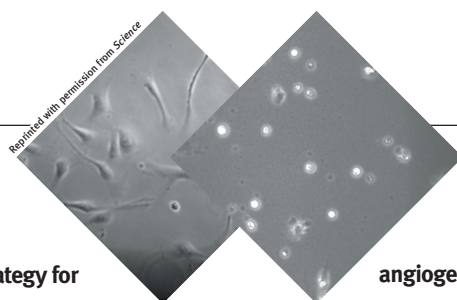


Image courtesy of Carol Robinson

Trypanosomes Synthesize to a Different Drummer

Trypanosoma brucei, the parasite that causes sleeping sickness, exploits two hosts during its life cycle, insects and mammals. In the mammalian bloodstream, *T. brucei* cleverly eludes immune detection by changing its surface coating of variant surface glycoprotein (VSG) molecules. Fatty acids (FAs) are a critical component of the glycosylphosphatidylinositol anchor that tethers the VSGs to the plasma membrane. Interestingly, the insect form and the mammalian form of *T. brucei* have distinct FA needs. Lee *et al.* (*Cell* 2006, 126, 691–699) have discovered that although most organisms, both prokaryotic and eukaryotic, use type I or II synthases to synthesize FAs, *T. brucei* uses microsomal elongases to generate its FAs.

Several pieces of circumstantial evidence suggested that *T. brucei* did not use the typical pathway for FA synthesis. This prompted investigation of the potential role of elongases (ELOs), which are known to extend FAs to longer-chain FAs in other organisms. A cell-free system containing *T. brucei* membranes was used to evaluate FA synthesis, and it was observed that RNAi silencing of ELO1 caused a dramatic reduction in FA synthesis. In addition, knockout strains for each of the four *ELO* genes in *T. brucei* were generated, and thin-layer chromatography and phosphorimaging analysis revealed that *T. brucei* uses a sequential pathway for FA synthesis. ELO1 is responsible for extending a 4 carbon chain (C4) to a 10 carbon chain (C10), ELO2 extends C10 to C14, and ELO3 extends C14 to C18. ELO4 does not appear to be involved in FA synthesis; rather, it elongates the unsaturated long-chain FA arachidonate. It was further demonstrated that the ELO pathway is responsible for FA synthesis in cultured *T. brucei*. Notably, culturing *T. brucei* in a low-lipid environment induced FA synthesis, revealing that mechanisms are in place for regulation of this pathway. The authors propose that *T. brucei* has purposely evolved a unique pathway for FA synthesis that is readily adaptable to the vastly different environments in which trypanosomes must thrive. **EG**

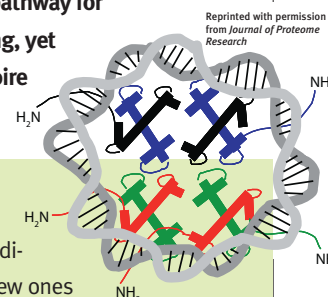


p53 Partners with Collagen

Antiangiogenic therapy is a promising strategy for treating cancer, essentially depriving a tumor of its lifeline by cutting off its blood supply. Understanding the underlying mechanisms of angiogenesis, along with the discovery of novel antiangiogenic agents, will greatly contribute to this field. Several types of collagen proteins possess C-terminal fragments, such as endostatin and tumstatin, that have promising antiangiogenic properties. The seemingly unrelated tumor suppressor protein p53 has also recently been implicated in the regulation of angiogenesis. Teodoro *et al.* (*Science* 2006, 313, 968–971) now report an intriguing link between p53 and α (II) collagen prolyl-4-hydroxylase (α (II)PH), an essential enzyme in collagen biosynthesis, that provides insight into the impact of angiogenesis regulation on cancer.

Various techniques, including polymerase chain reaction based subtractive hybridization and chromatin immunoprecipitation, were used to establish that the α (II)PH gene is a direct target of p53 transcriptional activation. Several human cancer cells lines expressing variants of p53, collagen, or α (II)PH were then created to investigate the roles of these proteins in

angiogenesis regulation. It was demonstrated that p53 dependent expression of α (II)PH ultimately results in an increase in the generation of the antiangiogenic collagen fragments, presumably because of increased collagen synthesis and subsequent processing. Inhibition of α (II)PH with the small-molecule inhibitor ethyl-3,4-dihydroxy benzoate or an antisense oligonucleotide against α (II)PH resulted in decreased endostatin levels, indicating that α (II)PH is a necessary component of endostatin production. Moreover, expression of α (II)PH in the absence of p53 was sufficient to stimulate the emergence of endostatin and tumstatin in conditioned media. Implications of these findings on angiogenesis were demonstrated when conditioned media containing antiangiogenic collagen fragments selectively triggered apoptosis in human umbilical vein endothelial cells. Furthermore, when cells expressing α (II)PH were xenografted into nude mice, tumor growth was dramatically suppressed. The authors propose that p53 induction of α (II)PH expression initiates a pathway for increased collagen synthesis and processing, yet another mechanism to add to p53's repertoire of tumor suppressor activities. EG



Histone-eomics

In the eukaryotic nucleus, DNA is compacted, in part, by the histone family of proteins. This locked-down configuration of DNA, known as chromatin, must rapidly respond to regulatory cues in the cell and open up specific gene regions to allow processes like transcription or DNA repair to occur. The histone proteins themselves play a key role in such transformations. Each histone carries a charged amino-terminal peptide tail that protrudes from the compact chromatin structure. These peptides can recruit or repel trans-acting factors, which might alter the compact state of the DNA. In a new study by Dirksen *et al.* (*J. Proteome Res.* 2006, 5,

2380–2388), a short consensus peptide that resembles all of the histone protein tails was synthesized and used as bait in a nuclear fishing expedition. The peptide was immobilized on a resin, and nuclear extracts from human immune cells were applied to this matrix. Factors that bound to the histone-like peptide were identified in a comprehensive manner *via* mass spectrometry of eluted material. Next, the authors employed the same proteomics but with nuclear extracts from cells that were treated with bleomycin, an agent that induces double-stranded breaks in the DNA. Interestingly, they found that 40 different proteins were no longer recovered

in DNA damage conditions, but also 44 new ones appeared to bind in response to damage. The study goes further and identifies the phosphorylation state of many of the bound proteins. These modifications may play key roles in modulating factor binding to the histone tails. This study demonstrates that a relatively simple experiment combined with a sensitive detection method can yield a host of interesting candidates and new directions for more careful inquiry. Also, because DNA damage is just one of the cellular phenomena to which chromatin must adapt, this methodology may prove useful to those who look closely at gene expression and cell signaling. **JU**

Spotlights written by Eva Gordon and Jason Underwood

Symbiosis: Chemical Biology at Wisconsin

Laura L. Kiessling*, Sally Garbo Wedde, and Ronald T. Raines

Department of Biochemistry and Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, Wisconsin 53706

Chemical biology has been percolating at the University of Wisconsin (UW)–Madison for most of the university's history (www.chemicalbiology.wisc.edu; 1–4). UW–Madison played a predominant role in early research on vitamins—those small molecules that are essential in trace amounts for human life. There, Harry Steenbock developed the irradiation process for the production of vitamin D and thereby eliminated the scourge of rickets (5, 6). Vitamin A was discovered at Wisconsin, as were the vitamin B complex and the hormonal form of vitamin D. Karl Paul Link isolated the potent anticoagulant dicoumarol there and then synthesized warfarin, an analogue that is still a common chemotherapeutic agent. Microbial fermentation methods developed at UW–Madison enabled the large-scale biosynthesis of penicillin and other antibiotics. H. Gobind Khorana carried out the first chemical synthesis of a gene there (7), and W. S. Johnson and Eugene van Tamelen devised synthetic routes to steroids inspired by the biosynthesis of this critical class of natural products (8). The use of natural protease inhibitors to devise highly effective inhibitors of aspartyl proteases was an insight that arose at Wisconsin and was used to design potent HIV protease inhibitors (9). These and other triumphs demonstrate an early symbiosis of chemistry and biology, which fostered the training of eminent interdisciplinary scientists such as Carl Djerassi and Ralph Hirschmann.

Seeking to coalesce the well-established success in research at the interface between

chemistry and biology, Dan Rich sought one of the initial Chemistry–Biology Interface (CBI) Training Grants from the National Institutes of Health (NIH) in 1993. His vision was to provide graduate students with a multitude of opportunities for interdisciplinary training in chemistry and biology. The process of preparing the grant application launched the UW–Madison program in chemical biology (www.chemicalbiology.wisc.edu). The now long-standing NIH training grant remains a core component of graduate training in chemical biology at the university.

Since the inception of the training program, UW–Madison's commitment to graduate-student training in chemical biology has increased. For example, the university has made the recruitment of faculty members with research interests in chemical biology a priority. As a result, graduate students have a wide variety of research options from which to choose. In addition, a powerful infrastructure for conducting chemical biology research has been built. Third, courses in chemical biology have been developed that employ innovative teaching methods. These courses are designed not only to introduce students to concepts in chemical biology but also to build their skills in critical thinking, creative problem selection, and communication. Lastly, students affiliated with



*Corresponding author,
kiessling@chem.wisc.edu.

Published online September 15, 2006

10.1021/cb600368z CCC: \$33.50

© 2006 by American Chemical Society



Image courtesy of L. Winderplögg

UW–Madison CBI trainees (L to R) Jared Mays, Christopher Marvin, Joseph Binder, Rachael Carpenter, Emily English, Matthew Shoulders, Paola Mera, Kelly Gorres, Kimberley Peterson, and Nicholas George.

the training program have an opportunity to gain research experience in industry. Such experiences allow graduate students to make truly informed career decisions.

The aforementioned components were designed to provide graduate students with an education that goes beyond teaching them to design and implement experiments. The Wisconsin approach includes helping students develop the skills to formulate exciting research questions, design appropriate experiments, critically evaluate results, and explain to others why their conclusions are important. It is with this backdrop that the program components are discussed.

RESEARCH OPPORTUNITIES

In 1998, UW–Madison launched an innovative program—the Cluster Hiring Initiative. This bold effort was designed to foster collaborative research, education, and outreach by creating new interdisciplinary areas of knowledge that cross the boundaries of existing academic departments. As an indicator of the UW–Madison’s commitment to chemical biology, it is the focus of one of the selected Cluster Hiring Initiatives (www.clusters.wisc.edu/clusters/show/8), selected by the Provost with input from faculty. A central aspect of this initiative is the recruitment of faculty members with

interdisciplinary research interests. With the new faculty members hired under the initiative added to existing faculty, the university now has >36 research groups with chemical biology research interests. Thus, graduate students have an opportunity to engage with a wide variety of chemical biology faculty members at the cutting edge of the field. Because chemical biology research groups not only are

located in the core departments of chemistry, pharmaceutical sciences, and biochemistry but also are interspersed throughout the university, from bacteriology to pharmacology to chemical and biological engineering. This spectrum of research opportunities allows incoming students to choose a research adviser from a variety of departments. Because chemical biology students are located in many different departments, extensive cross-fertilization of ideas, approaches, and expertise occurs.

RESEARCH INFRASTRUCTURE

UW–Madison has recognized the importance of building the proper infrastructure for conducting research in chemical biology. Graduate students have ready access to state-of-the-art equipment with personnel to help them implement experiments in areas that are new to them and their research group. This infrastructure is especially important to graduate students in chemical biology, who often need access to a wide variety of experimental methods and instrumentation. Some of UW–Madison facilities of interest to chemical biologists are the Keck Laboratory for Biological Imaging (KLBI), the Biophysics Instrumentation Facility (BIF), and the Keck Center for Chemical Genomics (KCCG).

Imaging methods are invaluable for monitoring protein or small-molecule localization

and function. Indeed, chemical biologists continue to make major contributions to this area. Moreover, many chemical biologists benefit from applying modern imaging methods. The KLBI provides the necessary facilities and expertise to the campus. Similarly, the BIF offers instruments for evaluating the strength and stability of biomolecular interactions that require a wide variety of methods. Both centers are staffed by highly experienced scientists who advise students and faculty on implementing experiments.

The KCCG is especially valuable for chemical biologists, because it provides equipment for researchers to synthesize and screen libraries of small molecules. The center is composed of a Chemical Genomics Research Facility and a Compound Screening Facility (CSF). The former has equipment for library generation (liquid handlers and microwave ovens that accelerate library synthesis) and instrumentation used to develop and implement screens (plate readers, surface plasmon resonance imaging, *etc.*). The CSF is integrated with the Comprehensive Cancer Center, providing a conduit for chemical biologists to mix with scientists from different disciplines. The facility contains compound libraries, screening robotics, data-analysis tools, plate readers, and high-throughput microscopy; chemical biologists and biologists are able to identify compounds with biological activity that could serve in drug development or as research tools.

An Interdisciplinary National Cooperative Drug Discovery Group works closely with the KCCG. This UW–Madison consortium of natural-products researchers is using natural products as blueprints to develop new medicines to treat colon, breast, cervical, and pancreatic cancer. The campus-wide group, led by Ben Shen, is producing and testing analogues of natural compounds from microorganisms. These engineered “natural products” provide an alternative

Box 1. Comment from the 2006 Evaluations for the Chemical Biology Course

“This was, by far, the most interesting class I have ever taken—undergraduate or otherwise. With few exceptions, I never noticed the amount of time that had elapsed during lecture until we were dismissed. The topics discussed were very intriguing, the papers used to augment the learning process were applicable and represented leaders in the field, and the instructors were very knowledgeable and interested in engaging the students. This class was an excellent example of a graduate-level course with just the right pace and had a degree of difficulty (for material covered) that served to appropriately stretch the student’s mind. Finally, the proposals were a good addition to the class requirements. Few classes ever truly challenge the student to perform in this capacity—while this was not the first proposal I’ve written, it was the first necessary for a normal class, as opposed to senior-level research. Furthermore, the requirement that the hypothesis utilize chemical-biology-related techniques to answer present, novel questions helped me to grapple with the field effectively.”

to the libraries produced by chemical synthesis.

EDUCATION

The chemical biology curriculum at UW–Madison includes two core courses that have been tailored to interdisciplinary researchers—chemical biology and a chemical biology seminar. These courses have unique features devised to promote independent thinking and creativity. In addition, each course has aspects that help to hone graduate-student skills in written and oral communication.

Chemical Biology Course. The beginning of the chemical biology course is taught from a perspective of how to merge chemical and biological concepts to explore biological systems. The course is organized around the flow of information in biological systems (DNA to RNA to protein) and emphasizes how to use chemical approaches to intervene in each step to elucidate and control that flow. A major goal is to empower scientists: to give chemists relevant novel targets and to offer biologists useful new tools and approaches. Examples of topics include creating small molecules that act to inhibit or enhance transcription and using genetic methods to synthesize proteins containing non-natural residues. An introduction to the chemical concepts underlying

catalysis is given, and chemical approaches to controlling signaling pathways are discussed. In addition to using specific examples, the course focuses on common features among different approaches. For example, many aspects of chemical biology rely on modularity: proteins have modular units, and molecules composed of modular units can be used to alter protein function, localization, degradation, and so forth. When students are shown that approaches that address very different biological questions can be based on similar fundamental concepts, they can begin to recognize how to devise new strategies to solve the biological questions that interest them.

Student participation is a key component of this course. Because the student participants have heterogeneous backgrounds (*i.e.*, some come from biology, others from chemistry), the lively in-class discussion allows students to appreciate different scientific perspectives. Perhaps the most valuable (and fun!) aspect of this course, however, is student participation in peer review through in-class study sections. Each student writes an original research proposal, which is reviewed anonymously in an NIH-style study section composed of a subset of other students in the course. The study sections, in which a faculty member acts as the chair and scientific-review

administrator, offer students the unique opportunity to experience the review process from the perspectives of both an applicant and a reviewer. Students typically find that serving as a reviewer is challenging, and the faculty chairs provide feedback at the meetings on their reviews (*e.g.*, do the criticisms focus on the central issues of the grant?). The students’ feedback reveals that through this exercise, they have developed a deeper appreciation of the importance of clarity in writing (see Box 1). In addition, they also build their communication skills in the study-section meetings; in these venues, the students must present their ideas to a group with heterogeneous scientific backgrounds. The skills that students gain from this experience are critical for success, whether they later choose a career in academia, government, or industry.

Chemical Biology Advanced Seminar.

The participants in the chemical biology seminar are typically advanced graduate students who already have completed the course in chemical biology. In the seminar, students discuss key publications from the past year. The format is designed to encourage an active dialogue. Students are given (or choose) a specific paper to present. A few days before the class meets, the presenter sends three to five discussion questions to other members of the class so that they can consider them as they read the assigned paper. During the course period, the student presenter provides the group with background information that summarizes the key findings and puts them into context. The students then break up into small groups to talk about the discussion questions. During the last 10–15 min, the entire class comes together again to share the ideas that were raised in the small groups. What makes this course interesting and stimulating is that it prompts the discussion of larger issues: the importance of the paper and the questions it seeks to address, the benefits or drawbacks of the approaches used, and the potential future

Chemical biology students benefit from their varied research and training experiences when they select a career.

directions of the research field under discussion. Thus, this seminar allows students to practice their formal presentation skills and to develop their ability to think critically about science.

CAREER CHOICES

Chemical biology students benefit from their varied research and training experiences when they select a career. The CBI trainees, who typically have three consecutive one-year appointments under the training grant, participate in industrial internships. The internships, usually ~12 weeks long, give students an opportunity to gain research experience in a new environment. Many students elect to pursue their internships in the pharmaceutical or biotechnology industries; others carry out research in a government laboratory (e.g., NIH or Los Alamos). Students are uniformly positive about their experiences, and they draw on them in making postgraduation career decisions. This type of training opportunity is available to all chemical biology students.

As described earlier, the CBI Training Grant (GM008505) from the National Institute of General Medical Sciences catalyzed graduate training in chemical biology at UW–Madison. Indeed, this grant, which has been running since 1993, helps fund the training of 10–12 outstanding chemical biology students every year. Moreover, several additional affiliated students receive funding from other sources. The training grant director (Laura Kiessling), the deputy director (Jon Thorson), and an advisory committee (five faculty trainers and one graduate-student trainee) develop and oversee the program in conjunction with the chemical biology trainer faculty. Trainees take the chemical biology course and a course on ethics for scientists and teachers, attend the chemical biology advanced seminar and the chemical biology colloquium, help recruit and welcome new students, participate in industrial research internships, receive an annual allowance

to travel to present their research at scientific conferences, and receive invitations to special and routine seminars by relevant scientists. The CBI committee considers appointments of incoming graduate students nominated by departments and of current graduate students early in their programs nominated by trainers and have made a substantial impact. Although the CBI training grant has been a catalyst for chemical biology graduate education at Wisconsin, the educational initiatives described are open to all graduate students at the university.

Current trends indicate that the boundaries between traditional scientific disciplines will continue to blur (10). We anticipate that educational initiatives like those at Wisconsin and elsewhere (11–17) will continue to facilitate the evolving symbiosis between chemistry and biology.

Note added after print publication: Because of a production error, the following references were misformatted: 5, 7, 10–17. These errors do not affect the scientific integrity of the article. This paper was originally posted September 15, 2006, and the electronic version was corrected and reposted to the web on October 20, 2006. An Addition and Correction may be found in *ACS Chem. Biol.* 1(9).

REFERENCES

1. Curti, M., and Carstensen, V. (1949) *The University of Wisconsin: A History, 1848–1925, Vol. 1*, University of Wisconsin Press, Madison, WI.
2. Curti, M., and Carstensen, V. (1949) *The University of Wisconsin: A History, 1848–1925, Vol. 2*, University of Wisconsin Press, Madison, WI.
3. Cronon, E. D., and Jenkins, J. W. (1994) *The University of Wisconsin: A History, 1925–1945, Politics, Depression, and War, Vol. 3*, University of Wisconsin Press, Madison, WI.
4. Cronon, E. D., and Jenkins, J. W. (1994) *The University of Wisconsin: A History, 1945–1971, Renewal to Revolution, Vol. 4*, University of Wisconsin Press, Madison, WI.
5. Jenkins, J. W. (1991) *A Centennial History: A History of the College of Agricultural and Life Sciences at the University of Wisconsin–Madison*, College of Agricultural and Life Sciences, University of Wisconsin–Madison.
6. Nelson, D. L., and Soltvedt, B. C., Eds. (1989) *One Hundred Years of Agricultural Chemistry and Biochemistry at Wisconsin*, Science Tech Publishers, Madison, WI.
7. van Helvoort, T. (2002) Institutionalizing biochemistry: The Enzyme Institute at the University of Wisconsin, *J. Hist. Med. Allied Sci.* 57, 449–479.

8. Ihde, A. J. (1990) *Chemistry as Viewed from Bascom's Hill*, Department of Chemistry, University of Wisconsin–Madison, Madison, WI.
9. Buckner, C., Connors, K. A., Parascandola, J., Sondecke, G., and Zografi, G. (1997) *The University of Wisconsin School of Pharmacy: Its First Century*, University of Wisconsin–Madison, Office of University Publications, Madison, WI.
10. Kiessling, L. L. (2006) Fostering Major Breakthroughs, *ACS Chem. Biol.* 1, 1–2.
11. Jabri, E. (2006) Teaching a Chemical Biologist, *ACS Chem. Biol.* 1, 397–399.
12. Arndt, H.-D., Niemeyer, C. M., and Waldmann, H. (2006) Chemical Biology Education at Dortmund: A Joint Endeavor with a Max Planck Institute, *ACS Chem. Biol.* 1, 407–410.
13. Kritzer, J. A. (2006) When Undergraduates Ask “Why,” Chemical Biology Answers, *ACS Chem. Biol.* 1, 411–413.
14. Bohle, D. S. (2002) The Evolution and Refinement of a Chemical Biology Training Program: A Canadian Perspective, *ACS Chem. Biol.* 1, 485–486.
15. Bucholtz, T. J., Palfey, B., Mapp, A. K., and Glick, G. D. (2006) Graduate Education Chemical Biology at the University of Michigan, *ACS Chem. Biol.* 1, 487–488.
16. Sulikowski, M. M., and Bachmann, B. O. (2006) Designing a Curriculum That Goes Beyond a List of Topics, *ACS Chem. Biol.* 1, 489–491.
17. Osada, H., Hagiwara, M., Toshima, K., and Imoto, M. (2002) The Education System for Chemical Biology in Japanese Universities, *ACS Chem. Biol.* 1, 492–494.

The Evolution and Refinement of a Chemical Biology Training Program: A Canadian Perspective

D. Scott Bohle*

Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal H3A 2K6, Quebec, Canada

In April 2005, the McGill University chemistry department completely reorganized its graduate education curriculum by throwing out the traditional distinctions of organic, inorganic, physical, and analytical chemistries. They were replaced with the new subdisciplines of chemical biology, materials, synthesis/green/catalysis, and environmental/atmospheric chemistry. McGill is not alone in instituting this type of reorganization, but those familiar with academic inertia will appreciate the magnitude of the change. A year later, we are now well along in our experiment in education, and this article will attempt to synthesize the experience: its successes and outcomes, intended or not.

Throughout this process, we have been surprised at how often the issue of what constitutes chemical biology comes up. Of course, many definitions exist about what chemical biology is (and is not), but in terms of establishing a curriculum, we have found these discussions often dead-end in semantic cul-de-sacs. This is by no means a new problem for emerging subdisciplines in chemistry; as a fledgling science, organometallic chemistry was once described as “concerning those compounds with metal-carbon bonds.” However, an examination of a recent issue of the ACS journal *Organometallics* clearly demonstrates that this restrictive early definition has aged, and many species described in that journal have no metal-carbon bonds or contain those of

only ancillary interest. In a parallel manner, the textbooks of organometallic chemistry have evolved over the past 50 yr. We have every reason to believe that a similar fate awaits contemporary attempts to define chemical biology in an overly narrow or restrictive way.

Perhaps a more enlightened way to proceed toward devising a graduate curriculum is phenomenological. What do the students of chemical biology study? What is it that the researchers in this field find important for their students to know? Here again, in our experience, the answers are as varied as the groups engaged in the field. Current graduate students enrolled in our chemical biology program typically take two of the three classes in bioorganic, biophysical, or bioinorganic chemistry. Students have taken classes in advanced spectroscopy, diffraction, organometallic chemistry, immunology, materials chemistry, and supramolecular chemistry, in addition to those three classes. These courses naturally reflect those currently being offered, and the challenge is determining what new or combined courses best meet the students' needs.

An important component of our chemical biology program is a fellowship scheme



McGill University

Image courtesy of McGill University

*Corresponding author,
scott.bohle@mcgill.ca.

Published online September 15, 2006

10.1021/cb600316j CCC: \$33.50

© 2006 by American Chemical Society

The demand from the students for courses in these new interdisciplinary areas is high.

funded by the Canadian Institutes of Health in 2002 to support, attract, and promote graduate students and postdoctoral researchers in chemical biology. This is a joint interdisciplinary program between the Faculty of Medicine's departments of biochemistry and pharmacology and the Faculty of Science's department of chemistry; ~20 one-year fellowships are available to students who are working in this area. These fellowships provide 100% support for graduate trainees and 50% support for postdoctoral researchers. A 50% match by the principal investigator is required for the latter. It is interesting trying to marry the divergent cultures found in medical and chemistry departments, and an almost equal split occurs in the numbers of students supported between the schools. To determine whether a project is eligible for these fellowships, we apply the operational, old-fashioned definition of chemical biology (1): "Chemical biology can be defined as the design or identification and the exploitation of novel small molecules as tools to investigate questions in biology" (Chemical Biology Research, McGill University, www.medicine.mcgill.ca/biochem/cihr/index_big.html). However, we have found that a wide range of excellent students and their projects are accommodated by this definition. In terms of requirements beyond the core chemistry, the course load is minimal, limited to attendance of selected departmental seminars and two or three workshops over the course of the year. Four years after the program was instituted, the results are positive in terms of the number of students, papers, and presentations. However, the students voice lingering concerns: the cascade of acronyms used in many biomedical talks can be bewildering for the uninitiated, whereas the arcane subtleties of chemical mechanisms seem a trivial exercise for many of the biochemists. It remains a challenge to find the right participants for the seminar series. On the other hand, the workshops have been more successful and

well-received and have included high-throughput screening, RNA silencing techniques, computational chemistry, and career choices upon graduation. The students have benefited from these workshops; in addition, the workshops have been well-received by volunteers from academia and industry who have made presentations.

Where does this leave undergraduate education? A recent bold suggestion by the authors of one of the chemical biology textbooks is that "today physical chemistry is one of three pillars of teaching and research in chemistry departments worldwide, and it is foreseeable that in the not too distant future a branch focusing on the borderline between chemistry and biology will be the fourth pillar" (2). While chemical biology may very well evolve into the fourth pillar of chemical education, the current consensus here is that students need a strong fundamental background in their chosen science. We have not instituted a similar reform for undergraduate education; less latitude for change exists at this level because the courses are subject to accreditation by external agencies. Once again, though, we can look to what courses the students take outside of chemistry. Here, we find that chemistry students, both medicine- and science-bound, have opted to take rigorous courses in areas such as biochemistry and immunology. In one exceptional case, a student with a double major in chemistry and immunology has gone on to graduate school in chemical biology. Thus, the demand from the students themselves for courses in these new interdisciplinary areas is high: the challenge is to integrate this demand with the established programs. Perhaps this is the real issue: when will the established programs recognize the legitimacy of the students' perceived educational needs?

Another perspective on the construction of a chemical biology program is that of the future employers of our students: the

academy and the biotechnical and pharmaceutical industries. Here, the response to the chemical biology subdisciplines has been mixed. The reception by traditional big pharma here in Montreal has been surprising. Some firms welcome our students and provide support for key parts of our lecture series. On the other hand, some firms repeat a mantra: "If we want a chemist to make things, we will hire a synthetic chemist who will make them and a biologist to evaluate them" or "There's no point in coming out of graduate school knowing nothing more about synthesis than methods to produce amide bonds." In spite of this dour assessment of chemical biology, we have found that many of our graduates have gone on to careers in the private sector—often to prepare new amides.

In the future, we intend to refine our graduate education so that it further reflects the needs of our students in this area. In a dynamic, formative discipline such as ours, this process is by definition active as well as iterative. A chemist is motivated to work in chemical biology because of the simple fact that a sizeable amount of chemistry can be learned from our growing understanding of biology. We are in a remarkable period of discovery in this new subdiscipline, and no doubt many new ways of thinking about traditional chemistry will soon emerge from our collective efforts. Along the way, we may have to make and break a few amide bonds.

Acknowledgment: The author gratefully acknowledges support from the CIHR, NSERC, CFI, and the CRC councils and schemes for their support of chemical biology at McGill.

REFERENCES

1. Schreiber, S. L. (1992) Using the principles of organic chemistry to explore cell biology, *Chem. Eng. News* 1992, 22–32.
2. Waldmann, H., and Janning, P. (2004) *Chemical Biology*, Wiley-VCH, Weinheim, Germany.

Graduate Education in Chemical Biology at the University of Michigan

Tonia J. Buchholz[†], Bruce Palfey[‡], Anna K. Mapp[§], and Gary D. Glick^{||,*}

[†]Second-Year Student, [‡]Associate Program Director, [§]Steering Committee Member, ^{||}Program Director, Chemical Biology Doctoral Program, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055

Around the country, chemical biology is increasingly recognized as a focal point for understanding biology at the molecular level—where modern biological and biomedical research is pursued from a uniquely chemical perspective.

Research in chemical biology has been conducted at the University of Michigan for many years in several departments, including biological chemistry, biophysics, chemistry, medicinal chemistry, and pharmacology. Historically, no single program provided large numbers of faculty working in this area. Moreover, graduate training in chemical biology within individual departments often forced students to conform to existing departmental guidelines that may not have been appropriate for them. This meant that in some cases it was difficult for students to get the training they needed.

In response to the need for specific graduate education in chemical biology, a new interdepartmental doctoral program was initiated in 2004. The program is essentially a virtual department and leads to a Ph.D. in chemical biology. Alternative models, in which students apply to an umbrella recruiting program that allows them to select from several departments to do their Ph.D. work after they matriculate, are used by several institutions (including Michigan for the biomedical sciences). Although such a model opened the door for more thesis advisers for students to choose from, we selected the virtual department to ensure that students in the program receive

a high-level, cohesive graduate experience; that the program would have full control over the graduate curriculum; and that a level of research collaboration and mentorship is enabled that is not possible in existing departments. Our program faculty (now ~40) is drawn from seven departments across campus. Students entering the Michigan chemical biology doctoral program are free to choose any of these faculty or groups of faculty as thesis mentors, and this allows them access to most laboratories on campus that do chemical biology research. Our faculty members have a diverse group of interests covering nearly all areas of what can be described as chemical biology.

A major challenge in designing our program is that “chemical biology” means different things to different people. Thus, entering students have varying backgrounds and diverse interests. Rather than trying to impose a specific vision of what chemical biology should be, we have taken a very open attitude. The chemical biology doctoral program at Michigan is structured to maximize the flexibility a student has in designing his or her curriculum and to place the key decision points in the hands of the students. A sequence of two core courses is required; a minimal list of topics is covered that we believe every chemical biologist



*Corresponding author,
gglick@umich.edu.

Published online September 15, 2006

10.1021/cb600315r CCC: \$33.50

© 2006 by American Chemical Society

Students are free to tailor their coursework to suit their own interests and visions of what chemical biology is.

should know, including macromolecular structure and folding, molecular recognition of small and large molecules, catalysis, protein biosynthesis and degradation, signal transduction, combinatorial synthesis, screening, and chemical genetics. In this course, each topic taught includes examples of how both chemists and biologists have addressed the problems being discussed (*e.g.*, use of small molecules and structure-function studies). Two literature-based discussion courses are also required. Beyond these prescribed courses, students fill their remaining credit requirements with whatever courses interest them from the broad range offered throughout the university. In this way, students are free to tailor their coursework to suit their own interests and visions of what chemical biology is.

Although a flexible curriculum is important, research is at the heart of any Ph.D. program. The structure of the chemical biology Ph.D. program emphasizes research training from the beginning. Students are supported by the program in their first year as research fellows, during which time they complete a minimum of two research rotations, which can begin the summer before they matriculate. These rotations generally last a semester, but students have the option of splitting them into half-rotations, maximizing their exposure to research in the labs of different faculty. By April of the first year, students select a faculty thesis mentor and start their thesis work. Students are not required to work as teaching assistants at any time in the program, although arrangements can be made for those desiring that experience. The progress of the students toward their degree is monitored by annual meetings with their thesis committee and with the program leadership team. The program is structured to minimize distractions from research with the goal of having students defend their thesis research within 5 years. Another challenge for a program that spans a large, diverse campus like Michigan is keeping a sense of identity and

coherence both for the faculty and students. This problem is somewhat mitigated because the laboratories of the participating faculty are within walking distance of each other. To further build esprit de corps, the program sponsors monthly lunches for the students. In addition, the members of the steering committee have dinner with all the students twice a year and involve the students in decisions about their training. For example, many of our students are interested in working in more than one lab for their thesis research, so we recently established mechanisms for joint mentorship and collaboration within the program. Such joint mentorship fosters collaborations among program faculty and allows students to create thesis projects that span specialties and reflect the cross-disciplinary nature of research in chemical biology.

How successful are we? Only time will tell, but as we move into the second year, results are extraordinarily encouraging. A class of 6 from around the U.S. matriculated in the first year, and a class of 16 will be entering in the fall of 2006. This puts our program on track to reach its target of 22 *new* students per year in 2007. Our students come from top undergraduate institutions throughout the U.S. and include two students from Western Europe and one from Asia. Their academic credentials as a group place them among the best graduate students in the sciences at Michigan! Much of the success must be attributed to the active involvement of all program faculty in teaching, recruiting, mentoring, and social events. Constant feedback from the students has also been extremely important, allowing us to develop a program that truly meets the needs of our graduate cohort. Finally, our program could not have flourished without the support of the chairs of the participating departments, their respective deans, and the senior administration at Michigan, plus very significant funding to get the program started.

What was attractive about Michigan's chemical biology program? The inaugural

students enrolled in chemical biology were drawn here for numerous reasons. The exciting science being pursued by individual professors, the reputation of the University of Michigan, and the enthusiasm shown by the faculty members and their strong desire to work together across departments to create a successful Ph.D. program are just a few of the key factors. To paraphrase one student: "I feel like I've made an excellent choice. The scientific diversity of the faculty in the program challenges you to become fluent in both the language of biology and that of chemistry. Additionally, the faculty members are as committed to our professional development and success as they are to growing a top-notch chemical biology program that will continue to make important scientific contributions at the crossroads of chemistry and biology."

Note added after print publication: The name of author Tonia J. Buchholz was inadvertently misspelled. This error does not affect the scientific integrity of the article. This paper was originally posted September 15, 2006, and the electronic version was corrected and reposted to the web on October 20, 2006. An Addition and Correction may be found in *ACS Chem. Biol.* 1(9).

Designing a Curriculum That Goes Beyond a List of Topics

Michelle M. Sulikowski* and Brian O. Bachmann

Department of Chemistry and the Vanderbilt Institute of Chemical Biology, Vanderbilt University, 7330 Stevenson Center, Nashville, Tennessee 37235

Institutions with rapidly expanding chemical biology programs are faced with the task of designing a curriculum for a field that resists self-definition. The Vanderbilt Institute of Chemical Biology (VICB) has designed a program for chemical biology education that attempts to mirror the strengths of the qualities of contemporary multidisciplinary research: flexibility and collaboration.

Though the recent proliferation of new journals may suggest otherwise, “chemical biology” is not a new term in the lexicon of scientific research. In fact, the origins of chemical biology can be traced back at least as far as 1946. Linus Pauling and George Beadle first created a joint research program at the California Institute of Technology that aimed for “the analysis of physiological processes in terms of the nature and structure of the chemical substances which are involved in them”; they completed a “laboratory of chemical biology” by 1954 (1). Today, chemical biology has reemerged as a term that means many things to many people, but the points of agreement are a recognition of the inseparability of modern chemistry and biology and an appreciation of the benefits of bringing in-depth expertise in both disciplines together to solve problems of significance in the life sciences. But questions remain about how to cultivate an appreciation for chemical biology in students and how to design a curriculum that provides a meaningful educational experience in this large, diverse field.

Interdisciplinary programs like chemical biology face a unique set of challenges in terms of training. Students typically come to chemical biology with a degree in a traditional science such as chemistry. A program must build on that knowledge base in addition to exposing students to completely new areas for which they may not have a basis for understanding at the graduate level. The balance between stimulating students and overwhelming them is delicate. We have addressed these concerns at VICB in both the foundations of chemical biology course (Figure 1) and the chemical biology seminar (Figure 2), which are part of the newly created Ph.D. in chemical and physical biology.

The chemical biology seminar (Figure 2), taken during the first year, is a vigorous series that features prominent speakers from academia and industry as well as pairs of researchers from within the VICB. Seminars can be challenging for first-year students, but this series offers the additional challenge of absorbing and understanding material that is often out-of-field, and students may not have an appropriate background to scaffold that learning. We have initiated an interactive, peer-led learning



Vanderbilt University

*Corresponding author,
michelle.m.sulikowski@vanderbilt.edu.

Published online September 15, 2006
10.1021/cb6003178 CCC: \$33.50

© 2006 by American Chemical Society

Pauling put it best: “No one method is good enough to solve the problem, and every method must be applied as effectively as possible.”

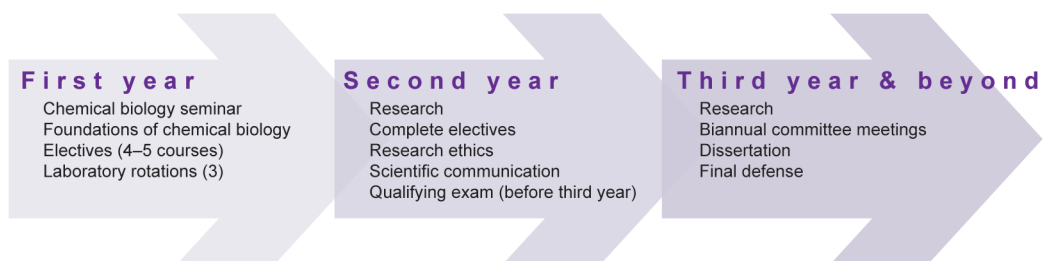


Figure 1. Chemical and physical biology Ph.D. curriculum.

environment that uses BlackBoard, a web-based learning tool, to meet this challenge. Each student chooses one speaker per semester for in-depth learning and becomes the local expert or peer tutor for that seminar. Through a PowerPoint tutorial and interactive discussions, students teach themselves and each other in a supportive environment that meets them where they are in the learning curve. Students learn at their own pace on a level that is appropriate to their scientific backgrounds. After completion of the course, students often report that they form lasting professional and social relationships with peers, achieve a significantly better understanding of the seminar content, feel more at ease asking for and receiving help, and acquire a greater sense of self-directed learning.

The foundation of a chemical biology course (Figure 1) is challenged by the amorphous nature of chemical biology. No clear consensus exists on the definition, but we must offer coursework that captures the essence of chemical biology. The difficulty of teaching such a course is finding bridges between the reductionism used in the study of biological systems and that used to describe molecules (e.g., the chemical bond). Students with a chemistry background are often overwhelmed by the complexity and vocabulary of biological systems, and students with a biology background are often overwhelmed by the vocabulary and science

of chemical structures, reactions, and interactions.

One solution to this dilemma is to assign team-based homework and presentations, which require multidisciplinary collaboration for an integrated solution. In this way, students with different strengths can all bring something to the table in a group that, only as a whole, has enough expertise to solve an integrated chemical biology problem. This mimics the very nature of chemical biology research, which relies on distributed expertise to solve problems. Pauling put it best: “No one method is good enough to solve the problem, and every method must be applied as effectively as possible.”

In our pedagogy, we choose topics that use chemical principles, methods, tools,

and insights to address problems in biology. The course is taught by a small number of faculty from the College of Arts and Science and the Vanderbilt University School of Medicine to provide breadth of coverage while maintaining continuity among lecturers. A risk

of teaching a class entitled chemical biology is to devolve into a series of topics with no clear unifying theme. Hence, we centered our course around two case studies: natural product biosynthesis and drug discovery. Natural product biosynthesis encompasses genomics, biochemistry, chemistry, and pathways, and many of the methods used in the study of biosynthetic pathways are applicable to other areas, such as drug discovery and development. Similarly, modern drug-discovery efforts span many disciplines, from cell biology and signaling pathways to chemical synthesis. These two areas are unified by having a chemical basis and a shared central theme of molecular structure and biological effects of both large molecules and small.

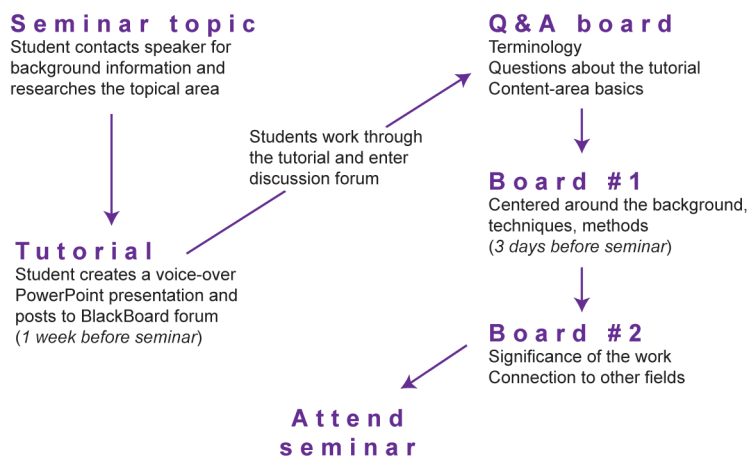


Figure 2. Format of the chemical biology seminar course.

An important goal for any new program is to offer a clear identity to a student or a researcher in the field of chemical biology; this is particularly important to Vanderbilt where students disperse into multiple departments across campus after their first year. The chemical biology course and seminar series, which require students to work with each other across disciplinary lines on a daily basis, have resulted in a sense of community that is maintained by continued attendance at the chemical biology seminar series and the VICB's annual research retreat.

Trends in graduate education and research often translate into areas of undergraduate interest, so we have begun teaching a freshman writing course in chemical biology. The course provides a basic foundation in chemical structures, cell biology, and biomolecules and uses it as a starting point to explore the contemporary literature. We have introduced regular research talks into the course, with several VICB members presenting their work at a level appropriate for freshmen. Students report that although they may not fully understand the details of each seminar, the big picture makes such an impact on their thinking about research and interdisciplinary fields of study that >70% of them actively seek research experiences within the university. Communication is an important focus in the course; 25% of instructional time is committed to discussion about the process of research, orphan drugs, drug discovery as a business, U.S. Food and Drug Administration regulations, and moral and ethical issues related to drug discovery. Students regularly write about their understanding of research lectures and topics that are presented in the course rather than take tests. Students prepare a research report on a topic of their choosing in the area of drug discovery; this undertaking serves as a capstone project. The report emphasizes chemical and biological aspects in addition to social and ethical implications. Students report that

they achieve a greater understanding of the process, promises, and pitfalls of research; become more aware of the complexity of the drug-discovery process; and feel that the course makes their studies in other science courses seem more relevant.

The chemical biology program at the graduate and undergraduate levels has at its core the goal of producing graduates who are self-directed learners capable of both creative and critical thinking across traditional disciplinary lines. We feel that we are building a curriculum that achieves these goals through the use of innovative teaching techniques and that is flexible enough to meet students at their level.

The solutions to the problems of training students in chemical biology will reflect the unique strengths and philosophies of the faculty at a given university. All the colleges, including the medical center, of Vanderbilt University are juxtaposed on a single campus. This greatly facilitates academic and research collaborations between chemists and biologists. Productive interactions are further supported by the existence of the VICB, which was created in 2002 by an investment from the Chancellor's Academic Venture Capital Fund. The VICB has recruited faculty and students, established critical core facilities, and generated substantial extramural support for multidisciplinary research and training in chemical biology. The creation of a Ph.D. program in chemical and physical biology was the inevitable next step in this growth cycle.

REFERENCE

1. Chemical Biology at Caltech (1954) *Pasadena California Institute of Technology Engineering and Science* 17, 10–11.

The Education System for Chemical Biology in Japanese Universities

Hiroyuki Osada^{†,*}, Masatoshi Hagiwara[‡], Kazunobu Toshima[§], and Masaya Imoto[¶]

[†]Discovery Research Institute, RIKEN, Wako-shi, Saitama 351-0198, Japan, and visiting professor, Tokyo Medical and Dental University and Keio University, [‡]School of Biomedical Science, Tokyo Medical and Dental University, Bunkyo-ku Tokyo 113-8510, Japan, [§]Department of Applied Chemistry, and [¶]Department of Biosciences and Informatics, Keio University, 3-14-1Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan

In Japan, organic chemistry has been the traditional approach to elucidating biological phenomena. The department of agricultural chemistry and the department of pharmacology have been the backbone of chemical biology, and many students graduating from these departments are employed

by pharmaceutical companies. Therefore, Japanese universities have not needed to reorganize their departments to conduct research at the interface of chemistry and biology.

As the field matures and grows, however, some universities do plan to establish a department for chemical biology, and

chemical biologists in Japan have organized the Japanese Association of Chemical Biology to facilitate research. To place the education system in perspective for all readers, it's important to note that Japan has national, public, and private universities. As examples of how chemical biology is taught at these universities, we will discuss the graduate program at Tokyo Medical and Dental University (TMDU; national) and the undergraduate and graduate programs at Keio University (private).

School of Biomedical Science, TMDU.

Research emerging from the decoding of the human genome opened the door to practical research and collaboration between

fields. The fields now working together include molecular and cellular biology, structural biology and immunology, neuroscience, pharmacology, and bioinformatics. In response to the social and academic demands of the new era, the School of Biomedical Science of TMDU was established in 2003; it is the first postgraduate school engaged in the education and research of postgenome medicine in Japan. In this graduate school, we started the biomedical science Ph.D. program to create scientists who have the ability to manage expanding information, resolve practical problems, and promote innovation in the life sciences. This program has two courses, bioinformatics and functional biology, and accepts undergraduate students educated in medicine, biology, bioscience, chemistry, pharmacology, informatics, and systems engineering from all over the world. To support the widening research fields, outstanding institutes in the Tokyo area, such as the National Cancer Center, the Tokyo Metropolitan Institute, and RIKEN, are contributing their resources and scientific expertise to this new program. For detailed information on our Ph.D. program, go to www.tmd.ac.jp/mri/SBS/index_e.html.

Chemical biology is one of the major subjects of our Ph.D. program. Chemical biology involves solving biological problems at the molecular level and using the techniques, knowledge, and ideas of chemistry to regulate biological systems. We developed a



*Corresponding author,
hisyo@riken.jp.

Published online September 15, 2006
10.1021/cb6003614 CCC: \$33.50

© 2006 by American Chemical Society

TABLE 1. Where students from Keio University go after graduation^a

	Undergraduate students	Graduate students
Employed at companies		
Pharmaceutical	5	30
Food	5	10
Chemical	7	30
Other	3	20
Work at institutes	0	5
Work at universities	0	5
Attend graduate school	80	

^aA total of 100 undergraduate students and 100 graduate students were surveyed.

lecture series and a practice course. The lecture series provides an overview of the chemical biology field and topics of recent research. *Bioprobes* (Osada, H., Ed.; Springer: New York, 2000) and *Molecular Cell Biology* (Lodish, H., Berk, A., Zipursky, L., Matsudaira, P., Baltimore, D., and Darnell, J., Eds.; W. H. Freeman and Co.: New York, 2000) are used as primary texts. We invite prominent researchers from Astellas Pharma, Inc., a major pharmaceutical company in Japan, as adjunct professors to teach the practical application of chemical biology for drug discovery. In the practice course, students learn molecular design and the syntheses of bioprobes, the structural analyses of bioprobes, and their applications to the biological systems in an experimental and hands-on manner. In this course, students also experience high-throughput screening with a chemical robot.

Department of Biosciences and Informatics, Keio University. As the sequencing of the human genome approaches completion, we must work to elucidate this new life science. To enter this new era, Keio University established the department of biosciences and informatics in 2002. Research activity at our new department involves unraveling the mysteries of life systems not only by using our knowledge of biology, molecular biology, and chemical biology but

also by explaining the solution in terms of informatics.

In their first year, undergraduate students are required to take biology, physics, chemistry, and mathematics, which are the common ground of the natural sciences and technology. In the second year, they study the three basic disciplines on which our department is founded: biology, chemistry, and bioinformatics. Students are encouraged to expand their studies beyond the traditional framework of biology by looking at life from the viewpoint of molecular theory and by conducting computerized analyses in the lab to process data on biological function.

Third- and fourth-year students can systematically investigate various areas of biology, including chemical, molecular, and cell biology; biochemistry; and genome technology. They also study data mining and pattern recognition, two essential tools in informatics for understanding biological phenomena and function.

Department of Applied Chemistry, Keio University. The department of applied chemistry covers many areas of pure and

applied chemistry, including chemical biology. In the field of chemical biology, the research activity at this department involves both the synthesis of bioactive molecules and the analysis of life systems.

In their first year, students are required to learn fundamental subjects, such as mathematics, physics, chemistry, and biology. In the second year, they study the four basic disciplines on which our department is founded: physical, inorganic, organic, and biochemistry. Undergraduates are encouraged to take their studies beyond the traditional framework of chemistry: to combine and modify subjects to create new ones, such as chemical biology. Third- and fourth-year students can systematically learn about the various realms of chemistry, such as analytical, environmental, material, synthetic organic, polymer, and biochemistry.

Graduate School of Science and Technology, Keio University. Advanced courses of chemical biology are taught by the staff from our graduate school as well as instructors from other graduate institutes. Students

can investigate natural product chemistry, chemical genetics, and signal transduction systems for cell function and their regulation by using small molecules, drug discovery, medicinal chemical synthesis, and so on. They also



Tokyo Medical and Dental University Students

study systems biology from a chemical biology perspective. A graduate student who wants to study chemical biology first selects an adviser. The student then conducts his or her research under the guidance of the adviser during the two-year master's program. When they complete the program, students choose to either graduate with a master's degree or to advance to the 3-yr Ph.D. program.

The data in Table 1 indicate that 80% of undergraduate students go on to pursue

graduate degrees. The data also show that students who graduate from the chemical biology course will work for chemical companies as well as pharmaceutical companies.

One student who attends the graduate school at Keio University and studies cell biology stated that “the chemical biology approach enables us to easily study signal-transduction systems involved in cellular events of interest.” Another graduate student in the chemistry field stated that “chemical biology education tells me the importance of chemistry in life sciences and prompts me to participate in biosciences from the viewpoint of chemistry.”

With the growing interest in chemical biology, our chemical biology education program at Keio University will progress by promoting interdisciplinary research and training in systems biology. This program will train students to be scientists and engineers who have new viewpoints different from those of students of classical chemistry or traditional biology and who will have the skills to flourish in various fields around the world.

MLL Core Components Give the Green Light to Histone Methylation

Brendan D. Crawford and Jay L. Hess*

Department of Pathology, University of Michigan Medical School, 1301 Catherine Road, M5240 MS1, Ann Arbor, Michigan 48109-0602

Over the past decade, considerable progress has been made in understanding how transcription is regulated through covalent modifications of histones. It is now clear that in addition to acetylation, phosphorylation, and ubiquitination, lysine methylation of histone tails plays a fundamental role in transcriptional regulation (1). Defining how the “readers, writers, and erasers” of this epigenetic code interact and how their activity is regulated is currently a major area of interest for many in the transcription field. A recent paper by Dou *et al.* (2) provides important insights into regulation of histone H3 Lys4 (H3K4) methylation.

Methylation of lysines on histone H3 and H4 tails confers either activating or silencing effects on transcription, depending on the specific modified residue (1). Dimethylation and trimethylation of H3K4 are associated with the coding region of actively transcribed genes (3). In contrast, histone H3 Lys9 (H3K9) and Lys27 (H3K27) methylation is associated with transcriptional repression. With the exception of Dot1, an H3 Lys79 methyltransferase (MT), all of the lysine MTs share an evolutionarily conserved catalytic domain, the SET domain, whose name is derived from the MTs Su(var)3-9, Enhancer of Zeste (EZH), and Trithorax. Only one H3K4 MT, Set1, has been identified in yeast; however, many human Set1 homologues have been identified, including Set1a, Set1b, and four members of the Mixed-Lineage Leukemia (MLL) family

(4–9). Of these, MLL1, which is homologous to the *Drosophila* protein Trithorax, has been the most intensively studied because of its involvement by chromosomal translocations in a variety of acute lymphoid and myeloid leukemias (10). Several years ago, the bacterially expressed MLL SET domain was shown to have modest MT activity itself (11). However, like other Set1 family members, MLL1 exists as part of an MT complex with enhanced MT activity (6) that includes a number of other proteins, such as the histone acetyltransferase MOF (males absent on the first) and three core components, WDR5, RbBP5, and Ash2L (12, Figure 1). The finding that these core components are evolutionarily conserved from yeast to humans and are shared among the different human Set1 family members suggests that they play important roles in regulating MT activity. Indeed, recent work by Wysocka and colleagues (13) shows that WDR5 is required for histone H3K4 trimethylation. In this report, WDR5 was found to bind preferentially to dimethylated H3K4, and its knockdown resulted in decreased expression of MLL1 target genes without affecting binding of MLL1 complexes to these targets. Several recent manuscripts (14–18), notably the study by Dou and colleagues (2), provide a clearer picture of the roles that not only WDR5 but also RbBP5 and Ash2L play in regulating MT activity.

An MT Structural “Presentation” Platform. Dou and colleagues (2) used baculovirus expression in insect cells and

ABSTRACT Trimethylation of histone H3 Lys4 (H3K4) is associated with transcriptional activation. One of the chief effectors of H3K4 methylation is mixed-lineage leukemia 1 (MLL1), a gene that is disrupted by chromosomal translocation in acute leukemia and a master regulator of *Hox* and other genes. In a recent paper, core components of the human MLL histone methyltransferase (MT) complex were found to form a structural platform, with one component (WDR5) mediating association between the specific histone H3K4 substrate and the MT. This novel regulatory mechanism, which is conserved from yeast to human, is required for both methylation and downstream target gene transcription.

*Corresponding author,
jayhess@umich.edu.

Published online September 15, 2006

10.1021/cb600367v CCC: \$33.50

© 2006 by American Chemical Society

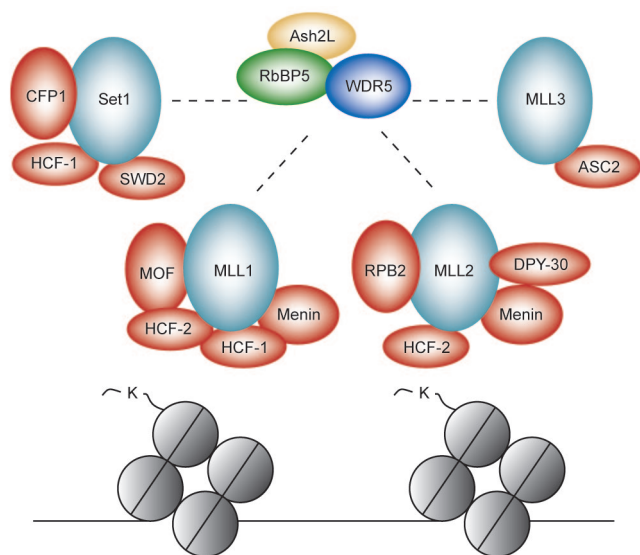


Figure 1. MT structural platform. WDR5, RbBP5, and Ash2L, core components of the MLL1 MT complex, form a structural platform for the different Set and MLL-family MTs that are required for trimethylation of histone H3K4.

immunoaffinity purification to reconstitute a functional MT complex composed of WDR5, RbBP5, Ash2L, and the catalytic C-terminus of MLL1 (MLL-C) *in vitro*. Various combinations of the three subunits were expressed through the baculovirus system, and interactions were examined through immunoprecipitation assays to determine the structural organization of the core complex. WDR5 and RbBP5 were shown to jointly mediate association with MLL-C. Interestingly, WDR5, RbBP5, and Ash2L formed an independent complex in the absence of MLL-C, an indication that together they form a structural platform for association with the different Set1 or MLL-family H3K4 MTs (Figure 1). After Dou and colleagues (2) determined the structural contributions to the core complex, they examined the functional contribution of each core component to H3K4 MT activity. *In vitro* histone MT (HMT) assays revealed that the absence of RbBP5 or Ash2L in the complex significantly reduces H3K4 trimethylation. In contrast, the absence of WDR5 completely abolishes methylation activity.

groups have reported the crystal structure of WDR5, which contains seven WD40 repeats organized in a whorl of β -“propeller blades”, each composed of four antiparallel pleated sheets (15–18). These studies show that a central depression formed by the WD40 repeats specifically recognizes N-terminal H3 peptides *via* interactions primarily with residues A1, R2, and T3. Mutations of any of these three residues inhibit binding. All studies show the H3K4 is solvent-exposed and thus available for further methylation.

Most studies that analyzed binding of WDR5 to histone tails report higher affinity of WDR5 for dimethylated H3K4 than the unmodified peptide (15, 18), although Couture *et al.* (16) report that WDR5 has virtually identical affinities for mono-, di-, or trimethylated H3K4. Subtle kinetic differences may account for why dimethylated H3K4 is a preferred binding partner in *in vitro* assays (15, 18), which are likely to be magnified in more *in vivo* situations. In support of this, pull-down studies in the context of intact MLL complexes show preferential interac-

Subsequent *in vivo* studies based on RNAi on transfected cells by Shilatfard and colleagues (14) show very similar findings.

The WDR5 Link.

The above results show that MLL1 requires the WDR5-RbBP5-Ash2L structural platform for full catalytic activity and indicate a central role for WDR5 in not only MLL1 MT activity but also likely for the other H3K4 MTs as well. Four

tation with K4 dimethylated tails (19). One take-home message is clear: WDR5 is required for full MLL1-complex mediated H3K4 trimethylation, possibly through regulating its activity.

Dou *et al.* (2) explored this possibility by constructing WDR5 mutants with a disruption of each of the three residues that make contact with histone H3 peptides, as determined by the structural studies above. Two of the WDR5 mutants, S91K and F133A, disrupted the interaction of MLL1 with the complex. The Y191F mutant was particularly instructive because it preserved the integrity of the MLL1 complex and dramatically reduced both binding of the complex to histone and MLL1 HMT activity. Together with the structural studies, Dou *et al.*'s (2) work indicates that WDR5 plays a unique role by linking the catalytic subunit of MLL-C and its histone H3K4 substrate in such a way that the complex can bind mono- or dimethylated H3K4 while “presenting” K4 for further methylation (Figure 2), a concept suggested from structural studies (15, 18).

Effects on MLL1-Mediated Transcription.

After Dou and colleagues (2) identified a functional role for WDR5 in MLL1-mediated HMT activity, they examined the role of

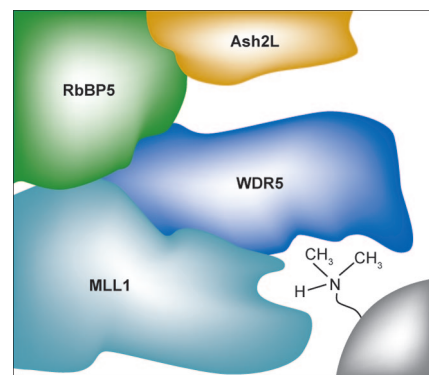


Figure 2. WDR5 link. WDR5 interacts with methylated H3K4, presenting it for further methylation and, with MLL1, linking the catalytic MT domain with its substrate. This interaction is necessary for H3K4 trimethylation and subsequent transcription of MLL1 target *Hox* genes.

WDR5-RbBP5-Ash2L in regulating the transcription of two well-characterized MLL1 target genes, *HOXA9* and *HOXC8*. An analysis of separate small interfering RNA (siRNA) knockdowns of each of the components revealed reduced expression of *HOXA9* and *HOXC8*, compared with the core complex. Chromatin immunoprecipitation experiments revealed that trimethylated H3K4 was reduced at the *HOXA9* locus when each of the three components was individually knocked down with siRNA. Trimethylated H3K4 was also reduced at the *HOXC8* locus with RbBP5 or WDR5 knockdown and dimethylated H3K4 with RbBP5 knockdown. Of note is a very recent report by Shilatifard and colleagues (14), who used a similar approach and showed that ASH2L is also pivotal for trimethylation of MLL1 targets. These combined siRNA knockdown results demonstrate that all three core components, WDR5, RbBP5, and Ash2L, are needed for full MLL1 target gene expression and that this requirement lies in the regulation of H3K4 dimethylation and/or trimethylation. Notably, in the study by Dou *et al.* (2), MLL1 recruitment to either the *HOXA9* or *HOXC8* locus was not disturbed by knockdown of WDR5, RbBP5, or Ash2L, an indication that these core components are not involved in MLL1 recruitment but rather in regulating MLL1 complex MT activity.

Looking Ahead: Biological and Chemical Implications. This study provides important insights into how histone methylation is regulated by bringing together specific MTs and histone tails *via* scaffolding components with WD40 repeats such as WDR5. One area of interest will be to determine whether the expression of the core subunits, or perhaps their post-translational modifications (PTMs), are developmentally regulated and whether, for example, this plays a role in the regulation of target genes such as the *Hox* genes. It will also be important to identify the specific sites of interaction of the SET domain with core components and whether these interactions are developmen-

tally regulated. The function of many of the other components of the MLL complexes, such as HCF-1 and -2, Dpy-30, and menin, remain to be determined. Using approaches similar to those employed by Dou *et al.* (2) should be a powerful tool for resolving these issues. It is intriguing that the amino acids in the histone H3 tail recognized by WDR5, in particular R2 and T3, are known to be methylated and phosphorylated *in vivo*. Couture *et al.* (16) showed that these PTMs inhibit WDR5 binding, and this lends further credence to the concept of a “phosphomethyl switch” in which the modification status of the histone tail regulates its ability to undergo subsequent PTMs (20).

Another important question will be to determine whether similar mechanisms apply to other HMTs involved in transcriptional activation and repression. Other WD40-containing proteins, such as Embryonic ectoderm development (Eed) and RbAp46/48, have been shown to associate with repressive EZH2 complexes that methylate Lys27 (21, 22). Furthermore, Eed knockout mice show global defects in histone H3K27 methylation, an indication that Eed plays a role similar to that of WDR5 in facilitating histone methylation (23). Because considerable amounts of WDR5 are present in lower-molecular-weight complexes compared with those identified for either the MLL family or EZH2 (14), determining what roles WDR5 and related WD40 repeat proteins may play beyond histone methylation will be very interesting.

Acknowledgment: We thank C. David Allis, Alexander J. Ruthenburg (Rockefeller University), and Joanna Wysocka (Stanford University) for critical review of this article and Jinron Min (University of Toronto) for permission to communicate results prior to publication.

REFERENCES

- Martin, C., and Zhang, Y. (2005) The diverse functions of histone lysine methylation, *Nat. Rev. Mol. Cell Biol.* 6, 838–849.
- Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components, *Nat. Struct. Mol. Biol.* 13, 713–719.

- Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999) Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena, *Proc. Nat. Acad. Sci. U.S.A.* 96, 14967–14972.
- Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. (2001) The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4, *EMBO J.* 20, 7137–7148.
- Lee, J. H., and Skalnik, D. G. (2005) CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex, *J. Biol. Chem.* 280, 41725–41731.
- Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L., and Roeder, R. G. (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF, *Cell* 121, 873–885.
- Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacherjee, A., Biondi, C. A., Kay, G. F., Hayward, N. K., Hess, J. L., and Meyerson, M. (2004) Menin associates with a trithorax family histone methyltransferase complex and with the *hoxc8* locus, *Mol. Cell* 13, 587–597.
- Goo, Y. H., Na, S. Y., Zhang, H., Xu, J., Hong, S., Cheong, J., Lee, S. K., and Lee, J. W. (2004) Interactions between activating signal cointegrator-2 and the tumor suppressor retinoblastoma in androgen receptor transactivation, *J. Biol. Chem.* 279, 7131–7135.
- Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N., and Herr, W. (2003) Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1, *Genes Dev.* 17, 896–911.
- Hess, J. L. (2004) Mechanisms of transformation by MLL, *Crit. Rev. Eukaryotic Gene Expression* 14, 235–254.
- Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., and Hess, J. L. (2002) MLL targets SET domain methyltransferase activity to *Hox* gene promoters, *Mol. Cell* 10, 1107–1117.
- Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C. M., and Canaani, E. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation, *Mol. Cell* 10, 1119–1128.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development, *Cell* 121, 859–872.
- Steward, M. M., Lee, J. S., O'Donovan, A., Wyatt, M., Bernstein, B. E., and Shilatifard, A. (2006) Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes, *Nat. Struct. Mol. Biol.* published online Aug 6, <http://dx.doi.org/10.1038/nsmb1131>.

15. Ruthenburg, A. J., Wang, W., Graybosch, D. M., Li, H., Allis, C. D., Patel, D. J., and Verdine, G. L. (2006) Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex, *Nat. Struct. Mol. Biol.* **13**, 704–712.
16. Couture, J. F., Collazo, E., and Trievel, R. C. (2006) Molecular recognition of histone H3 by the WD40 protein WDR5, *Nat. Struct. Mol. Biol.* **13**, 698–703.
17. Han, Z., Guo, L., Wang, H., Shen, Y., Deng, X. W., and Chai, J. (2006) Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD40 protein WDR5, *Mol. Cell.* **22**, 137–144.
18. Schuetz, A., Allali-Hassani, A., Martin, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A. N., Arrow-smith, C. H., and Min, J. (2006) Structural basis for molecular recognition and presentation of histone H3 by WDR5, in press.
19. Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C., and Allis, C. D. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodeling, *Nature* **442**, 86–90.
20. Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., Allis, C. D. (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation, *Nature* **438**, 1116–1122.
21. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing, *Science* **298**, 1039–1043.
22. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein, *Genes Dev.* **16**, 2893–2905.
23. Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P., and Magnuson, T. (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation, *Curr. Biol.* **15**, 942–947.

Glyco-Stripping and Glyco-Swapping

Charles E. Melançon III[†], Christopher J. Thibodeaux[‡], and Hung-wen Liu^{§,*}

[†]Department of Chemistry and Biochemistry, [‡]Institute for Cellular and Molecular Biology, and [§]Division of Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712

Many bioactive natural products derive their activity from the sugar components of their structures. Variations in the structures of these sugars can have profound impacts on the biological activity, selectivity, and pharmacokinetic properties of the parent compounds (1, 2). Because of the important role sugars play in natural product bioactivity, the development of robust chemical and enzymatic methods to derivatize natural products with diverse sugar moieties has attracted much attention. Several complementary strategies, including semisynthesis (3, 4), pathway engineering (5–7), and *in vitro* enzymatic glycosylation techniques (4, 8, 9), have emerged from recent efforts as effective means of varying natural product sugar structures. The feasibility of biosynthetic and chemoenzymatic glycodiversification methods relies on the substrate tolerance of glycosyltransferases (GTs), enzymes that catalyze the coupling of sugar donors to aglycon acceptors. Accumulating evidence suggests that GTs involved in natural product biosynthesis are generally substrate-flexible, which bodes well for future engineering work. Although glycodiversification with these techniques has enjoyed much success, sugar randomization by any method remains challenging. In particular, generation of activated nucleotide diphosphate (NDP)-sugars and appropriate acceptor substrates often need elaborate chemical or enzymatic syntheses (10), and *in vivo* engineering methods require detailed information about biosynthetic pathways as well as well-behaved genetic

systems for gene disruption and heterologous expression (6).

The recent *Science* article by Thorson and coworkers (11) presents findings that have the potential to revolutionize the way researchers alter the structures of glycosylated natural products. First, they established that several GT-catalyzed reactions are reversible. Hence, with the addition of excess NDP, they were able to generate NDP-activated sugars from glycosylated natural products by running the GT reaction in reverse. This finding is significant because it provides access to NDP-sugars in a single enzyme-catalyzed step, bypassing the need for any tedious chemical or enzymatic syntheses. Second, they performed one-pot sugar exchange and aglycon exchange reactions, in which a GT can couple either product of its reverse reaction (the NDP-sugar or aglycon) to a non-natural cosubstrate (a different NDP-sugar or aglycon). Notably, this methodology yielded >70 glycosylated natural product analogues in a relatively straightforward manner.

After the proposed role of CalG1, the 3-*O*-methyl-L-rhamnosyl transferase involved in calicheamicin (CLM) biosynthesis, was verified, 10 other thymidine 5'-diphosphate (TDP) sugars (Figure 1, top, a–j) were identified from a TDP sugar library as alternative substrates of CalG1. The enzyme was incubated with CLM α_3 (1) and an alternative sugar substrate, TDP-3-deoxy-D-glucose (c), to test the regiospecificity of the CalG1-catalyzed reaction. Serendipitously, two new products, which were subsequently

ABSTRACT Glycosyltransferases (GTs) are ubiquitous enzymes that catalyze the transfer of a sugar moiety from an activated donor to an acceptor and thus play important roles in natural product biogenesis, virulence, and biomolecular recognition. Sugars are often critical for bioactivity of natural products, and methodologies for creating diverse glycoforms of these compounds are highly desirable. A recent study demonstrates that several GTs involved in natural product biosynthesis catalyze reversible reactions. Sugar exchange and aglycon exchange strategies were used to exploit this reversibility to generate >70 calicheamicin analogues.

*Corresponding author,
h.w.liu@mail.utexas.edu.

Published online September 15, 2006

10.1021/cb600365q CCC: \$33.50

© 2006 by American Chemical Society

identified as deglycosylated CLM α_3^1 and an analogue with the 3-*O*-methyl-L-rhamnose moiety replaced by 3-deoxy-D-glucose (**1c**), were detected. Control reactions showed that the observed sugar exchange is a result of CalG1-catalyzed deglycosylation (reverse glycosylation resulting in the formation of a true sugar nucleotide intermediate) of 3-*O*-methyl-L-rhamnose followed by re-glycosylation with 3-deoxy-D-glucose. Taking advantage of the reversible property of CalG1-catalyzed reaction, the authors used the 10 alternative sugar substrates (**a–j**) to judiciously explore the ability of this enzyme to catalyze 3-*O*-methyl-L-rhamnose removal and subsequent glycodiversification of 8 CLM analogues (**1–8**).

In all cases, the sugar-exchanged product was identified, and a library of 72 new CLM derivatives (**1a–j**, **2a–j**, **3a–j**, **4a–j**, **5a–j**, **6a–j**, **7a–j**, **8a**, **8b**) was generated, with an average yield of 60% (Figure 1, inset box). These experiments nicely illustrate the power of using the reverse GT reaction for glycodiversification of natural products.

Also demonstrated in this work is the elucidation of GT function *via* a reverse glycosylation assay. Several CLM derivatives were used to assay CalG4, a GT proposed to be

responsible for the attachment of the aminopentose moiety of CLM. The fact that CalG4 is capable of removing the attached sugar in all cases confirms its assigned role. This enzyme, like CalG1, can also catalyze sugar exchange. The development of this assay based on the reversibility of GT-catalyzed reactions has significant implications for the study of GTs. Because the GT is assayed in the reverse direction, the NDP-sugar and aglycon substrates for the forward reaction are not required. For compounds with mul-

tipule sugar appendages, this assay could be used not only to identify which GT is responsible for a given glycosyltransferase step but also to reveal the order of glycosylation events.

To better assess the prevalence of synthetically useful reverse GT reactions in nature, the authors assayed the well-studied GTs from the vancomycin biosynthetic pathway, GtfD and GtfE (12), and found that both catalyze reversible reactions. GtfD could also catalyze aglycon

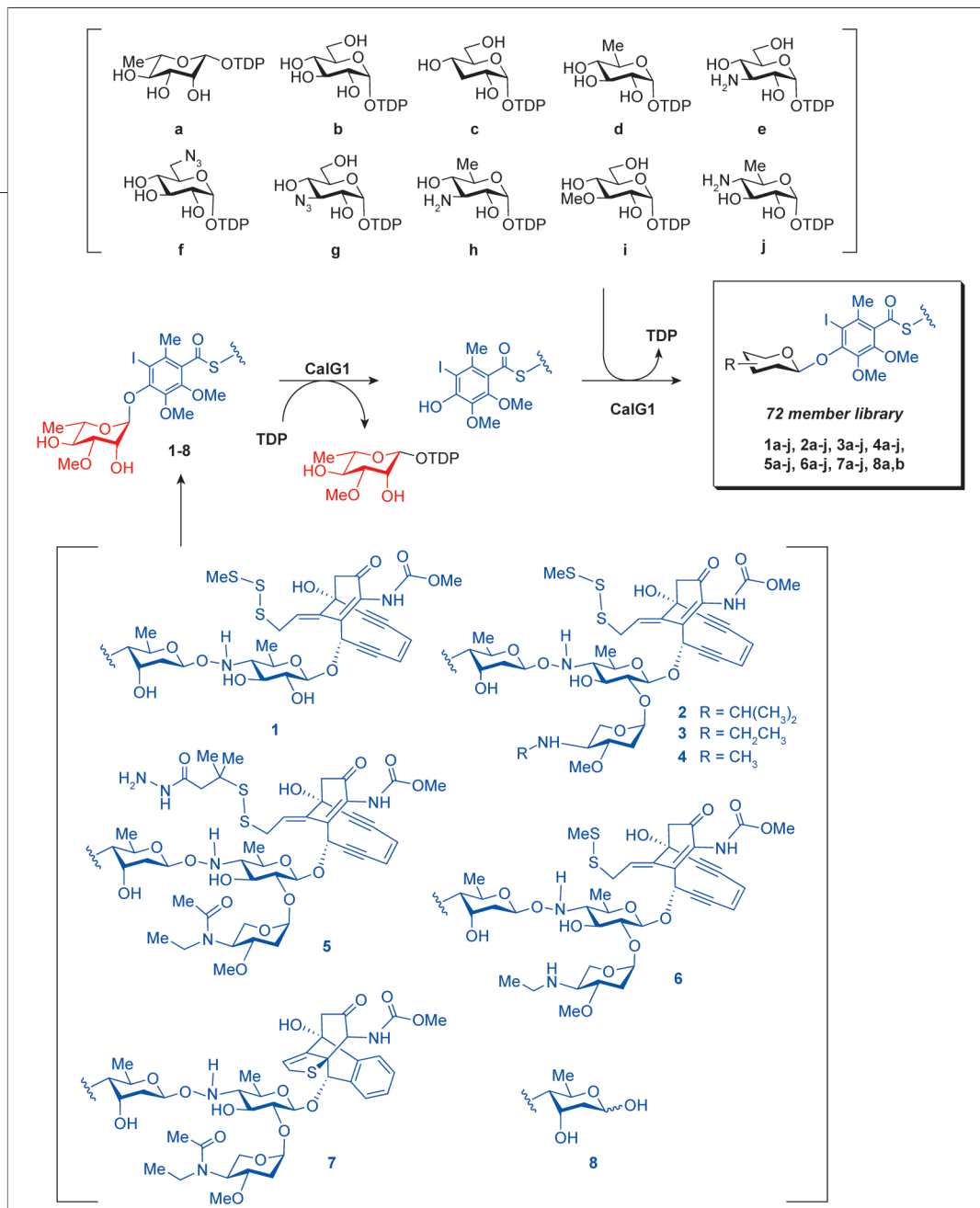
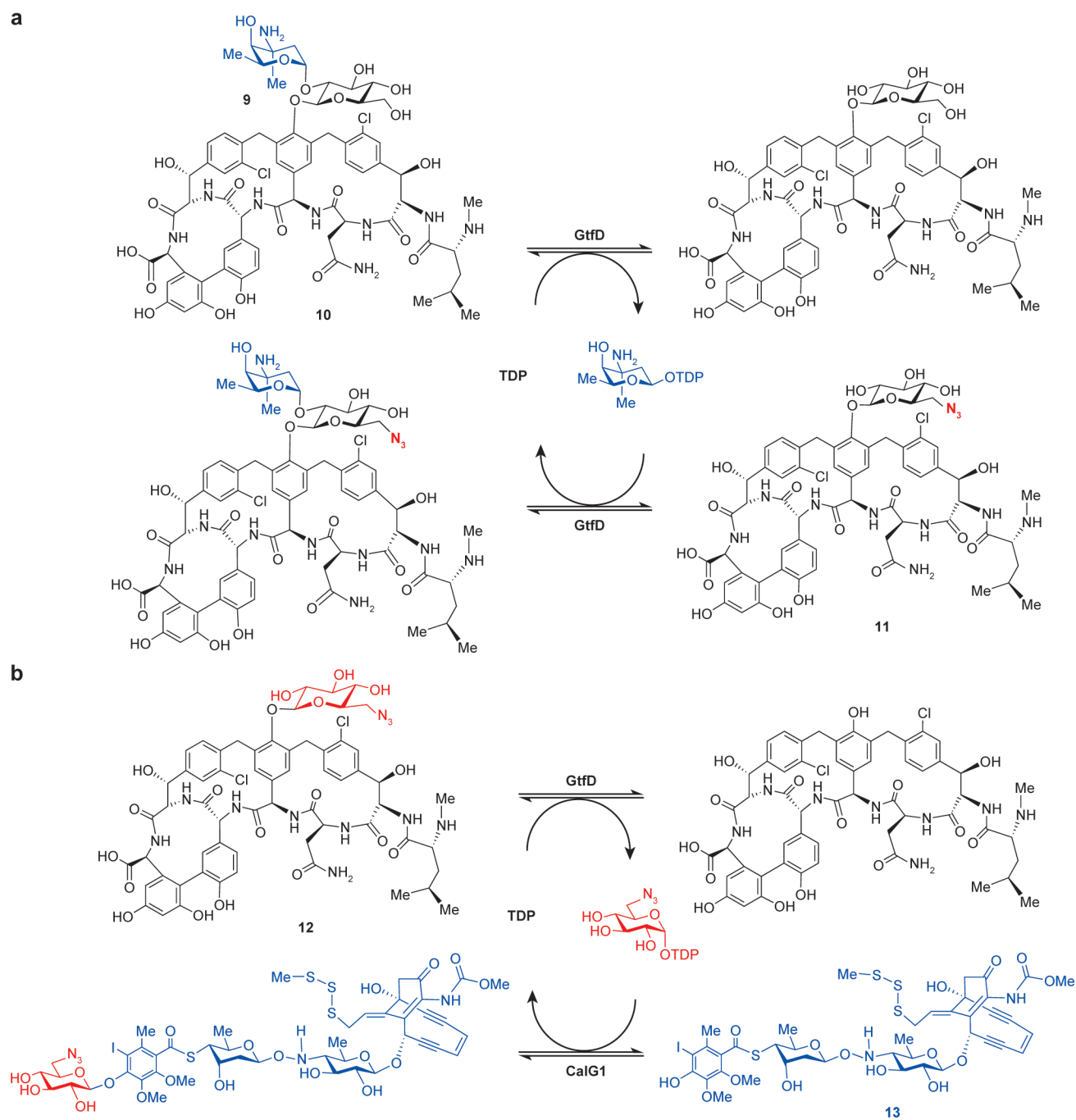


Figure 1. Reverse GT-catalyzed reactions used for glycorandomization of CLM analogues. These reactions are described in ref 11.



Thorson and coworkers have demonstrated the feasibility of several important new applications of glycosyltransferase-catalyzed reactions.

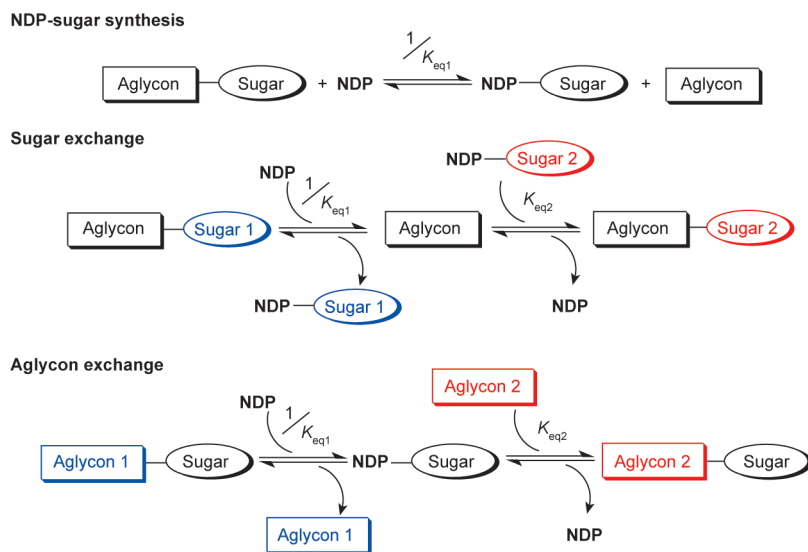


Figure 3. The utility of the reverse glycosylation reaction. Three reaction types were presented by Zhang *et al.* (11) The equilibrium constants (K_{eq}) represent the forward (glycosylation) reactions. Here, as in the text, all equilibrium constants for reverse glycosylation steps are written as $1/K_{eq}$.

exchange, stripping L-vancosamine (9) from vancomycin (10) and coupling it to a non-natural vancomycin derivative (11) (Figure 2, panel a). A final elegant demonstration of the utility of the reversibility of the GT reactions was an experiment in which GtfE and CalG1 were used in a tandem one-pot aglycon exchange reaction in which a non-natural azido sugar (12) was removed from the vancomycin aglycon by GtfE and then transferred to the CLM aglycon (13) by CalG1 (Figure 2, panel b). These experiments demonstrate a convenient way to transfer “exotic” sugars between structurally related and between unrelated natural product scaffolds.

Applications based on the reversibility of GT-catalyzed reactions that have previously been reported for the enzymatic synthesis of uridine 5'-diphosphate-glucose (13) and, more recently, in “aglycon switch” reactions catalyzed by VinC, the GT found in the vicenistatin biosynthesis pathway (14). Also, a plant GT, flavonol 3-O-galactosyltransferase, was found to catalyze the forward and

reverse reactions with K_{eq} values near unity (15). So neither the concept nor the demonstration of the reversibility of GT-catalyzed reactions is novel. Yet, what Thorson and coworkers (11) have done is to significantly expand upon these previous findings, elegantly demonstrating the feasibility of several important new applications of GT-catalyzed reactions, such as one-pot generation of glycodiversified natural product libraries, two-GT-mediated aglycon exchange, GT functional elucidation *via* reverse assay, and facile synthesis of exotic NDP-sugars.

This work provokes two questions: What in the way of reversibility and *in vitro* synthetic utility can be expected from GT-catalyzed reactions involved in biosynthesis of other bioactive natural products, polysaccharides, and glycoproteins? Are all GTs amenable to reverse catalysis? We investigated the four reaction types presented in this report: NDP-sugar synthesis, sugar exchange, one-GT aglycon exchange, and two-GT aglycon exchange (Figure 3),

with kinetic simulations to determine both the advantages and limitations of each reaction type. Importantly, all of the simulations revealed that the values of the equilibrium constants (K_{eq}) of GT-catalyzed reactions are the single most critical factor governing reaction efficiency. While simulating each reaction type, we assumed that K_{eq} values for glycosyltransfer reactions ranged from 4.5 to 159, which are the values reported for the GtfE- (11) and OleD- (16) catalyzed reactions, respectively. The simulations suggest that in each of the reactions (Figure 3), conditions can be manipulated to give the desired product in >90% yield from the limiting substrate. In the one-step NDP-sugar synthesis reaction (Figure 3), the concentration of NDP can be raised to push the equilibrium toward products. Typical outcomes for sugar and aglycon exchange reactions (Figure 3) are illustrated by the sugar exchange simulations presented in Figure 4. In the most favorable scenario (Figure 4, panel c), the sugar removal reaction is reversible (such as in the GtfE-catalyzed reaction in which $1/K_{eq1} = 1/4.5$), whereas the coupling reaction is essentially irreversible ($K_{eq2} > 50$). Here, high percent conversions of the limiting substrate (Figure 4, panel a, A_1S_1) to the desired glycoform (Figure 4, panel a, A_1S_2) can be readily obtained even at low molar equivalents of the variable substrate (Figure 4, panel a, NDP- S_2). When the K_{eq} values of the sugar and aglycon exchange reactions are not conducive to high product yields (such as those in Figure 4, panel d), the concentration of either of the species in the denominator of the $K_{sugar\ exchange}$ expression (Figure 4, panel b) can be raised to improve the product yield. Unlike the NDP-sugar synthesis reactions, only catalytic amounts of NDP are needed for efficient conversion in sugar and aglycon exchange reactions (Figure 4, panel e). Similar results were obtained in the simulations of both one- and two-GT aglycon exchanges. Thus, the simulations suggest that it will be feasible to use a

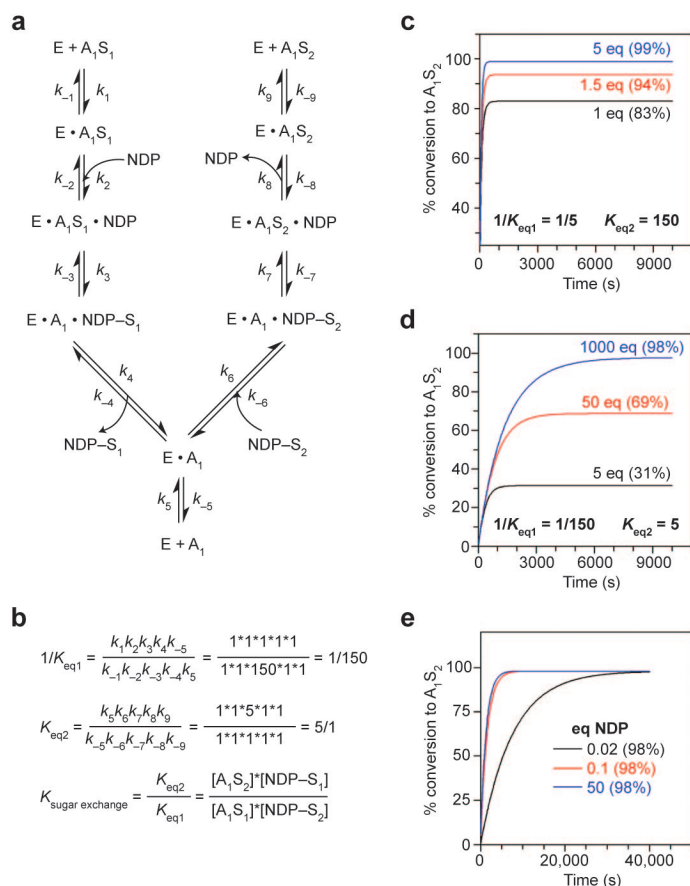


Figure 4. Controlling reaction efficiency in a hypothetical sugar exchange reaction. **a**) The kinetic scheme was simulated with KinTekSim (E = GT; A = aglycon; S = sugar). All of the substrate binding (k_1 , k_2 , k_5 , and k_6) and product release (k_{-4} , k_{-5} , k_8 , and k_9) steps were set to $1 \mu\text{M}^{-1} \text{s}^{-1}$ or 1s^{-1} , respectively, and the K_{eq} values were adjusted by varying the rates of the chemistry steps (k_3 , k_{-3} , k_7 , and k_{-7}). **b**) The set of rate constants used to generate the curves in panel d along with the equilibrium expression for the overall two-step sugar exchange reaction ($K_{sugar\ exchange}$). The GT concentration was set to $12 \mu\text{M}$, and the limiting substrate [A_1S_1] concentration was set to $100 \mu\text{M}$. The equilibrium constants for the glycosylation reactions were assumed to be either 5 or 150 (i.e., k_{-3}/k_3 and $k_7/k_{-7} = 5$ or 150) and either [$NDP-S_2$] or [NDP] was varied. In panels c and d, the molar equivalents of $NDP-S_2$ and the percent conversion obtained are shown next to their respective curves. In panel c, the equilibrium constants are favorable for A_1S_2 formation, and only small amounts of variable substrate (~5 equiv) are needed for nearly complete conversion. In panel d, the equilibrium constants are not favorable for A_1S_2 formation (i.e., $1/K_{eq1} \ll 1$ and $K_{eq2} < 10$), and 100 equiv of $NDP-S_2$ are needed to achieve >95% conversion to A_1S_2 . **e**) When [$NDP-S_2$] is sufficiently high, the [NDP] only affects the rate at which equilibrium is reached and not the percent conversion. Identical trends as those presented in c–e were observed in simulations of aglycon exchange reactions and when simpler kinetic schemes (which did not include substrate binding or product release steps) were used. However, the minimal realistic model shown in panel a can be adapted to account for different rates of substrate binding, product release, and chemistry to better match the available kinetic information on the system being studied. Finally, because the rate constants in panel b used to generate the simulated curves were arbitrarily chosen, actual reactions will reach equilibrium on different time scales.

variety of GTs whose reactions exhibit different degrees of reversibility as catalysts in both NDP-sugar synthesis and sugar and aglycon exchange reactions; Thorson's methodology will thus be applicable to numerous systems.

This work also brings to the forefront the dearth of detailed kinetic and mechanistic studies on GTs and highlights the need for this type of information in order for scientists to take full advantage of the tremendous opportunities this new technology offers. As more GTs from diverse sources are characterized, quantitative data on reaction reversibility and specificity will prove invaluable in assessing their synthetic potential. Structural analysis of GTs will also facilitate engineering to enhance their substrate flexibility, potentially allowing custom design of GT catalysts. However, the effective demonstration and exploitation of the reversibility of GT-catalyzed reactions presented in this work have clearly opened up many new avenues for exploration. Whether for the elucidation of GT function or the enzymatic synthesis of antibiotics or polysaccharides of defined structure, these methods will likely enjoy widespread use in future glycoscience research.

Acknowledgment: This work is supported in part by the National Institutes of Health Grants GM35906 and 40541.

REFERENCES

- Kren, V., Martinkova, L. (2001) Glycosides in medicine: The role of glycosidic residue in biological activity, *Curr. Med. Chem.* 8, 1303–1328.
- Weymouth-Wilson, A. C. (1997) The role of carbohydrates in biologically active natural products, *Nat. Prod. Rep.* 14, 99–110.
- Zhang, G., Shen, J., Cheng, H., Zhu, L., Fang, L., Luo, S., Muller, M. T., Lee, G. E., Wei, L., Du, Y., Sun, D., Wang, P. G. (2005) Synthesis and biological activities of rebeccamycin analogs with uncommon sugars, *J. Med. Chem.* 48, 2600–2611.
- Griffith, B. R., Langenhan, J. M., Thorson, J. S. (2005) "Sweetening" natural products via glycorandomization, *Curr. Opin. Biotechnol.* 16, 622–630.
- Melançon, C. E., III, Yu, W.-I., and Liu, H.-w. (2005) TDP-mycaminose biosynthetic pathway revised and conversion of desosamine pathway to mycaminose pathway with one gene, *J. Am. Chem. Soc.* 127, 12240–12241.

6. Blanchard, S., Thorson, J. S. (2006) Enzymatic tools for engineering natural product glycosylation, *Curr. Opin. Chem. Biol.* **10**, 263–271.
7. Salas, J. A., Mendez, C. (2005) Biosynthesis pathways for deoxysugars in antibiotic-producing Actinomycetes: isolation, characterization and generation of novel glycosylated derivatives, *J. Mol. Microbiol. Biotechnol.* **9**, 77–85.
8. Borisova, S. A., Zhang, C., Takahashi, H., Zhang, H., Wong, A. W., Thorson, J. S., Liu, H.-w. (2006) Substrate specificity of the macrolide glycosylating enzyme pair DesVII/DesVIII: opportunities, limitations, and mechanistic hypotheses, *Angew. Chem., Int. Ed.* **45**, 2748–2753.
9. Oberthur, M., Leimkuhler, C., Kruger, R. G., Lu, W., Walsh, C. T., Kahne, D. (2005) A systematic investigation of the synthetic utility of glycopeptide glycosyltransferases, *J. Am. Chem. Soc.* **127**, 10747–10752.
10. Ruprath, C., Schumacher, T., Elling, L. (2005) Nucleotide deoxysugars: essential tools for the glycosylation engineering of novel bioactive compounds, *Curr. Med. Chem.* **12**, 1637–1675.
11. Zhang, C., Griffith, B. R., Fu, Q., Alberman, C., Fu, X., Lee, I.-K., Li, L., Thorson, J. S. (2006) Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions, *Science* **313**, 1291–1294.
12. Losey, H. C., Peczu, M. W., Chen, Z., Eggert, U. S., Dong, S. D., Pelczar, I., Kahne, D., Walsh, C. T. (2001) Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides, *Biochemistry* **40**, 4745–4755.
13. Cardini, C. E., Leloir, L. F., Chiriboga, J. (1955) The biosynthesis of sucrose, *J. Biol. Chem.* **214**, 148–155.
14. Minami, A., Kakinuma, K., Eguchi, T. (2005) Aglycon switch approach toward unnatural glycosides from natural glycoside with glycosyltransferase VinC, *Tetrahedron Lett.* **46**, 6187–6190.
15. Miller, K. D., Guyon, V., Evans, J. N. S., Shuttleworth, W. A., Taylor, L. P. (1999) Purification, cloning, and heterologous expression of a catalytically efficient flavonol 3-O-galactosyltransferase expressed in the male gametophyte of *Petunia hybrida*, *J. Biol. Chem.* **274**, 34011–34019.
16. Quiros, L. M., Carbajo, A. F., Brana, A. F., Salas, J. A. (2000) Glycosylation of macrolide antibiotics. Purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*, *J. Biol. Chem.* **275**, 11713–11720.

Clearing the Skies over Modular Polyketide Synthases

David H. Sherman^{†,*,S,*} and Janet L. Smith^{†,*}

[†]Departments of Medicinal Chemistry, ^{*}Microbiology and Immunology, ^SChemistry, and [†]Biological Chemistry, Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109

Macrolide antibiotics are among the most effective and successful natural product anti-infective agents and are featured in ongoing efforts toward the development of new therapeutics that target drug-resistant pathogens. Erythromycin was the first macrolide to be introduced into human use, and its structure continues to be a crucial template for production of semisynthetic antibiotics, particularly the new ketolide anti-infective agents introduced recently into the clinic (1).

The erythromycin biosynthetic pathway (2, 3) has also been a key model system for understanding the intricate series of steps involved in assembly of the 14-membered macrolactone ring system and the glycoside appendages that together form the macrolide class of antibiotic agents. Assembly of the core macrolactone of erythromycin is prescribed by a series of multifunctional enzymes called modular polyketide synthases (PKSs), which have been studied genetically and biochemically over the past 15 yr. The frontier in new efforts to understand and engineer these fascinating megaenzymes requires detailed atomic-level structures of the individual catalytic domains and domain partners that comprise the modular PKS system. Recently, Tang *et al.* (4) provided penetrating new insights into the workings of this biochemical assembly line in reporting the first crystal structure of a modular PKS ketoacyl synthase-AT di-domain. Together with a recent crystal structure of a modular PKS

ketoreductase (KR) domain (5) and crystallographic analyses of mammalian and fungal fatty acid synthases (6, 7), a unified and unexpected picture is emerging that offers a refined understanding of the versatile capabilities of these extraordinary biological catalysts.

Since the first studies of polyketide biosynthesis a century ago (8), chemists have recognized the striking parallels between mechanisms involving assembly of this hugely diverse family of bioactive natural products and their cousins, the more chemically modest, but essential, fatty acids (9). Both are constructed by linking a short-chain acyl-coenzyme A (CoA) starter unit with several additional short-chain acyl-CoA extender units to form a more complex product. For both polyketides and fatty acids, the center of this process is the set of carbon-carbon bond-forming reactions that link the acyl subunits, with both systems depending on an acyl carrier protein (ACP) to mediate the reaction sequences for chain elongation and subsequent keto group processing. In terms of structural outcome, the similarities between polyketide metabolites and fatty acids stop here. The products of the fatty acid biosynthetic machinery (fatty acid synthase or FAS) are limited with respect to chain length, involve a small number of biosynthetic subunits or appended functional groups, and usually result in a metabolic product devoid of stereochemistry. By contrast, enzymes involved in construction of complex polyketides (the

ABSTRACT Modular polyketide synthases (PKSs) are large multifunctional proteins that synthesize complex polyketide metabolites in microbial cells. A series of recent studies confirm the close protein structural relationship between catalytic domains in the type I mammalian fatty acid synthase (FAS) and the basic synthase unit of the modular PKS. They also establish a remarkable similarity in the overall organization of the type I FAS and the PKS module. This information provides important new conclusions about catalytic domain architecture, function, and molecular recognition that are essential for future efforts to engineer useful polyketide metabolites with valuable biological activities.

*Corresponding authors,
davidhs@umich.edu,
janetsmith@umich.edu.

Published online September 15, 2006

10.1021/cb600376r CCC: \$33.50

© 2006 by American Chemical Society

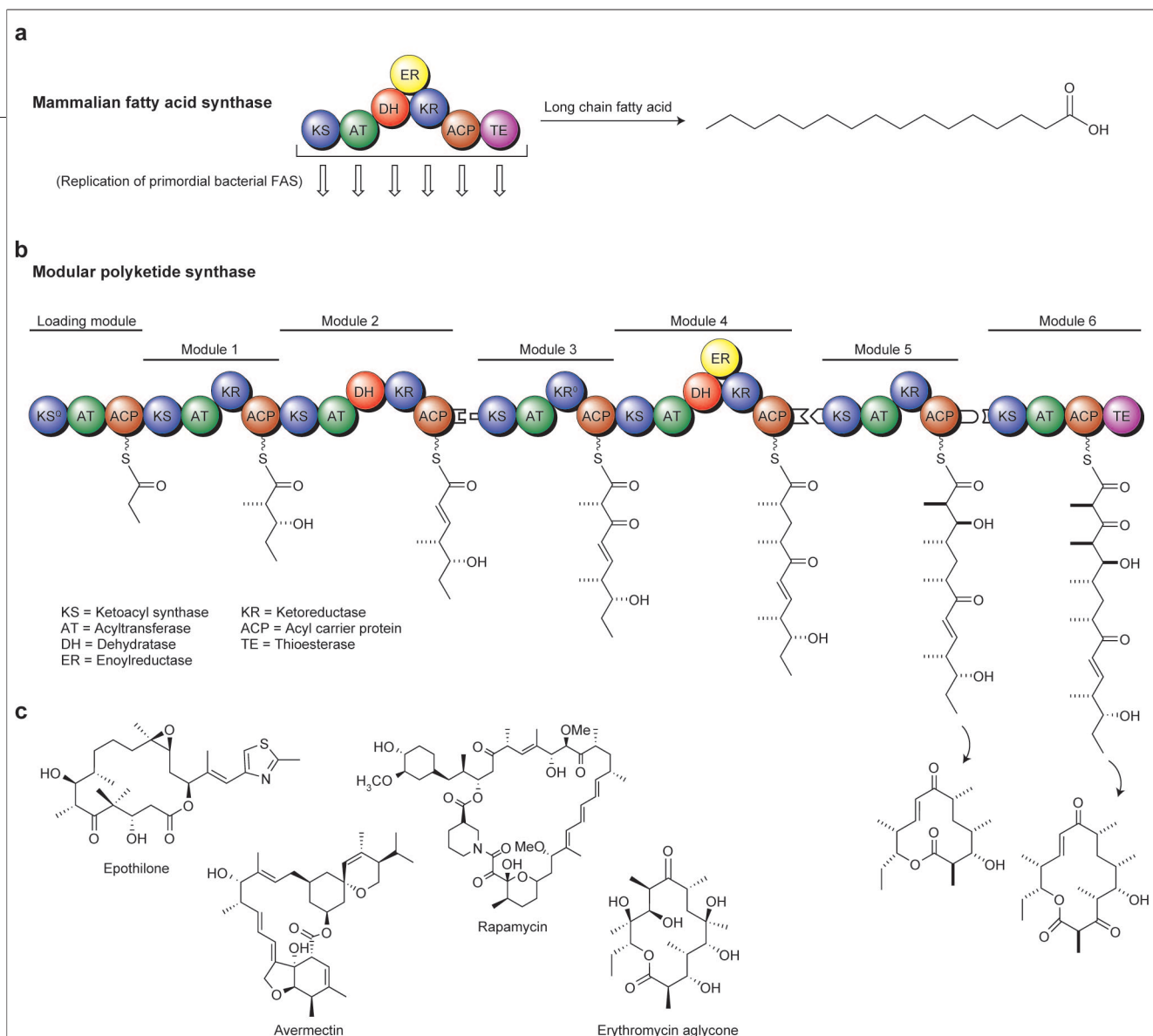


Figure 1. Comparison of type I FAS and modular PKS. **a)** Mammalian FAS functions iteratively to generate a long-chain saturated fatty acid chain. **b)** Bacterial modular PKS, derived from primordial bacterial type I FAS (that in turn evolved into mammalian FAS) is composed of a series of modules bearing specific combinations of catalytic domains that operate sequentially. The metabolic outcome is a complex linear polyketide chain elongation intermediate that is finally cyclized by the terminal TE domain. **c)** Examples of polyketide chemical diversity, including (from left to right) epothilone, avermectin, rapamycin, and erythromycin aglycone.

modular PKSs) are renowned for their ability to create chemical diversity involving starter and extender units, regio- and stereochemistry, and chain length. An elaborate “programming” process that resides within the basic synthase unit, or PKS module, enables this diversity.

A fundamental enigma exists when comparing the basic enzymatic subunit of a type I FAS with a fully equipped type I PKS module (Figure 1). Both are huge multifunctional proteins composed of an equal set of

catalytic domains, including a keto acyl synthase (KS), acyltransferase (AT), dehydratase (DH), enoylreductase (ER), KR, ACP, and terminal thioesterase (TE). In the case of FAS, these catalytic domains exist on a single multifunctional enzyme that operates iteratively, where the growing fatty acid chain cycles through the prescribed set of catalytic reactions until the proper chain length is achieved and is finally off-loaded by the terminal TE. In the modular PKSs, this same series of seven catalytic domains (KS-AT-KR-

DH-ER-ACP-(TE)) participates as part of a larger composite complex involving many modules that include sequential chain extensions and keto group processing reactions, where the growing acyl chain is passed from one module to another until it is off-loaded and (typically) cyclized as a macrolactone by the terminal TE (10). From this more complex arrangement, polyketide metabolites can display an almost limitless chemical diversity. The PKS module and the type I FAS have a clear evolutionary connec-

tion, apparent both in the similar sequences of their catalytic domains and in the order of these domains in the multifunctional proteins. However, the wide distribution of PKSs in the microbial world and the extreme chemical diversity of their products are the result of considerable divergence from their common FAS ancestor (Figure 1). Nature has removed catalytic domains from many modules, rendered domains inactive in others, rearranged domains in some, and occasionally imported non-FAS catalytic domains (*e.g.*, the curacin (11) and jamaicamide (12) PKSs). The linking sequences in PKS modules and FAS appear dissimilar, in contrast to the sequences of the catalytic domains that they connect. In light of such divergence, what relationship, if any, remains in the overall structures of the PKS module and the FAS?

Over the past few months, several crystal structures have been reported that focus on distinct aspects of these multifunctional proteins. Three seemingly disparate contributions (4–6) provide a new, unifying view of the FAS and PKS systems (Figure 2). In reality, the great divergence of PKS modules has clouded the fundamental similarity in overall architecture of the PKS module and its type I FAS ancestor. The nearly simultaneous publication of these three new crystal structures clears the clouds, sheds a brilliant light on their structure and mechanisms, provides directions for additional studies, and reminds us once again that protein structure is far more conserved than primary sequence.

First to appear was the structure of the porcine FAS from Ban and coworkers (6). Crystallization of this 540-kDa dimeric megaenzyme was a major achievement. Although the resolution of the electron density map was too low (4.5 Å) for detailed fitting of the amino acid sequence, previously reported structures of five enzymes (KS, AT, DH, ER, KR) could be fit to the map with a high degree of confidence. The overall structure has upper and lower halves

connected by a narrow “waist” (Figure 2). The KS and AT domains reside in the lower half of the protein, and the DH, ER, and KR domains reside in the upper half. The fundamentally dimeric KS, DH, and ER domains reside along the dimer axis, and the monomeric AT and KR domains are at the periphery. Even after all the catalytic domains had been placed, some regions of electron density were unaccounted for. These were ascribed to long linking sequences (between KS and AT domains, between DH and ER domains) and to the terminal TE.

One month later, Keatinge-Clay and Stroud (5) reported the structure of the KR domain from module 1 of the erythromycin PKS. Crystals were derived from a tryptic fragment that included a large linker region preceding the assigned KR domain, as well as nearly 100 residues following it. The structure contained a huge surprise; the entire region forms a single protein structure consisting of an amino-terminal “structural” domain intimately associated with the previously assigned catalytic domain at the carboxy terminus. It does not appear that the KR catalytic domain can exist in the absence of its structural domain partner. Keatinge-Clay and Stroud then examined sequences of many PKS modules and found a KR structural domain upstream of nearly all assigned KR catalytic domains. The KR structural domain is less well-conserved than the KR catalytic domain but is easily recognized by an amino-terminal peptide that makes critical contacts with the catalytic domain. The newly assigned KR structural domain is nearly 200 amino acids in length, and thus, a large region of blank “linker” sequence can now be assigned across the family of type I PKS modules. One of the most surprising findings is that, in PKS modules containing both ER and KR domains (erythromycin PKS module 1 lacks an ER domain), the ER domain is between KR structural and catalytic domains in the polypeptide sequence.

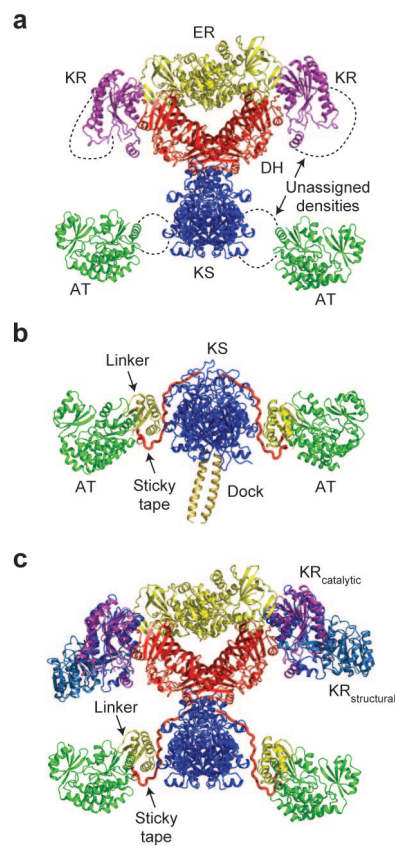


Figure 2. FAS and PKS multifunctional protein structures. **a)** Crystal structure of porcine FAS (6, PDB code 2CF2). Individual domains are colored separately and linked sequentially as shown in Figure 1 (KS-AT-DH-ER-KR). Dotted lines indicate the locations of unassigned electron density in the FAS map. **b)** Crystal structure of the KS-AT di-domain from module 5 of the erythromycin PKS (4, PDB code 2HG4). The sequence of domains along the polypeptide is N-dock-KS-link-AT-“sticky-tape”. The overall position and orientation of KS (blue) and AT (green) domains are strikingly similar in the PKS module and in the type I FAS shown in panel a. **c)** FAS structure with missing domains modeled from PKS structures. The KS-AT linker (yellow) and the “sticky-tape” (red) are from the structure of the KS-AT di-domain structure shown in panel b. The FAS KR (magenta) is overlaid with the KR catalytic domain (dark blue) from module 1 of the erythromycin PKS (5, PDB code 2FR1); the partner PKS KR structural domain is shown in light blue.

The emerging model of catalytic domains is also consistent with the emerging model of a common architecture for PKS modules and the type I FAS.

Each of the FAS and KR crystal structures provides important clues about missing regions of the other. One of the regions of unassigned electron density in the FAS map is adjacent to the KR catalytic domain and seems likely to be a KR structural domain (5), given the superposition of the PKS and FAS KR catalytic domains (Figure 2, panel c). Indeed, the FAS sequence includes a large unassigned “linker” region upstream of the ER domain, in an analogous position to the newly assigned KR structural domain of PKS modules. If the type I FAS and the PKS module have a common organization of KR structural and catalytic domains, then the dimeric organization of DH and ER domains is also likely to be common.

The connection of PKS and FAS architecture is even stronger in the third new crystal structure, by Khosla and colleagues (4), of the KS-AT region from module 5 of the erythromycin PKS, the first view of a modular PKS di-domain. The di-domain is an extended dimer with KS domains at the center joined to peripheral AT domains by a linker domain of ~100 amino acids. The structure also includes 30 amino acids of the linker sequence following the AT domain. This extended C-terminal peptide acts as a crucial “sticky-tape” in the overall structure, making extensive hydrophobic contacts as it wraps around the linker and KS domains ending near the dimer axis. One of the most exciting aspects of this structure is its striking similarity to the lower half of the type I FAS structure in overall size and shape of the dimer and in the positions of the KS and AT domains (Figure 2). These structural similarities, as well as sequence similarities in the KS-AT linker domains and in the sticky-tape peptides of PKS modules and type I FASs, imply a similar architecture for the two multifunctional proteins (4). In accord with this idea, unassigned electron density in the FAS map lies between KS and AT domains in exactly the same position as the linker domain in the PKS. Likewise, the sticky-tape peptide, mapped onto the FAS structure,

would end exactly at the waist of the structure where the lower half (KS-AT) connects to the upper half (DH-ER-KR-ACP-TE). Thus, it appears that PKS modules may also have DH-ER-KR-ACP-TE upper and KS-AT lower halves joined at a waist.

With the significant new insights gained in these recent studies, what structural or functional questions remain to be clarified in the type I FAS and PKS systems? Moreover, what opportunities for practical applications exist given the new information gained from the combined information of these crystal structures? Now that the close structural and functional similarities between type I mammalian FAS and bacterial modular PKSs are so evident, a mystery remains. What is the basis for the protein biochemical metamorphosis of a primordial FAS (that operates as an iterative system to generate an 18-carbon saturated fatty acid) to a multimodular PKS (where an individual catalytic domain functions only once in a sequential series of steps to generate a complex polyketide molecule)? The most salient unique element of modular PKSs is the N-terminal and C-terminal docking domains (Figure 3) that mediate specific module–module interactions. This feature enables cognate pairs of PKS proteins to transfer sequentially an individual linear acyl chain elongation intermediate from one module to another in a process called channeling (Figure 1, panel b). The molecular details of this fascinating process remain obscure, but again, the new structures provide a coherent view. The PKS di-domain structure includes a portion of the N-terminal docking domain, which is its “ticket” to molecular recognition and association with the preceding PKS module. The coiled-coil helices at the dimer axis are consistent with an NMR structure of the isolated docking domains (13), demonstrating that the basis for successful docking is formation of a four-helix bundle with the C-terminus of the previous module. The emerging model of docking domains is also consistent with

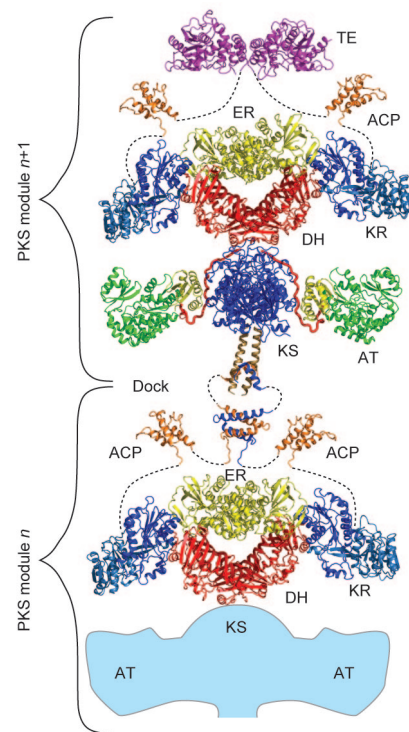


Figure 3. Model of docked PKS modules. A hypothetical PKS module $n+1$ is assembled from the PKS and FAS structures shown in Figure 2, with catalytic domains colored as in Figure 1. The terminal module $n+1$ also includes an ACP domain from the solution structure of the actinorhodin PKS ACP (18, PDB code 2AF8) and a terminal TE domain from module 6 of the pikromycin PKS (15, PDB code 1HFK). The N-terminal docking domain (brown) of module $n+1$ associates with the C-terminal docking domain (gold and blue) of module n , as seen in the solution structure of fused docking domains from modules 2 and 3 of the erythromycin PKS (13, PDB codes 1PZQ and 1PZR). Flexible connections between domains are shown as dotted lines. The lower half of module n is shown in outline form.

the emerging model of a common architecture for PKS modules and the type I FAS. The docking domain coiled-coil extends from the di-domain on the opposite side from the sticky-tape linker peptide. This corresponds to the bottom of the lower half of an intact module. The complementary docking

domain at the C-terminus of the preceding module would presumably extend from the ACP domains of its upper half (Figure 3).

The remarkable insights provided by these new structural studies were accomplished with native proteins in the absence of natural substrates or substrate mimics. Further understanding, particularly details about modular PKS protein–protein interactions, substrate channeling, dynamics, and catalysis, will require cocrystal structures with appropriate molecular probes or affinity labels, in conjunction with protein NMR analysis. Several new cocrystal structures of the pikromycin PKS terminal TE domain involving affinity labels based on natural substrates have recently demonstrated the power of this approach (14, 15).

Ultimately, understanding the basis for molecular recognition and substrate channeling in modular PKSs will be of wide-ranging practical importance for scientists working with these metabolic systems. Although progress has been made toward empirical engineering of PKSs that are functional and able to generate new compounds, these examples are limited to two or three modules (specifically, heterologous modules or catalytic domains derived from phylogenetically related bacteria) that generate tri- or tetraketide metabolites in modest yields (16). When these approaches enable *de novo* design of polyketide pathways that are versatile, are scalable, and provide efficient access to structurally diverse products, the ability to create new natural product chemical entities will go sky high (17).

Acknowledgment: We thank David L. Akey for assistance in building the structural models displayed in Figures 2 and 3. Work on modular PKSs in the authors' laboratories is supported by National Institutes of Health grants GM076477 (D.H.S.) and DK042303 (J.L.S.).

REFERENCES

1. Nguyen, M., and Chung, E. P. (2005) Telithromycin: the first ketolide antimicrobial, *Clin. Ther.* 27, 1144–1163.
2. Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*, *Nature* 348, 176–178.
3. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) Modular organization of genes required for complex polyketide biosynthesis, *Science* 252, 675–679.
4. Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase, *Proc. Natl. Acad. Sci. U.S.A.* 103, 11124–11129.
5. Keatinge-Clay, A. T., and Stroud, R. M. (2006) The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases, *Structure* 14, 737–748.
6. Maier, T., Jenni, S., and Ban, N. (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution, *Science* 311, 1258–1262.
7. Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006) Architecture of a fungal fatty acid synthase at 5 Å resolution, *Science* 311, 1263–1267.
8. Collie, J. (1907) Derivatives of the multiple keten group, *J. Chem. Soc.* 91, 1806–1813.
9. Hopwood, D. A., and Sherman, D. H. (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis, *Annu. Rev. Genet.* 24, 37–66.
10. Aldrich, C. C., Venkatraman, L., Sherman, D. H., and Fecik, R. A. (2005) Chemoenzymatic synthesis of the polyketide macrolactone 10-deoxymethynolide, *J. Am. Chem. Soc.* 127, 8910–8911.
11. Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., Sherman, D. H., and Gerwick, W. H. (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*, *J. Nat. Prod.* 67, 1356–1367.
12. Edwards, D. J., Marquez, B. L., Nogle, L. M., McPhail, K., Goeger, D. E., Roberts, M. A., and Gerwick, W. H. (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*, *Chem. Biol.* 11, 817–833.
13. Broadhurst, R. W., Nieltispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003) The structure of docking domains in modular polyketide synthases, *Chem. Biol.* 10, 723–731.
14. Giraldes, J. W., Akey, D. L., Kittendorf, J. D., Sherman, D. H., Smith, J. L., and Fecik, R. A. (2006) Structural and mechanistic insights into polyketide macrolactonization from polyketide-based affinity labels, *Nat. Chem. Biol.* 2, 531–536.
15. Akey, D. L., Kittendorf, J. D., Giraldes, J. W., Fecik, R. A., Sherman, D. H., and Smith, J. L. (2006) Structural basis for macrolactonization by the pikromycin thioesterase, *Nat. Chem. Biol.* 2, 537–542.
16. Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., Hopwood, D. A., and Santi, D. V. (2005) Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes, *Nat. Biotechnol.* 23, 1171–1176.
17. Gonzalez-Lergier, J., Broadbelt, L. J., and Hatzimaniatis, V. (2005) Theoretical considerations and computational analysis of the complexity in polyketide synthesis pathways, *J. Am. Chem. Soc.* 127, 9930–9938.
18. Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A., and Simpson, T. J. (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from *Streptomyces coelicolor* A3(2), *Biochemistry* 36, 6000–6008.

Single-Insect NMR: A New Tool To Probe Chemical Biodiversity

Aaron T. Dossey[†], Spencer S. Walse[‡], James R. Rocca[§], and Arthur S. Edison^{†,§,¶,*}

[†]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, Florida 32610-0245, [‡]Center for Medical, Agricultural and Veterinary Entomology, USDA-ARS, Gainesville, Florida 32604, [§]McKnight Brain Institute, University of Florida, Gainesville, Florida 32610, and [¶]National High Magnetic Field Laboratory, University of Florida, Gainesville, Florida 32610

Individual organisms often produce natural products in very small quantities (1). Accordingly, their isolation and identification traditionally require large amounts of starting material and a significant effort in sample preparation. Analytical techniques such as mass spectrometry (MS), capillary electrophoresis, and fluorescence spectroscopy are now extremely sensitive and are being used to expedite this process. The use of NMR on the other hand, has lagged behind due to large sample requirements. Although notoriously insensitive, NMR is indispensable to natural product identification because it provides structural information that is not accessible with other techniques. Microcoil (2–6) and cryogenic (7) technology for NMR probes has significantly reduced sample mass requirements and enhanced several natural product studies (3, 8–12). We recently combined the advantages of small-diameter samples with cryogenic technology in a 1-mm-diam NMR probe made from high-temperature superconducting (HTS) material to achieve ~25× greater sensitivity than a conventional probe (13). Here we have used this novel probe to characterize the defensive secretions of individual walking stick insects.

Anisomorpha buprestoides (order Phasmatodea) is common in the southeastern U.S. and is often found in pairs with the smaller male riding on the back of the female (14). When threatened, it accurately

sprays a secretion at predators (14, 15). Following the extraction of >1000 *A. buprestoides* “milkings” into methylene chloride, Eisner and Meinwald (15) identified its active component as a cyclopentanoid monoterpene dialdehyde that they named anisomorpal. At about the same time, Cavill and Hinterberger (16) identified a similar compound in ants that they named dolichodial. Anisomorpal had lower optical activity than dolichodial, suggesting that *A. buprestoides* secretions contained a mixture of isomers or an optically active impurity (15). Subsequently, two related stereoisomers were identified from a plant in the mint family, *Teucrium marum* (cat thyme) (17–19). The minor isomer from *T. marum* was assigned to anisomorpal (17). For clarity, we will refer to any of the stereoisomers with the covalent structure as “dolichodial-like” (Scheme 1); we will suggest specific assignments at the end of this work.

Without purification or additional preparation, we were able to collect the 1D ¹H NMR spectrum (Figure 1, panel a) within ~10 min following the milking of a single midsized *A. buprestoides* male. The spectrum was more complicated than expected for a compound with only 10 carbon atoms, so we extracted the sample with an equal volume of deuterated chloroform (CDCl₃) and collected ¹H NMR data on the respective aqueous (Figure 1, panel b) and organic (Figure 1, panel c) fractions.

ABSTRACT Because of analytical limitations, multiple animals or plants are typically required to identify natural products. Using a unique 1-mm high-temperature superconducting NMR probe, we directly examined the chemical composition of defensive secretions from walking stick insects. Individual milkings were dissolved in D₂O without purification and examined by NMR within 10 min of secretion. We found that *Anisomorpha buprestoides* secretes similar quantities of glucose and mixtures of monoterpene dialdehydes that are stereoisomers of dolichodial. Different individual animals produce different stereoisomeric mixtures, the ratio of which varies between individual animals raised in the same container and fed the same food. Another walking stick, *Peruphasma schultzei*, also secretes glucose and a single, unique stereoisomer that we are naming “peruphasmal”. These observations suggest a previously unrecognized significance of aqueous components in walking stick defensive sprays. Single-insect variability of venom demonstrates the potential importance of chemical biodiversity at the level of individual animals.

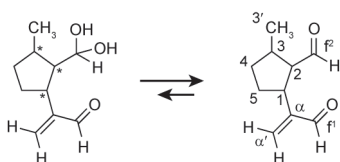
*Corresponding author,
art@mbi.ufl.edu.

Received for review July 27, 2006
and accepted August 14, 2006.

Published online September 15, 2006
10.1021/cb600318u CCC: \$33.50

© 2006 by American Chemical Society

We confirmed that the aqueous fraction contains essentially pure glucose by adding 0.9 μL of 50 mM D-glucose with 0.11 mM 3-(trimethylsilyl) propionate-2,2,3,3- d_4 (TSP) in D_2O to a similarly prepared sample. Only the peaks corresponding to those in the aqueous fraction increased in intensity (Figure 1, panel e), no additional resonances were detected, and the resonances observed within the aqueous fraction were identical with those of aqueous D-glucose (Figure 1, panel f). HPLC/MS of aqueous fractions supplemented with $^{13}\text{C}_6$ D-glucose also supports this conclusion (Supplementary Figure 1). Using HPLC and colorimetric (20) assays, we estimate that an *A. buprestoides* secretion contains between 140 and



Scheme 1. Aqueous equilibrium between the dialdehyde (right) and diol forms (left) of dolichodial-like structures. Chiral carbons are identified by asterisks. The numbering scheme is according to Chemical Abstracts Service.

280 mM glucose. By NMR, we find roughly equal amounts of glucose and dolichodial-like isomers (Figure 1), but the exact ratio varies between animals. We are unaware of any previous reports of glucose in phasmid insect secretions.

In order to assign the NMR resonances (Supplementary Table 1), 2D datasets were recorded from a single walking stick milking (Figure 2). We were able to collect high-quality COSY, TOCSY, ROESY, and natural abundance ^{13}C HMQC and HMBC datasets in the time typically used for conventional 600 μL samples (Supplementary Figure 2). From the 1D and 2D NMR data on *A. buprestoides*, we identified two major dolichodial-like isomers with the corresponding diols that are expected in aqueous solution

(Scheme 1). Each major isomer could be fully assigned through ^1H - ^1H and/or ^1H - ^{13}C scalar coupling correlations. Diols were recognized by the disappearance of the formyl² (Scheme 1) aldehyde proton in water and were verified by extracting the sample into CDCl_3 , which essentially eliminates the diol. We estimate that in water the diols are about 14% of the concentration of the dialdehydes based on integration of NMR peaks in the aldehyde and vinyl regions. Using gas chromatography (GC) with mass spectrometry detection (GC/MS) we also identified two major isomers as well as a minor isomer (Figure 3, panel c and Supplementary Figure 3).

The isomeric heterogeneity of dolichodial-like isomers led us to examine the composition of single milkings from different individual *A. buprestoides* as a function of time. We separated four half-grown males from our culture into their own containers, collected a sample from each, and analyzed them by 1D ^1H NMR. We similarly collected and analyzed milkings from the same four animals 2 and 8 d later. An expansion of the vinyl region of the NMR spectrum of each milking for each animal is shown (Figure 3). The chemical shifts of the vinyl protons are different for each isomer and thus provide a direct indication of the heterogeneity of the samples. To our surprise, different individual *A. buprestoides* raised under identical conditions produce different mixtures of dolichodial-like isomers. Furthermore, the composition of the isomeric mixture changed with time for some individuals.

Peruphasma schultei, a recently described walking stick from Peru, also produces a defensive secretion (21). We obtained a pooled sample of three *P. schultei* milkings and found by NMR that it also contains glucose but only one dolichodial-like isomer and corresponding diol; it was distinct from either of the two major isomers found in *A. buprestoides* based on a comparison of NMR chemical shifts (Figure 2 and Figure 3) and by GC (Figure 3, panel c).

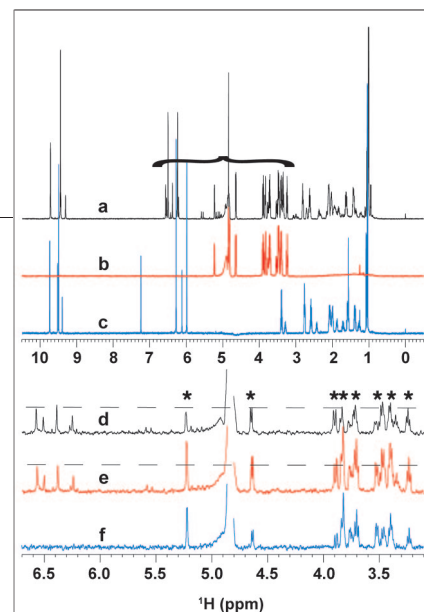


Figure 1. One-dimensional ^1H NMR spectra of single *A. buprestoides* milkings. All spectra were collected at 600 MHz using a 1-mm HTS probe, and sample temperatures were 27 $^\circ\text{C}$. Each spectrum was collected with eight scans. a) About 1 μL was collected from a single insect on a glass pipet tip and added to 10 μL of D_2O containing 0.11 mM TSP. The sample was loaded into a 1-mm capillary NMR tube without purification, and the spectrum was obtained within ~ 10 min of the sample collection. Sample a was extracted with 15 μL of chloroform- d_3 , and the aqueous b) and organic c) fractions were collected and recorded. d) Expansion of a second sample that includes the aqueous component and the vinyl organic region of the spectrum. e) 0.9 μL of pure 50 mM D-glucose was added to sample d. The region of the expansions in d-f is indicated by a bracket in spectrum a. The horizontal dashed lines in spectra d and e indicate the constant vinyl peak intensities, and the asterisks indicate peaks that increased in intensity. f) NMR spectrum of pure glucose.

To compare and name the different dolichodial-like isomers identified in this study, we performed GC/MS on chloroform extracts of walking stick secretions and *T. marum*, reported previously to produce dolichodial and a small amount of anisomorphal (17). The two *T. marum* isomers are consistent with the two major *A. buprestoides* isomers (Figure 3, panel c), and on the basis of assignments of Pagnoni and coworkers (17), we are assigning the *A. buprestoides* isomers at 11.95 and 12.15 min to “dolichodial” and “anisomorphal”, respectively (Figure 3, panel c). These results are in apparent disagreement with more recent studies suggesting that *A. buprestoides* pro-

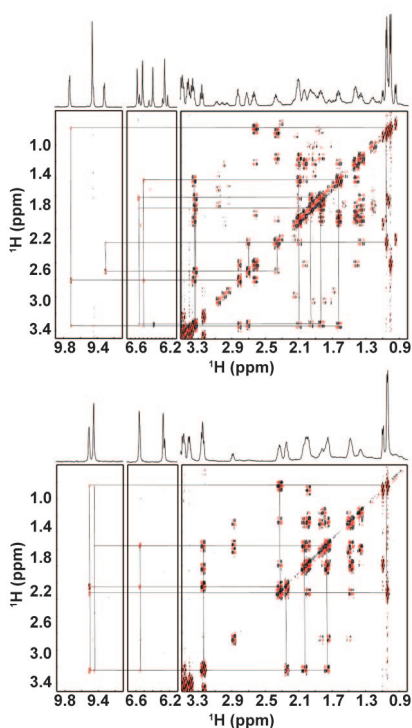


Figure 2. Two-dimensional expansions of COSY (right panels) and ROESY (left and center panels) from a single milking of *A. buprestoides* (top) and a pooled sample from three *P. schultei* milkings (bottom). One-dimensional ^1H spectra from the same samples are shown along the top. All data were collected at 600 MHz using the 1-mm HTS probe. The COSY experiments were collected in ~ 2.5 h with 8 scans and 512 complex indirect data points. The ROESY experiments were collected in ~ 9 h with 32 scans, 512 complex points, and a 400-ms mixing time.

duces a single dolichodial-like isomer (22). This could be due to improvement of analytical methods, genetic variability, or environmental factors. The *P. schultei* and minor *A. buprestoides* isomers at 11.78 min (Figure 3, panel c) appear to be the same and, we believe, are previously unreported. We are naming this isomer “peruphasmal”.

Previous studies using MS, electrophoresis, or LC have reported individual variation in polypeptide toxins from snakes (23–25),

cone snails (26), and a variety of arthropods (27–30). To our knowledge, it has never before been possible to perform a detailed molecular study of a mixture of natural products from an individual insect using NMR. This new capability provides the possibility of elucidating chemical variation, such as stereochemistry, in greater detail. Three major findings on walking stick defense secretions were enabled by high-sensitivity NMR (13): (i) the heterogeneity of defensive dolichodial-like stereoisomers that varies between *A. buprestoides* individuals and with time, (ii) a new dolichodial-like isomer called peruphasmal from *P. schultei*, and (iii) the identification of glucose in phasmid secretions. The quantity of glucose suggests a biological or chemical role in walking stick venom that merits further investigation.

METHODS

Insect Rearing and Sample Preparation. Adult *A. buprestoides* were collected at night in Gulf Hammock, FL, during the fall of 2005. Eggs produced by the insects were hatched in captivity. The young phasmids were fed a diet of only variegated *Ligustrum sinense* purchased from a local plant nursery. We were able to collect single milkings from half-grown males consisting of ~ 1 μL of a whitish fluid by gently touching the secretory duct with a glass pipet. To this we added 10 μL of D_2O containing 0.11 mM TSP as a chemical shift reference to the sample.

NMR. NMR experiments were done using a 600-MHz 1-mm triple-resonance HTS cryogenic probe that was developed through collaboration between the University of Florida, the National High Magnetic Field Laboratory (NHMFL), and Bruker Biospin (13). The total sample volume is ~ 8 μL , and each sample was loaded into a 1-mm \times 100-mm capillary NMR tube (Norell, Inc.) using a 10- μL syringe with a fixed 110-mm \times 30-gauge blunt needle. The capillary tube was held in a standard 10-mm spinner using a Bruker MATCH device, and the capillary–MATCH–spinner combination was lowered vertically into the magnet on an air column as usual. The sample temperature was regulated at 27 $^\circ\text{C}$. The spectrometer was a Bruker Avance 600 with Xwin-NMR software, and all other data acquisition was done using standard technology. Two-dimensional datasets were processed using NMRPipe (31) and manually assigned using NMRView (32).

GC–Flame Ionization Detector. A Hewlett-Packard (Palo Alto, CA) 5890 series II gas chromatograph and a flame ionization detector (GC–FID) with nitrogen make-up gas (1.5 mL/min) and helium carrier gas (1.3 mL/min) were used. Cool on-column and splitless injections (1 μL)

were at 40 and 200 $^\circ\text{C}$, respectively; the detector was maintained at 260 $^\circ\text{C}$. The oven program was as follows: isothermal for 5 min, heating from 40 to 200 $^\circ\text{C}$ at 11 $^\circ\text{C}/\text{min}$, isothermal for 10 min, heating from 200 to 250 $^\circ\text{C}$ at 25 $^\circ\text{C}/\text{min}$, and then isothermal for 15 min. GlasSeal connectors (Supleco) fused three silica columns in series: a primary deactivated column (8 cm long, 0.53 mm i.d.), an HP-1MS retention gap column (2 m long, 0.25 mm i.d., $df = 0.25$ μm), and a J&W DB-5 analytical column (30 m long, 0.25 mm i.d., $df = 0.25$ μm).

GC/MS. A Varian 3400 gas chromatograph and a Finnigan MAT Magnum ion trap mass spectrometer in electron impact ionization mode (70 eV) with a filament bias of 11765mV or chemical ionization mode (isobutane) were employed to acquire full-scan spectra over the ranges m/z 40–400 at 0.85 s per scan. Holox (Charlotte, NC) high-purity helium was used as a carrier gas (1.4 mL/min). Injection and oven conditions were as above. Transfer-line and manifold temperatures were 240 and 220 $^\circ\text{C}$, respectively.

Acknowledgments: A sample of defensive secretions from *P. schultei* was kindly provided by O. V. Conle of Bolsterlang, Germany. We thank Drs. W. W. Brey (NHMFL) and R. S. Withers and R. E. Nast (Varian NMR) for the collaboration and support on the 1-mm HTS probe. Dr. P. Teal (USDA Laboratory, Gainesville, FL) provided helpful encouragement and discussions. Supported by NIH P41RR016105, the Human Frontier Science Program (ASE), and the NHMFL. NMR data were collected in the Advanced Magnetic Resonance Imaging and Spectroscopy Facility at McKnight Brain Institute of the University of Florida.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Metcalf, R. L. (1998) Ultramicrochemistry of insect semiochemicals, *Mikrochim. Acta.* 129, 167–180.
2. Olson, D. L., Peck, T. L., Webb, A. G., Magin, R. L., and Sweedler, J. V. (1995) High-resolution microcoil H-1-NMR for mass-limited, nanoliter-volume samples, *Science* 270, 1967–1970.
3. Gronquist, M., Meinwald, J., Eisner, T., and Schroeder, F. C. (2005) Exploring uncharted terrain in Nature’s structure space using capillary NMR spectroscopy: 13 Steroids from 50 fireflies, *J. Am. Chem. Soc.* 127, 10810–10811.
4. Peti, W., Norcross, J., Eldridge, G., and O’Neil-Johnson, M. (2004) Biomolecular NMR using a microcoil NMR probe—new technique for the chemical shift assignment of aromatic side chains in proteins, *J. Am. Chem. Soc.* 126, 5873–5878.
5. Li, Y., Logan, T. M., Edison, A. S., and Webb, A. (2003) Design of small volume HX and triple-resonance probes for improved limits of detection in protein NMR experiments, *J. Magn. Reson.* 164, 128–135.

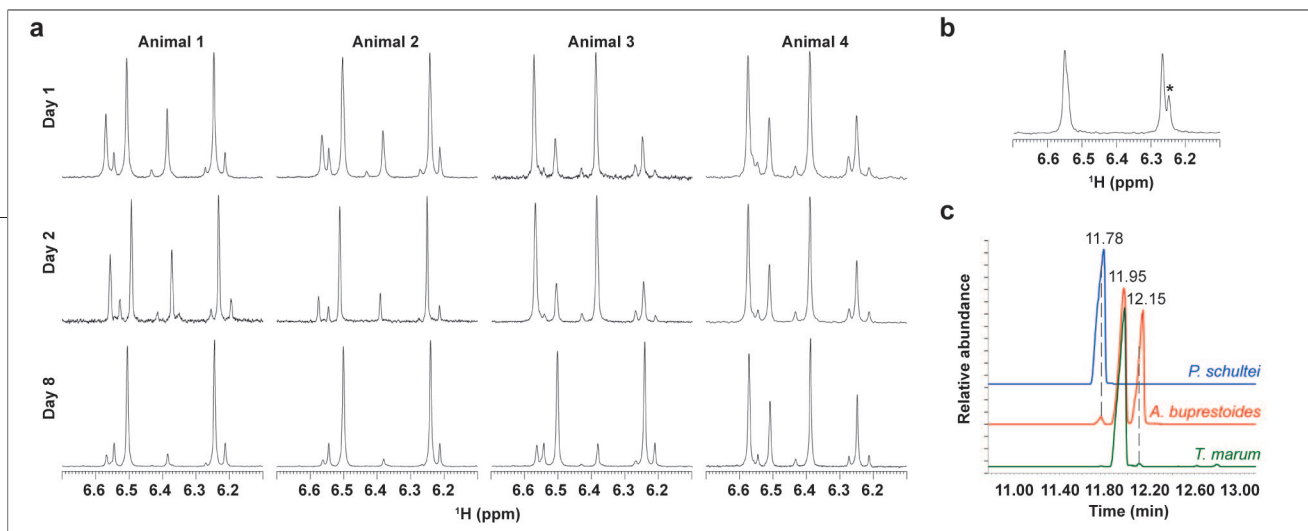


Figure 3. Isomeric variation of venom. **a)** Expansions of the vinyl region of NMR spectra of single milkings from individual *A. burestoides* collected on different days. Samples were dissolved in 10 μ L of D_2O + TSP without further purification. **b)** Same expansion from a mixture of three milkings from *P. schultei*. The shoulder marked with an asterisk corresponds to one of the vinyl peaks of the diol, and the broad peak at 6.54 ppm is an overlap of the diol and dialdehyde isomers (Scheme 1). **c)** GC analysis of chloroform extracts of *A. burestoides* (red) and *P. schultei* (blue) secretions, and *T. marum* (green). With both cool on-column (shown) and splitless injection, all isomers ionized with comparable efficiency (FID), fragmented similarly (electron impact MS), and had identical masses of 166 Da (chemical ionization MS).

- Li, Y., Webb, A. G., Saha, S., Brey, W. W., Zachariah, C., and Edison, A. S. (2006) Comparison of the performance of round and rectangular wire in small solenoids for high-field NMR, *Magn. Reson. Chem.* **44**, 255–262.
- Kovacs, H., Moskau, D., and Spraul, M. (2005) Cryogenically cooled probes—a leap in NMR technology, *Prog. Nucl. Magn. Reson. Spectrosc.* **46**, 131–155.
- Rogers, E. W., and Molinski, T. F. (2005) A cytotoxic carotenoid from the marine sponge *Prianos osiros*, *J. Nat. Prod.* **68**, 450–452.
- Wolters, A. M., Jayawickrama, D. A., and Sweedler, J. V. (2005) Comparative analysis of a neurotoxin from *Calliostoma canaliculatum* by on-line capillary isotachopheresis 1H NMR and diffusion 1H NMR, *J. Nat. Prod.* **68**, 162–167.
- McPhail, K. L., France, D., Cornell-Kennon, S., and Gerwick, W. H. (2004) Peyssonnenynes A and B, novel enediyne oxylipins with DNA methyl transferase inhibitory activity from the red marine alga *Peyssonnelia caulifera*, *J. Nat. Prod.* **67**, 1010–1013.
- Russell, D. J., Hadden, C. E., Martin, G. E., Gibson, A. A., Zens, A. P., and Carolan, J. L. (2000) A comparison of inverse-detected heteronuclear NMR performance: conventional vs cryogenic microprobe performance, *J. Nat. Prod.* **63**, 1047–1049.
- Saman, D., Cvacka, J., Svatos, A., Bouman, E. A. P., and Kalinova, B. (2006) Structural identification of an anthrasteroid hydrocarbon from the sheep tick *Ixodes ricinus*. *J. Nat. Prod.* DOI: 10.1021/np0680127.
- Brey, W. W., Edison, A. S., Nast, R. E., Rocca, J. R., Saha, S., and Withers, R. S. (2006) Design, construction, and validation of a 1-mm triple-resonance high-temperature-superconducting probe for NMR, *J. Magn. Reson.* **179**, 290–293.
- Eisner, T. (1965) Defensive spray of a phasmid insect, *Science* **148**, 966–968.
- Meinwald, J., Chadha, M. S., Hurst, J. J., and Eisner, T. (1962) Defense mechanisms of arthropods. 9. Anisomorphal, the secretion of a phasmid insect, *Tetrahedron Lett.* 29–33.
- Cavill, G. W., and Hinterberger, H. (1961) Chemistry of ants. 5. Structure and reactions of dolichodial, *Aust. J. Chem.* **14**, 143–149.
- Pagnoni, U. M., Pinetti, A., Trave, R., and Garanti, L. (1976) Monoterpenes of teucrium-marum, *Aust. J. Chem.* **29**, 1375–1381.
- Eisner, T., Eisner, M., Aneshansley, D. J., Wu, C. L., and Meinwald, J. (2000) Chemical defense of the mint plant, *Teucrium marum* (Labiatae), *Chemoecology* **10**, 211–216.
- Ricci, D., Fratemale, D., Giampieri, L., Bucchini, A., Epifano, F., Burini, G., and Curini, M. (2005) Chemical composition, antimicrobial and antioxidant activity of the essential oil of *Teucrium marum* (Lamiaceae), *J. Ethnopharmacol.* **98**, 195–200.
- Hendrix, D. L. (1993) Rapid extraction and analysis of nonstructural carbohydrates in plant-tissues, *Crop Sci.* **33**, 1306–1311.
- Conle, O. V., and Hennemann, F. H. (2005) Studies on neotropical Phasmatodea I: A remarkable new species of Peruphasma Conle & Hennemann, 2002 from northern Peru (Phasmatodea: Pseudophasmatidae: Pseudophasmatinae), *Zootaxa* **1068**, 59–68.
- Eisner, T., Morgan, R. C., Attygalle, A. B., Smedley, S. R., Herath, K. B., and Meinwald, J. (1997) Defensive production of quinoline by a phasmid insect (*Oreophoetes peruana*), *J. Exp. Biol.* **200**, 2493–2500.
- Menezes, M. C., Furtado, M. F., Travaglia-Cardoso, S. R., Camargo, A. C., and Serrano, S. M. (2006) Sex-based individual variation of snake venom proteome among eighteen *Bothrops jararaca* siblings, *Toxicon* **47**, 304–312.
- Francischetti, I. M., Gombarovits, M. E., Valenzuela, J. G., Carlini, C. R., and Guimaraes, J. A. (2000) Intraspecific variation in the venoms of the South American rattlesnake (*Crotalus durissus terrificus*), *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **127**, 23–36.
- Monteiro, R. Q., Yamanouye, N., Carlini, C. R., Guimaraes, J. A., Bon, C., and Zingali, R. B. (1998) Variability of bothrojaracin isoforms and other venom principles in individual jararaca (*Bothrops jararaca*) snakes maintained under seasonally invariant conditions, *Toxicon* **36**, 153–163.
- Jakubowski, J. A., Kelley, W. P., Sweedler, J. V., Gilly, W. F., and Schulz, J. R. (2005) Intraspecific variation of venom injected by fish-hunting Conus snails, *J. Exp. Biol.* **208**, 2873–2883.
- Borges, A., Garcia, C. C., Lugo, E., Alfonso, M. J., Jowers, M. J., and Op den Camp, H. J. (2006) Diversity of long-chain toxins in *Tityus zuliaanus* and *Tityus discrepans* venoms (Scorpiones, Buthidae): molecular, immunological, and mass spectral analyses, *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **142**, 240–252.
- Lai, C. C., and Her, G. R. (2000) Analysis of phospholipase A2 glycosylation patterns from venom of individual bees by capillary electrophoresis/electrospray ionization mass spectrometry using an ion trap mass spectrometer, *Rapid Commun. Mass Spectrom.* **14**, 2012–2018.
- Pimenta, A. M., de Lima, M. E., De Marco Almeida, F., Martin-Eauclaire, M. F., and Bougis, P. E. (2003) Individual variability in *Tityus serrulatus* (Scorpiones, Buthidae) venom elicited by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* **17**, 413–418.
- Escoubas, P., Corzo, G., Whiteley, B. J., Celerier, M. L., and Nakajima, T. (2002) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and high-performance liquid chromatography study of quantitative and qualitative variation in tarantula spider venoms. *Rapid Commun. Mass Spectrom.* **16**, 403–413.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe—a multidimensional spectral processing system based on Unix pipes, *J. Biomol. NMR* **6**, 277–293.
- Johnson, B. A., and Blevins, R. A. (1994) NMR View—a computer-program for the visualization and analysis of NMR data, *J. Biomol. NMR* **4**, 603–614.

Utilizing RNA Aptamers To Probe a Physiologically Important Heme-Regulated Cellular Network

Jacquin C. Niles[†] and Michael A. Marletta^{†,*,§,*}

Departments of [†]Chemistry and ^{*}Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720,

[§]Division of Physical Biosciences, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Recent advances in genomics and proteomics are increasing our understanding of transcriptional and post-transcriptional gene regulation and how various gene products integrate into networks (1–4). Understanding the functional importance of specific proteins in these contexts has been aided by several methods, including targeted gene knockouts/mutant collections, RNA interference in permissive organisms, and yeast two-hybrid studies, used in combination with microarray transcriptional profiling and mass spectrometry (3, 5–10).

While these strategies are facilitating the elucidation of protein function, less attention has been focused on the role biologically important small molecules play in modulating protein networks. It is known that small molecules can play important regulatory roles. For example, the cofactor NAD, in addition to a role in energy homeostasis, binds directly to Sir2 and orthologous histone deacetylases and regulates chromatin silencing, aging in response to caloric restriction, and repression of p53-mediated apoptosis in response to DNA damage (11, 12). The reduced form of this cofactor, NADH, binds the transcription repressor Rex, facilitating high-affinity DNA binding of this protein, and regulates expression of several respiratory genes (13). Heme is also known to interact specifically with heme-responsive motifs on the transcription factor Hap-1, leading to transcriptional activation of genes involved in oxidative respiration (*e.g.*, cytochrome *c*) and the response to oxidative stress (*e.g.*, catalase and flavohemoglobin genes) (14). Similarly, heme binding to the transcription repressor Bach1 derepresses gene expression (15). Given the integration of small molecules into these critical circuits, broadly

ABSTRACT Broadly applicable strategies facilitating direct and selective modulation of the intracellular levels of physiologically important small molecules are essential for dissecting their integral and multiple roles in cellular processes. Therefore, we have been exploring the suitability of RNA aptamers for this purpose. Using the *Escherichia coli* heme biosynthetic pathway as a simple model of a negative feedback regulated process, we show that heme-binding RNA aptamers, developed *in vitro* and expressed intracellularly, induce a heme-dependent growth defect in an *E. coli* heme auxotroph defective in converting δ -aminolevulinic (δ -ALA) acid into downstream products. Relative to a control oligonucleotide, the aptamers also induce δ -ALA accumulation in cells grown under heme-limiting conditions. Increasing the concentration of heme in the media completely reverses both the growth defect and δ -ALA accumulation, except for two aptamers for which reversal is partial. Thus, these aptamers specifically target their cognate ligand *in vivo* and functionally modulate its intracellular concentration, demonstrating that RNA aptamers are useful tools for elucidating the role of heme and possibly other small molecules in regulating cellular networks.

*Corresponding author,
marletta@berkeley.edu.

Received for review June 14, 2006
and accepted August 8, 2006.

Published online September 8, 2006
10.1021/cb6002527 CCC: \$33.50

© 2006 by American Chemical Society

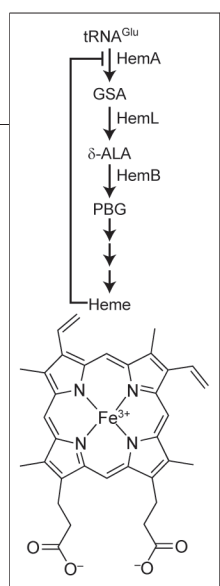


Figure 1. Schematic of the *E. coli* heme biosynthetic pathway. tRNA^{Glu} = glutamyl tRNA; GSA = glutamate-1-semialdehyde; δ-ALA = δ-aminolevulinic acid; PBG = porphobilinogen; HemA = glutamyl-tRNA reductase; HemL = glutamate-1-semialdehyde aminotransferase; HemB = δ-aminolevulinic acid dehydratase.

coli heme biosynthesis as a model for a product feedback inhibited system (Figure 1). Reduction of transfer RNA (tRNA)^{Glu} by glutamyl-tRNA reductase (HemA) is both the rate-determining (28) and regulated step in this pathway. There are at least two mechanisms governing bacterial HemA regulation, namely, (i) direct interaction of heme with HemA, causing a decrease in its enzymatic activity (29, 30), and (ii) heme-dependent proteolytic HemA degradation when intracellular heme is non-limiting (31). In both cases, the net effect is that HemA is highly active and flux through the biosynthetic pathway is elevated when intracellular heme levels are low and conversely when the cell is replete with heme. Here we show that intracellular expression of *in vitro* evolved heme-binding RNA aptamers in *E. coli* predictably modulates this important biofeedback loop.

RESULTS AND DISCUSSION

***In Vitro* Aptamer Selection.** Our first objective was to establish high-affinity, expressible heme-binding RNA aptamers for *in vivo* studies. Single-stranded DNA and 2'-NH₂-modified RNA aptamers targeting other porphyrins, such as *N*-methylmesoporphyrin IX (NMM) and hematoporphyrin IX (HPIX) have previously been reported (32–35). While these aptamers bind heme, they are unsuitable for our purposes, since they cannot be expressed or loaded into *E. coli* at functionally important concentrations. Furthermore, attempts to convert both the DNA and 2'-NH₂ RNA aptamers into expressible 2'-OH aptamers have resulted in significantly decreased binding affinity (35, 36), making this strategy highly inefficient. Thus, a 2'-OH RNA heme-binding aptamer library had to be evolved *de novo* for our experiments.

Aptamers were generated using the *in vitro* SELEX method from a starting library containing 6×10^{13} unique sequences and a 50-nucleotide variable region. The RNA library, body-labeled with [α]-³³P-ATP to facilitate monitoring the selection process, was applied to a column containing immobilized mesoporphyrin IX (MPIX) to retain aptamers having the ability to bind the protoporphyrin IX (PPIX) scaffold. After the column was washed to remove nonspecifically bound sequences, heme-binding aptamers were specifically eluted with 2.5 mM heme in selection buffer. Library selection was judged complete after eight rounds, when the recovered RNA postselection was the same as in the previous round (Figure 2, panel a). The ability of the round 7 selected library to bind PPIX was qualitatively assessed

applicable strategies that facilitate the systematic elucidation of their roles in these contexts are required to improve our understanding of cellular physiology.

Nucleic acid aptamers have several characteristics that make them attractive tools in achieving this goal. First, aptamers naturally encoded in the 5'-untranslated region of several genes play integral roles in sensing and regulating gene expression in response to the intracellular concentrations of small molecules such as amino acids (16, 17), enzyme cofactors (18), purines (19, 20), and Mg²⁺ (21). Second, aptamers can be developed *in vitro* using the readily accessible systematic evolution of ligands by exponential enrichment (SELEX) procedure (22, 23) and can conceivably be evolved to bind any target, as exemplified by the wide range of small molecule, peptide, and protein aptamers previously described (24, 25). While proteins binding these small molecules may exist, aptamers offer greater flexibility since they can be “tailor-made” to have specifically desired properties (*e.g.*, dissociation constants). Third, RNA aptamers can be expressed intracellularly, and their impact on cellular physiology can be examined using simple phenotypic screens (*e.g.*, growth rate) or global approaches such as microarray, proteomic, and metabolomic analyses. Last, aptamers selectively targeting specific proteins have been used to investigate cell-signaling pathways, as demonstrated by inhibition of the mitogen-activated protein kinase (MAPK) pathway by aptamers targeting extracellular-regulated kinase (ERK1/2) *in vitro* (26) and cytohesin-2 *in vivo* (27), indicating that they are useful tools for probing important protein-based networks.

To evaluate the hypothesis that nucleic acid aptamers can be used to explore the role of small molecules in regulating cellular pathways, we have used *Escherichia*

relative to the round 2 library using fluorescence spectroscopy. Free PPIX, when excited at 400 nm gives an emission spectrum with a maximum at 620 nm, and this remains unchanged upon incubation with round 2 library. However, incubation of the round 7 selected library with PPIX results in a bathochromic shift in the emission maximum to 635 nm along with an increase in the emission intensity, indicative of a binding interaction (Figure 2, panel b).

The rounds 6 and 8 DNA libraries were blunt-end-cloned into a plasmid vector and transformed into *E. coli*, from which individual aptamers for sequencing, binding affinity characterization, and archiving were obtained. Sequence data obtained for two and 28 aptamers from the round 6 and 8 pools, respectively, are summarized in Table 1. Analysis using the MEME algorithm (37) revealed the occurrence of a highly conserved G-rich motif in the majority of sequenced clones (Table 1). Interestingly, **6-5** and several round 8 aptamers do not contain this consensus sequence. These may represent distinct heme-binding classes or RNA molecules with no significant heme affinity that bind well to the solid support in spite of negative selection. For **6-5**, the latter appears to be the case (*vide infra*). The majority of sequenced aptamers, however, contain this G-rich motif, suggesting that it represents a conserved aptamer element that is important in heme binding. The motif selected in our experiments, where heme (FePPIX) is used to specifically elute 2'-OH-RNA aptamers, resembles those obtained during the selection of DNA and 2'-NH₂-RNA aptamers to NMM and HPIX, respectively (32–34). This suggests that the PPIX scaffold is the predominant binding surface interacting with selected aptamers and that the electronic or steric changes upon metal coordination (heme), pyrrole N-methylation (NMM), and hydration of the vinyl side chains (HPIX) do not significantly alter this interaction. The observation that these aptamers bind interchangeably with the different PPIX derivatives supports this conclusion.

Next, the heme-binding affinity of selected aptamers was determined using UV–vis spectroscopy monitored heme titrations. Upon binding heme, aptamers induce a shift in the Soret maximum from 396 nm for free heme to ~405–408 nm, with slight hyperchromicity. By titrating heme into aptamer containing solutions and monitoring by difference UV–vis spectroscopy, we determined apparent dissociation constants (K_d) for the binding interaction, and these are summarized in

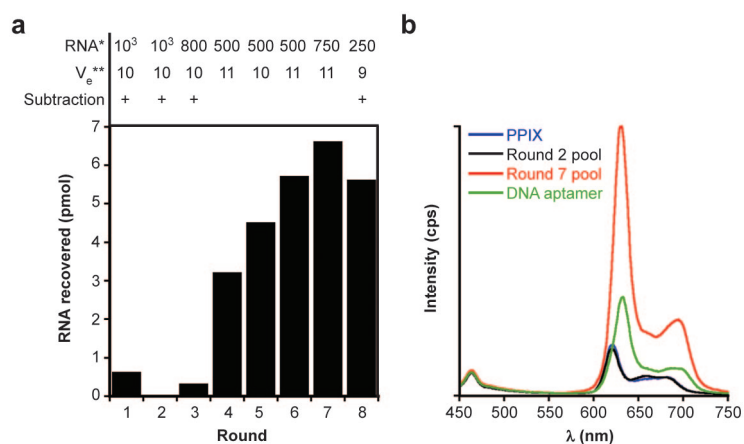


Figure 2. Monitoring of the aptamer selection progress. a) The *in vitro* selection protocol and progress monitored by ³³P-labeled RNA recovery at the end of each round; * indicates the amount of RNA (pmols) applied to the subtraction or selection column; ** indicates that V_e is the number of column volumes of selection buffer used during the washing step. b) Fluorescence emission spectra for PPIX only, PPIX + round 2 pool (negative control), PPIX + round 7 pool, and PPIX + PS2.M (positive control = DNA heme-binding aptamer reported in ref 25). The shift in emission maximum from 600 nm for free PPIX to 623 nm is indicative of a binding interaction between round 7 pool and PPIX.

Table 1. For the tested aptamers, the apparent K_d 's ranged between 190 and 450 nM, indicating a reasonably high-affinity interaction. A dissociation constant for **6-5** could not be determined, since the UV–vis spectrum for **6-5** with heme was indistinguishable from that of free heme (Figure 3). Co-incubating equimolar concentrations of **6-5** and **8-13**, which binds heme and induces a shift in the heme Soret maximum, with one heme equivalent produces a spectrum practically identical to that of heme and **8-13** alone (data not shown). Indeed, if the heme-binding affinity of **6-5** and **8-13** were comparable, a Soret shift toward the **6-5**/heme maximum would be expected. Thus, **6-5** does not contain the consensus G-rich motif and does not appear to have significant heme-binding affinity.

The selected G-rich motif, which may form G-quartets (34), is directly involved in heme binding, as demonstrated by the ability of the chemically synthesized G-rich elements from **8-8** and **8-34** to also bind heme (data not shown), and this finding is consistent with earlier reports (33, 34). Interestingly, the apparent dissociation constants for the motifs from **8-8** and **8-34** are 1.8 and 4.5 μ M, respectively, indicating that these constructs bind heme less tightly than the full-length aptamers by ~4–10-fold. This suggests that unidentified cis- or

TABLE 1. The randomized region of sequenced aptamers



Aptamer ^a	Screened ^b	Sequence ^c	K _d (nM) ^d
6-3	Yes	UGGUUUCAGCGACAGGAGGGGUGUAGGUGGAAUUGCUGUCCUUUGCGUGU	188
6-5	Yes	CUUAUGCAGUUUUACAGGGGUGAUUAGACUGCACCAGUAGGGGGAUGUUC	ND
8-1	Yes	UGCUGAGUUUGUAUGCACGUGGAGGAGGAGGCGUACACUUGCUUUGUGGU (2)	309
8-2		UGCUGAGUAUUGUAUGCACGUGGAGGAGGAGGCAGAAAAGCGCUUAGGGUGUUUG	
8-3		UGCAGAAGUUGCUGUGUGGAGGAGGUGGCAUGACUGCUUAGGGUAGUUG	
8-5		GGGAUCUGGUUGAAACUGGAGGCCUAUAGAAGUUGGUUGUGUUUUAG	
8-6		GUCCUGUGAUGGUUUUUUCGUUCCGCUUACUUCGACAUGAGGCCCGGAUCCAUCUGAAU	
8-7	Yes	GAUUGACCGUAUGGAGGAUGCAAAGGGAGGGAGGUCACUUGAGUUAGUUA	256
8-8	Yes	GCAGGAUGUGGAGGAGGCAUCUGCUGCAAUCGGGACUUGUGUCGAGUAUC (4)	
8-12	Yes	GCAUUGUCUGCGUGUGGAGGCAGGAGGCAAGAUAGAGGUGAUGCGGUUG	220
8-13	Yes	CAUGUUGGCGAUACGUCUAAAACGGUGGGUUGUGGAGGAAUGAUUUUAUCG	371
8-19	Yes	AGUAGUGUCAGCGUGUGGUGGAGGUUGGCGACAUUAGUAGGGUGCGAUUG	445
8-20		CGAAGGCACUUCAUUGGGGUGGAGGAGGCAUCGAGGUGUCCGGCGAGUGG	
8-26		CACACGUGACUGUGGAGGCAGCGGAGGCGAGUUUUGUGAUGUUAAGAGGU	
8-27		UAGGGUGAUUUGUUCUAGAGAUGGCAUGAAA	
8-28		UAUGUUAAAGAGGCCACUGAUGCGCGUAGGUCUCUGGGGAUUGAGGAAGGU	
8-30		AGGUUUCGCUAGGUGAGGAAGGAGGUGUAGGUACGGCCUUAUUGAGUGGGA	
8-31		CGUAGUCCAUGAGUGUCUUUAGCUAACGGUUGGUAGUGAACCAUAUCCUG	
8-32		GCCAAUGAGAGCUGUAGGAGGGCGGGACGUGCUUAGUGCGUGACACCGGA	
8-33		UUGUCCUGACUUGCUUGAACGCUUAGCGUGAUGCGUUUAGCCUUGGAUGGG	
8-34		CACCAUAGACGGGGUUUAAGACGGAGGGAGAUUGCAUCGGUGUGAAGCUGA	
8-35	Yes	UGCGCAAUACACGGUGAGGAGGUGGAGAGAUUGAGGUGCUUAGCAGUUGA	425
8-37		CGUGAACGCAUGUGGUGGAGGAGGCGAUUGCAGGUGGGACCGAGCAUUG	
8-39		GAUGUANCGGUGUCUUAGCCUUGUGGGANUAGGGUGCGUAUGGGGAUGNC	
8-40		UGGACCGCAGCACGGCGUCUGUGGUAAGGCCGUAUGCCCAUCGAAUGAAG	

^aAptamers with demonstrated PPIX scaffold binding are italicized, and δ -ALA levels were determined for cells expressing aptamers shown in boldface.

^bAptamers screened in the *E. coli* growth assay are indicated. ^cThe conserved motif identified using MEME is displayed as a sequence logo (generated online at weblogo.cbr.nrc.ca/logo.cgi). G-rich regions within aptamer sequences are shown in italic boldface; two poly-G regions in the non-heme-binding control oligonucleotide **6-5** that do not conform to the G-rich consensus are underlined. The number of clones harboring a given sequence is shown in parentheses. ^dThe heme dissociation constants in select cases. ND = not detected.

trans-acting structural elements in the full-length aptamer may be stabilizing heme binding to the G-rich motif.

Impact of Heme-Binding Aptamers on *E. coli* Growth.

We hypothesized that intracellular aptamer expression will lead to heme sequestration and impair aerobic bacterial growth. The heme-permeable *E. coli* heme auxotroph RP523 was selected for these studies for two

main reasons. First, prototrophic *E. coli* strains can significantly upregulate heme biosynthesis (38, 39), decreasing the likelihood that subtle growth phenotypes due to intracellular heme sequestration will be detectable. Second, growth of the auxotroph is strictly dependent upon heme added to the media, and this can be directly controlled experimentally. For testing in *E. coli*,

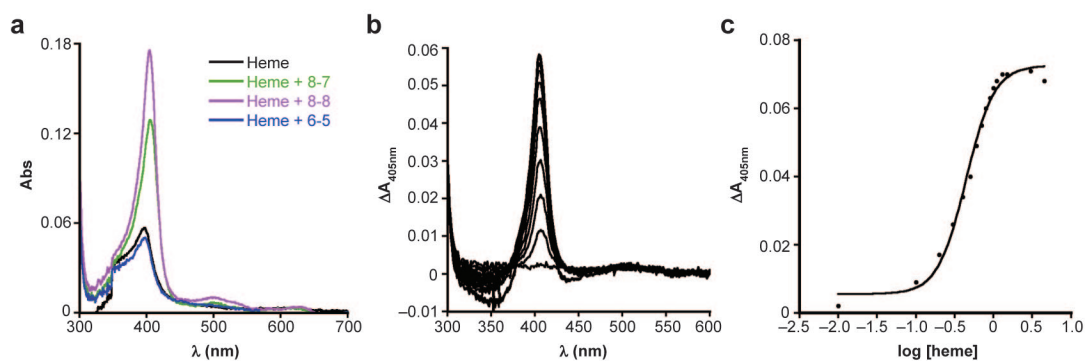


Figure 3. UV-vis spectral properties and titration curve for selected aptamers. **a)** UV-vis spectra for heme alone and with 6-5, 8-7, and 8-8. **b)** Representative difference spectra obtained during heme titrations with aptamers. **c)** Representative plot of $\Delta A_{405\text{nm}}$ versus $\log[\text{heme}]$ fit to eq 1 used to determine dissociation constants for heme binding to aptamers.

several characterized aptamers were cloned into the RNA expression vector pGFIB, which is a high-copy plasmid in which transcription is initiated by the strong and constitutively active *llp* promoter and terminated by the efficient *rmC* terminator (40). From this plasmid, tRNA^{Ala}, which is similar in size to our aptamers, is synthesized in high abundance, with levels $\sim 70 \times$ that of chromosomally encoded tRNA^{Ala} (41).

RP523 cells were transformed with several pGFIB-aptamer constructs, and production of full-length aptamer was ascertained by chain reaction (RT-PCR) for several aptamers (data not shown). To screen for an aptamer-induced growth defect, cells were grown aerobically overnight in 1 mL cultures supplemented with 0.5, 1, 2, and 4 μM heme added to the media. The extent of bacterial growth was determined by OD₆₀₀ readings and normalized to RP523 harboring the non-heme-binding construct pGFIB-6-5. Overall, final bacterial density attained increases directly with heme concentration, as expected. For RP523 not harboring plasmid, the EC₅₀ for maximal growth is $\sim 1.5 \mu\text{M}$, and heme concentrations $\geq 4 \mu\text{M}$ are no longer growth limiting (data not shown). These characteristics are also true for RP523 harboring pGFIB-6-5, although there is a slight decrease in the final bacterial densities achieved at a given heme concentration. The results of screening eight selected aptamers are summarized (Figure 4, panel a) and illustrate that RP523 harboring **8-12**, **8-13**, **8-19**, and **8-35** show dramatically less growth relative to **6-5** at 0.5–1 μM heme, while **6-3**, **8-1**, and **8-7** show a trend toward decreased growth at 0.5 μM heme, but this is not statistically significant. Growth of RP523 express-

ing **8-8** is not impaired relative to **6-5** at all of the tested heme concentrations. Thus, the aptamers are differentially effective at impairing bacterial growth. Importantly, the growth defect induced is fully reversed by increasing media heme concentration for cells expressing **6-3**, **8-1**, **8-7**, **8-13**, and **8-35** and partially so for **8-12** and **8-19** expressing cells. These data are consistent with the hypothesis that the aptamers are inhibiting growth by sequestering intracellular heme, and this effect can be completely or partially overcome by increasing the bioavailable heme concentration.

Next, growth curves for RP523 expressing the selected aptamers **8-1**, **8-7**, **8-12**, **8-13**, and **8-35** and control oligonucleotide **6-5** were conducted at limiting (1 μM) and saturating (10 μM) heme. At 1 μM heme, bacteria expressing aptamers grew more slowly and reached final optical densities $\sim 50\%$ less than those of the control culture (Figure 4, panel b). At 10 μM heme, the growth curves for RP523 expressing **8-1**, **8-7**, **8-13**, and **8-35** and the control oligonucleotide **6-5** were identical. Similar to the screening studies, this defect is not completely reversed for RP523 expressing **8-12**, even though the growth rate and final OD₆₀₀ increased with heme concentration. Overall, these data are consistent with the screening results and indicate that intracellular heme sequestration by aptamers slows bacterial growth and limits the total number of cell doublings. At nonlimiting heme concentrations, the heme-sequestration capacity of the aptamers is overcome, and sufficient heme is available to support growth comparable to that of control cultures.

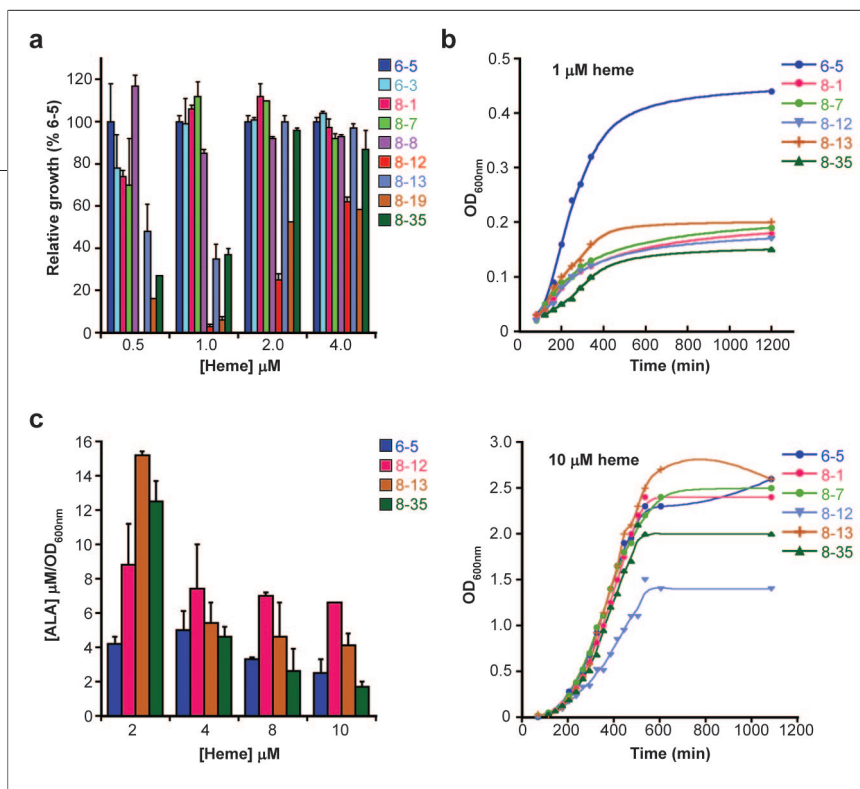


Figure 4. Impact of intracellularly expressed aptamers on RP523 growth and δ -ALA production as a function of heme concentration. **a)** Final cell density measurements obtained for small-scale cultures containing initial heme concentrations of 0.5, 1, 2, and 4 μ M. **b)** Representative growth curves for selected aptamers at limiting (1 μ M; top panel) and nonlimiting (10 μ M; bottom panel) heme. **c)** δ -ALA concentration normalized to OD_{600nm} for cells expressing control oligonucleotide 6-5 and aptamers 8-12, 8-13, and 8-35.

Determination of δ -ALA Levels in RP523 Expressing Aptamers. Heme auxotrophy in *E. coli* RP523 is due to a mutation(s) in the *hemB* locus leading to absent δ -aminolevulinic acid dehydratase (ALAD) activity (42). This enzyme is responsible for converting δ -ALA into porphobilinogen in the heme biosynthetic pathway (Figure 1). Since the *hemA* and *hemL* loci are intact in RP523, this strain synthesizes δ -ALA, and under heme-limiting conditions, HemA activity increases, leading to δ -ALA accumulation. Conversely, when heme is nonlimiting, HemA activity is repressed, leading to decreased δ -ALA levels. Therefore, if aptamers are limiting intracellular heme availability, δ -ALA levels should be higher relative to those in RP523 expressing control oligonucleotide 6-5. To test this, OD_{600} -corrected δ -ALA levels in RP523 grown at 2–10 μ M heme and expressing control oligonucleotide 6-5 and the aptamers 8-12, 8-13, and 8-35 were measured using LC/electrospray ionization (ESI)-MS in positive ions selected ion monitoring (SIM) mode, and these data are summarized (Figure 4, panel c). At 2 μ M heme, δ -ALA levels in aptamer expressing cells were \sim 1.5–4-fold higher than those in cells expressing 6-5, suggesting that at this heme concentration, expressed aptamers are limiting intracellular

heme concentrations. Between 4 and 10 μ M heme, δ -ALA levels among 6-5, 8-13, and 8-35 expressing cells are comparable, indicating that excess heme overcomes the sequestration effect of these aptamers. Interestingly, δ -ALA levels in 8-12 relative to 6-5, 8-13, and 8-35 expressing cells remain elevated, even at nonlimiting heme concentrations.

Altogether, there is a strong concordance between the heme dependence of growth and δ -ALA levels in aptamer-expressing cells. Relative to control cells, a growth defect is observed at low heme concentrations, δ -ALA levels are correspondingly high, and both parameters are normalized at high media heme concentrations. Aptamers are not all equally bioactive. Intracellular aptamer efficacy is a function of at least three variables, namely, (i) target affinity (K_d), (ii) intracellular concentration, and (iii) attaining the properly folded structure required for productive target binding. As

shown in Table 1, the aptamers all bind heme with similar affinity, so this variable cannot explain the different growth phenotypes. Using quantitative RT-PCR, we determined the intracellular levels of aptamers 8-1, 8-7, 8-12, 8-13, and 8-35 relative to control oligonucleotide 6-5 in cells grown at 2 μ M heme. The aptamers and control oligonucleotide are present at similar levels (Figure 5), indicating that dramatic differences in the intracellular concentration of the various aptamers do not account for their differential *in vivo* efficacy. *In vivo* folding efficiency, therefore, is most likely the factor

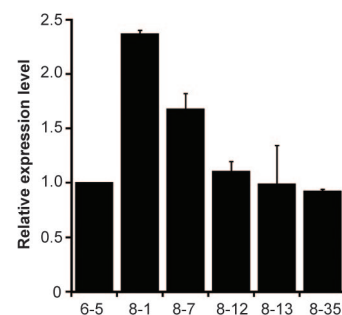


Figure 5. Relative aptamer expression levels as determined by quantitative RT-PCR.

responsible for the differential efficacy of aptamers at sequestering intracellular heme, with the aptamers demonstrating strong heme-dependent phenotypes reflecting those efficiently attaining folded structures competent for heme binding. Misfolding of *in vitro* selected aptamers when expressed intracellularly is a recognized challenge; however, sufficiently diverse libraries can be screened to identify those aptamers possessing the desired intracellular activity and hence the appropriately folded structure *in vivo*. Additionally, aptamers **8-12** and **8-19** are interesting in that the growth defect persists even at higher media heme concentrations. Similarly, in **8-12**-expressing cells, δ -ALA levels remain elevated. This may reflect the possibility that these aptamers fold intracellularly with the highest efficiency, and transport of heme into the cell becomes limiting prior to the aptamer heme-sequestration capacity being surpassed. While we cannot completely exclude additional “off-target” effects of these aptamers, there is a clearly demonstrated “on-target” effect as cells expressing **8-12** accumulate excess δ -ALA. This suggests that intracellular heme limitation remains a reasonable explanation for the persistent growth defect observed in these aptamer-expressing cells, so that even at high heme concentrations in the culture media, the sensed intracellular heme levels may remain low due to sequestration.

Overall, the δ -ALA data are significant for two reasons: First, they validate the hypothesis that the tested aptamers are specifically targeting heme *in vivo* and, therefore, are acting “on-target”. Second, the data indicate that the aptamers can sequester intracellular heme and perturb the heme-mediated inhibition of the

heme biosynthetic pathway predictably and in an objectively measurable way, thus confirming their utility as tools for probing this heme-dependent cellular circuit.

In this study, we have used *in vitro* SELEX to evolve and test the first reported expressible heme-binding RNA library in a cellular system. Selected aptamers, when expressed in an *E. coli* heme auxotroph, induce a growth defect and δ -ALA accumulation at low heme concentrations, and both phenotypes are reversed upon increasing heme concentration in the growth media. Together, these data indicate that aptamers are selectively binding their cognate ligand *in vivo*. Furthermore, by restricting the bioavailable intracellular heme pool, with respect to both inhibiting cell growth and modulating heme-dependent repression of the heme biosynthetic pathway, these aptamers impact an important cellular regulatory circuit predictably, thereby validating this approach in efforts to elucidate other heme-regulated networks and possibly other small molecule regulated cellular processes.

This strategy is inherently appealing because it can be used to explore a wide range of biologically important compounds, since the versatility of SELEX will enable discovery of suitable aptamers. While we have used an *E. coli* heme auxotroph to facilitate screening based on a growth phenotype, this is not an absolute requirement. Using aptamers to perturb intracellular pools of physiologically important small molecules in conjunction with global screening methods such as transcriptional profiling microarrays, proteomics, and metabolomics may augment and extend findings based on gene knockout and other available approaches.

METHODS

Preparation of Selection and Subtraction Columns. MPIX (Frontier Sciences) was immobilized on oxirane-activated acrylic beads (Sigma-Aldrich) as previously described (34). Briefly, 3.1 mg (4.4 μ mol) of MPIX was dissolved in 1 mL of DMSO, then diluted to 10 mL with 100 mM KH_2PO_4 , pH 9.5, and added to 1.07 g of activated acrylic beads. This slurry was gently rotated for 2 d in the dark; then the beads were washed extensively with (i) 10 mM KH_2PO_4 , pH 5, (ii) H_2O , (iii) 0.1 M NaOH, (iv) H_2O , (v) 100 mM KH_2PO_4 , pH 8.0, and finally, (vi) 100 mM KH_2PO_4 , pH 8, with 5% β -mercaptoethanol. The beads were gently rotated in this buffer for 2 d, before being washed extensively with H_2O . Subtraction beads were prepared by treating 1 g of activated beads with 100 mM KH_2PO_4 , pH 8, and 5% β -mercaptoethanol as above. Both selection and subtraction beads were stored in 10 mM KH_2PO_4 , pH 7.2, 150 mM NaCl, 0.02% sodium azide buffer at 4 °C in the dark.

In Vitro RNA Transcription Reactions. Oligonucleotides were obtained from Integrated DNA Technologies, Inc. The randomized single-stranded DNA library with sequence GCC GGA TCC GGG CCT CAT GTC GAA [N]₅₀T TGA GCG TTT ATT CTG AGC TCC C was synthesized on a 1 μ mol scale and polyacrylamide gel electrophoresis-purified prior to use. A 5'-*Hind*III primer (CCG AAG CTT AAT ACG ACT CAC TAT AGG GAG CTC AGA ATA AAC GCT CAA) and 3'-*Bam*HI primer (GCC GGA TCC GGG CCT CAT GTC GAA) were used for PCR amplification of the starting library and RT-PCR during library evolution.

Prior to the first selection round, a single-stranded DNA library (10 \times 100 μ L reactions, 10 pmol DNA per reaction) in 10 mM Tris-HCl, 50 mM KCl, 7.5 mM MgCl_2 , pH 8.3, 1 mM each dNTP, 2 μ M each 5'- and 3'-primers, and *Taq* DNA polymerase (2.5 U) was subjected to six cycles of PCR (pre-PCR, 94 °C \times 4 min, 57 °C \times 5 min; PCR, 94 °C \times 30 s, 57 °C \times 60 s, 72 °C \times 60 s, and a 7 min final extension). The double-stranded product was purified and extracted

from a 4% agarose gel, and 150 pmol was used as template for *in vitro* transcription using the Ampliscribe T7 Flash Kit (Epicentre Biotechnologies) spiked with 1 μL of $\alpha\text{-}^{32}\text{P}$ -ATP (3000 Ci mmol^{-1} , 10 μCi μL^{-1} , Perkin Elmer). Reaction times ranged from 4 h to overnight, at the end of which DNase I (1 U) was added at 37 $^{\circ}\text{C}$ for 30–60 min to digest the template DNA. RNA was purified by phenol–chloroform extraction followed by ethanol precipitation at -20 $^{\circ}\text{C}$.

Aptamer Selection. RNA (200–1000 pmol) in diethyl pyrocarbonate (DEPC)-treated water was denatured by heating to 70 $^{\circ}\text{C}$ for 5 min, allowed to cool to RT, and refolded in selection buffer (SB), with composition 100 mM Tris–acetate, 200 mM sodium acetate, 25 mM potassium acetate, 10 mM magnesium acetate, 0.05% Triton X-100, and 5% DMSO. For subtraction (rounds 1–3 and 8), RNA in ~ 200 μL of SB was added to ~ 100 μL of subtraction resin pre-equilibrated in SB and incubated at ambient temperature with gentle mixing for 30 min. The supernatant and 3×50 μL of SB washes of the subtraction resin were recovered and incubated with selection resin (200 μL) for 1 h at ambient temperature with gentle agitation. The selection resin was washed with 11 column volumes of SB and eluted with 2.5 mM hemin (Sigma-Aldrich) in SB (6×200 μL aliquots with 10 min between additions). The eluted RNA in each fraction was ethanol precipitated overnight at -20 $^{\circ}\text{C}$ with 20 μg of glycogen as a carrier, and amplified using 8 $\times 50$ μL Ready-To-Go RT-PCR tubes (Amersham) and ~ 300 pmol each 5'- and 3'-primers. RT was carried out at 42 $^{\circ}\text{C}$ for 40 min, and the reverse transcriptase was inactivated at 95 $^{\circ}\text{C} \times 5$ min, followed by 18 PCR cycles (94 $^{\circ}\text{C} \times 30$ s, 57 $^{\circ}\text{C} \times 60$ s, 72 $^{\circ}\text{C} \times 60$ s) and a 7 min final extension. The RT-PCR products were pooled, concentrated, and purified using 4% agarose gel, and the desired length product was extracted and ethanol precipitated. DNA was resuspended in DEPC-treated H_2O for the next round of *in vitro* transcription. At the sixth and eighth selection rounds, the evolved library was blunt end cloned into pSTBlue-1 vector (Invitrogen) and used to transform competent NovaBlue *E. coli* cells (Invitrogen). Single colonies were used for mini-prep cultures from which plasmid encoding a single aptamer was isolated for sequencing and archiving. Plasmids were sequenced at either Elim Biopharmaceuticals or the UC Berkeley Sequencing Facility using the SP6 primer.

Determining Aptamer Heme-Binding Properties. Binding of rounds two and seven libraries to the protoporphyrin IX scaffold was qualitatively determined by fluorescence spectroscopy using a FluoroMax-2. Refolded library RNA (~ 2 μM) and PPIX in the more physiologic SHMCK buffer (20 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (Hepes), 120 mM NaCl, 5 mM KCl, 1 mM MgCl_2 , 1 mM CaCl_2 , pH 7) were incubated at ambient temperature, and the fluorescence emission spectrum was measured after excitation at 400 nm. Binding of selected aptamers from the round 8 library to heme was determined by UV–vis spectroscopic difference titrations using a Cary 300 Bio spectrophotometer equipped with a dual cell Peltier accessory (Varian). Briefly, reference and sample cuvettes containing SHMCK only and refolded RNA (~ 1 – 2 μM) in SHMCK were prepared, and heme was added in 0.1–0.2 μM aliquots to each sample and reference cuvette pair at 25 $^{\circ}\text{C}$ while stirring continuously. Difference spectra were recorded every 5 min, and $\Delta A_{405\text{nm}}$ was plotted against $\log[\text{heme}]$ and fitted to eq 1 to determine aptamer apparent heme-binding dissociation constants.

$$\Delta A_{405\text{nm}} = m_1 + (m_2 - m_1) / \{1 + 10^{-(m_3 \log(m_4))}\} \quad (1)$$

The variables are as follows: $m_0 = \log[\text{heme}]$; $m_1 =$ minimum $\Delta A_{405\text{nm}}$; $m_2 =$ maximum $\Delta A_{405\text{nm}}$; $m_3 =$ Hill coefficient; $m_4 =$ apparent K_d .

Cloning of Aptamers into the RNA Expression Vector pGFB. Plasmid pGFB was obtained as a gift from Prof. William McClain (University of Wisconsin, Madison). With primers CCG GAA TTC AAT ACG ACT CAC TAT AGG GAG CTC AGA ATA AAC GCT CAA (5'-*EcoRI*) and GCC CTG CAG GGG CCT CAT GTC GAA (3'-*PstI*), full-length aptamers were PCR amplified from the archival plasmid, purified by

4% agarose gel, extracted, and resuspended in dd H_2O after ethanol precipitation. Aptamers (0.36–0.50 μg of DNA) were double-digested with *EcoRI* (20 U) and *PstI* (20 U) for 4 h at 37 $^{\circ}\text{C}$ in $1 \times$ *EcoRI* Unique Buffer (New England Biolabs), then ligated into pGFB (~ 4.5 μg digested with 20 U of *EcoRI* and 20 U of *PstI* for 5 h at 37 $^{\circ}\text{C}$, then treated with 0.6 U of calf intestinal phosphatase at 37 $^{\circ}\text{C}$ for an additional 30 min) using the Rapid Ligation Kit (Roche). DH-5 α *E. coli* cells were transformed using aliquots of the ligation reactions and grown overnight on Luria-Bertani plates supplemented with 50 μg mL^{-1} carbenicillin. Single colonies were selected for mini-culture and plasmid isolation, and aptamer insertion was verified by sequencing using the M13 Forward primer.

Screening Heme-Binding RNA Aptamers for *in Vivo* Function. pGFB/aptamer constructs were used to transform the *E. coli* heme auxotroph RP523 obtained from the *E. coli* Genetic Stock Center (<http://cgsc.biology.yale.edu/top.html>). Cells were grown overnight on LB plates supplemented with 50 μg mL^{-1} carbenicillin and 15 μM hemin at 37 $^{\circ}\text{C}$ and stored in the dark at 4 $^{\circ}\text{C}$. All growth experiments were done in LB containing 50 μg mL^{-1} carbenicillin and supplemented with the appropriate heme concentration as indicated. For high-throughput screening, overnight cultures (1 mL in 15 mL tubes) containing 0.5, 1, 2, and 4 μM hemin were inoculated with RP523 harboring specific pGFB/aptamer constructs from starter cultures in the early to mid-log phase of growth and incubated at 37 $^{\circ}\text{C}$ and 250 rpm. Optical density measurements (OD_{600}) at ~ 14 – 20 h were used to assess the extent of bacterial growth. For growth kinetics studies, 50 mL cultures in 250 mL Erlenmeyer flasks inoculated with mid-log phase starter cultures (~ 0.5 mL) and supplemented with 1 and 10 μM heme were grown at 37 $^{\circ}\text{C}$ and 250 rpm. OD_{600} readings were taken every 30–60 min to assess growth.

Measurement of δ -ALA Levels in RP523 *E. coli*. Cells were grown in 5 mL of LB containing between 2 and 10 μM heme and harvested by centrifugation after measurement of the OD_{600} . Generally, cells expressing control oligonucleotide **6-5** and aptamers were harvested at similar OD_{600} values, dictated by the maximum OD_{600} attained by aptamer-expressing cells. This ensured that the degree of media heme depletion was similar between control and experimental cultures. Cell pellets were resuspended in 200 μL of 4% heptafluorobutyric acid (HFBA) and lysed by three freeze–thaw cycles followed by sonication for 5 min. Supernatants containing δ -ALA were recovered after centrifugation at 14,000 rpm for 10 min and quantitated using an Agilent 1100 series liquid chromatograph/mass selective detector (LC/MSD) operated in positive ESI mode. A standard curve was constructed using authentic δ -ALA (Sigma-Aldrich) detected by monitoring the $m/z = 114$ $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$ and 132 $[\text{M} + \text{H}]^+$ ions in SIMS mode. For the LC, a 150 mm \times 3.9 mm, 5 μm Nova-Pak C18 column (Waters) with 25 mM HFBA, 5 mM ammonium acetate (solvent A) and 90:10 methanol/5 mM ammonium acetate (solvent B) as mobile phases were used. The column was eluted at a flow rate of 0.4 mL min^{-1} , according to the following gradient: 5% B for 3 min; 5–25% B over 12 min; 25–100% B over 2 min; an isocratic phase at 100% B for 7 min; 100–5% B over 1 min. For the MSD, the drying gas flow rate and temperature were 12 L min^{-1} and 350 $^{\circ}\text{C}$, respectively, and the nebulizer gas pressure was 35 psig. The capillary exit and fragmenter voltages were ~ 3000 and 70 V, respectively.

Quantitative RT-PCR. RP523 bacteria expressing **6-5**, **8-1**, **8-7**, **8-12**, **8-13**, and **8-35** were grown in LB containing 2 μM heme. Cells were harvested at $\text{OD}_{600} = 0.3$ by adding ice-cold 95:5 ethanol/water-saturated phenol to 11% v/v and centrifuging for 2 min at 4 $^{\circ}\text{C}$. Supernatants were aspirated, and cell pellets were frozen in liquid nitrogen, then stored at -80 $^{\circ}\text{C}$ until needed. Total RNA was isolated using TRIzol (Invitrogen) according to the supplier's protocol. Approximately 2 μg of RNA was digested with 5 U of RNase-free DNase I (Fermentas) in 10 mM Tris-HCl, pH 7.5, 2.5 mM MgCl_2 , 0.1 mM CaCl_2 at 37 $^{\circ}\text{C}$ for 1.5 h. DNase I was inactivated at 75 $^{\circ}\text{C}$ for

15 min after adding EDTA to 1 mM. cDNA syntheses with ~200 ng of DNase I-treated RNA were performed using the ThermoScript RT-PCR System (Invitrogen). Target-specific primers for aptamers (GCC GGA TCC GGG CCT CAT GTC GAA) and 16S rRNA (GGT TAC CTT GTT ACG ACT T) as an internal reference were used for cDNA synthesis. Q-PCR reactions were carried out in triplicate per RNA dilution in 50 μ L reactions containing 10 mM Tris-HCl, pH 8.5, 50 mM KCl, 1.5 mM MgCl₂, 200 μ M dNTPs, 0.1 \times SYBR Green I (Molecular Probes), 1.5 U of *Taq* (Fermentas), and 200 nM primers. The primers for amplifying the aptamer library were also used for aptamer quantitation by Q-PCR. For 16S rRNA, primers 16S_462F (GTT AAT ACC TTT GCT CAT TGA) and 16S_801R (ACC AGG GTA TCT AAT CCT GTT) were used. The temperature program used for both cDNA targets was 95 $^{\circ}$ C \times 10 min and 40 cycles of 95 $^{\circ}$ C for 30 s, 57 $^{\circ}$ C \times 1 min, 72 $^{\circ}$ C \times 1 min, followed by a melting curve. Buffer and no RT controls for each RNA sample were included to verify the absence of contaminating DNA in the DNase I-treated RNA samples. Expression levels reported are relative to an aptamer/16S rRNA ratio of 1 assigned to sample 6-5.

Acknowledgments: We thank Prof. William H. McClain (University of Wisconsin, Madison) for providing plasmid pGFIB and Dr. Jasper Rine's lab (University of California, Berkeley) for use of their Q-PCR instrument. This work is supported by NIH Grant 5 F32 AI058646 (J.C.N) and the Aldo DeBenedictis Fund (M.A.M).

Note added after print publication: Because of a production error, the following references were misformatted: 1–42. These errors do not affect the scientific integrity of the article. The electronic version was corrected and reposted to the web on October 20, 2006. This paper was originally posted September 8, 2006, and the electronic version was corrected and reposted to the web on October 20, 2006. An Addition and Correction may be found in *ACS Chem. Biol.* 1(9).

REFERENCES

- Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J., and Palsion, B. O. (2004) Integrating high-throughput and computational data elucidates bacterial networks, *Nature* 429, 92–96.
- Ge, H., Walhout, A. J. M., and Vidal, M. (2003) Integrating 'omic' information: a bridge between genomics and systems biology, *Trends Genet.* 19, 551–560.
- Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network, *Science* 292, 929–934.
- Volker, U., and Hecker, M. (2005) From genomics via proteomics to cellular physiology of the Gram-positive model organism *Bacillus subtilis*, *Cell. Microbiol.* 7, 1077–1085.
- Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference, *Nature* 408, 325–330.
- Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edlmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes, *Nature* 415, 141–147.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L. Y., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W. V., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry, *Nature* 415, 180–183.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proc. Natl. Acad. Sci. U.S.A.* 98, 4569–4574.
- Li, S.M., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalian, P. O., Han, J. D. J., Chesneau, A., Hao, T., Goldberg, D. S. Lin, N. Martinez, M., Rual, J. F., Lamesch, P., Xu L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q. R., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H. Y., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., van den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M. (2004) A map of the interactome network of the metazoan *C. elegans*, *Science* 303, 540–543.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamar, G., Yang, M. J., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*, *Nature* 403, 623–627.
- Guarente, L. (2000) *Sir2* links chromatin silencing, metabolism, and aging, *Genes Dev.* 14, 1021–1026.
- Luo, J. Y., Nikolaev, A. Y., Imai, S., Chen, D. L., Su, F., Shiloh, A., Guarante, L., and Gu, W. (2001) Negative control of p53 by *Sir2 α* promotes cell survival under stress, *Cell* 107, 137–148.
- Brekasis, D., and Paget, M. S. B. (2003) A novel sensor of NADH/NAD(+) redox poise in *Streptomyces coelicolor* A3(2), *EMBO J.* 22, 4856–4865.
- Zhang, L., and Hach, A. (1999) Molecular mechanism of heme signalling in yeast: the transcriptional activator *Hap1* serves as the key mediator, *Cell. Mol. Life Sci.* 56, 415–426.
- Ogawa, K., Sun, J., Taketani, S., Nakajima, O., Nishitani, C., Sassa, S., Hayashi, N., Yamamoto, M., Shibahara, S., Fujita, H., and Igarashi, K. (2001) Heme mediates derepression of *Maf* recognition element through direct binding to transcription repressor *Bach1*, *EMBO J.* 20, 2835–2843.
- Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S., and Breaker, R. R. (2003) An mRNA structure in bacteria that controls gene expression by binding lysine, *Genes Dev.* 17, 2688–2697.
- Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W. L., and Breaker, R. R. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression, *Science* 306, 275–279.
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA, *Chem. Biol.* 9, 1043–1049.
- Mandal, M., and Breaker, R. R. (2004) Adenine riboswitches and gene activation by disruption of a transcription terminator, *Nat. Struct. Mol. Biol.* 11, 29–35.
- Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C., and Breaker, R. R. (2003) Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria, *Cell* 113, 577–586.
- Cromie, M. J., Shi, Y., Latifi, T., and Groisman, E. A. (2006) An RNA sensor for intracellular Mg²⁺, *Cell* 125, 71–84.

22. Ellington, A. D., and Szostak, J. W. (1990) *In Vitro* selection of RNA molecules that bind specific ligands, *Nature* **346**, 818–822.
23. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment – RNA ligands to bacteriophage-T4 DNA-polymerase, *Science* **249**, 505–510.
24. Nimjee, S. M., Rusconi, C. P., and Sullenger, B. A. (2005) Aptamers: An emerging class of therapeutics, *Annu. Rev. Med.* **56**, 555–583.
25. Wilson, D. S., and Szostak, J. W. (1999) *In vitro* selection of functional nucleic acids, *Annu. Rev. Biochem.* **68**, 611–647.
26. Seiwert, S. D., Nahreini, T. S., Aigner, S., Ahn, N. G., and Uhlenbeck, O. C. (2000) RNA aptamers as pathway-specific MAP kinase inhibitors, *Chem. Biol.* **7**, 833–843.
27. Theis, M. G., Knorre, A., Kellersch, B., Moelleken, J., Wieland, F., Kolanus, W., and Famulok, M. (2004) Discriminatory aptamer reveals serum response element transcription regulated by cytohesin-2, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11221–11226.
28. Verderber, E., Lucast, L. J., VanDehy, J. A., Cozart, P., Etter, J. B., and Best, E. A. (1997) Role of the *hema* gene product and delta-aminolevulinic acid in regulation of *Escherichia coli* heme synthesis, *J. Bacteriol.* **179**, 4583–4590.
29. Rieble, S., and Beale, S. I. (1991) Purification of glutamyl-transfer RNA reductase from *Synechocystis Sp Pcc 6803*, *J. Biol. Chem.* **266**, 9740–9745.
30. Srivastava, A., and Beale, S. I. (2005) Glutamyl-tRNA reductase of *Chlorobium vibrioforme* is a dissociable homodimer that contains one tightly bound heme per subunit, *J. Bacteriol.* **187**, 4444–4450.
31. Wang, L. Y., Elliott, M., and Elliott, T. (1999) Conditional stability of the *HemA* protein (glutamyl-tRNA reductase) regulates heme biosynthesis in *Salmonella typhimurium*, *J. Bacteriol.* **181**, 1211–1219.
32. Teramoto, N., Ichinari, H., Kawazoe, N., Imanishi, Y., and Ito, Y. (2001) Peroxidase activity of *in vitro*-selected 2'-amino RNAs, *Bio-technol. Bioeng.* **75**, 463–468.
33. Okazawa, A., Maeda, H., Fukusaki, E., Katakura, Y., and Kobayashi, A. (2000) *In vitro* selection of hematoporphyrin binding DNA aptamers, *Bioorg. Med. Chem. Lett.* **10**, 2653–2656.
34. Li, Y. F., Geyer, C. R., and Sen, D. (1996) Recognition of anionic porphyrins by DNA aptamers, *Biochemistry* **35**, 6911–6922.
35. Kawazoe, N., Teramoto, N., Ichinari, H., Imanishi, Y., and Ito, Y. (2001) *In vitro* selection of nonnatural ribozyme-catalyzing porphyrin metalation, *Biomacromolecules* **2**, 681–686.
36. Travascio, P., Bennet, A. J., Wang, D. Y., and Sen, D. (1999) A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites, *Chem. Biol.* **6**, 779–787.
37. Bailey, T. L., and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers, *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36.
38. Zhang, L., Hach, A., and Wang, C. (1998) Molecular mechanism governing heme signaling in yeast: a higher-order complex mediates heme regulation of the transcriptional activator *HAP1*, *Mol. Cell. Biol.* **18**, 3819–3828.
39. Woodard, S. I., and Dailey, H. A. (1995) Regulation of heme biosynthesis in *Escherichia coli*, *Arch. Biochem. Biophys.* **316**, 110–115.
40. Masson, J. M., and Miller, J. H. (1986) Expression of synthetic suppressor transfer-RNA genes under the control of a synthetic promoter, *Gene* **47**, 179–183.
41. Gabriel, K., and McClain, W. H. (1999) A set of plasmids constitutively producing different RNA levels in *Escherichia coli*, *J. Mol. Biol.* **290**, 385–389.
42. Umanoff, H., Russell, C. S., and Cosloy, S. D. (1988) Availability of porphobilinogen controls appearance of porphobilinogen deaminase activity in *Escherichia coli K12*, *J. Bacteriol.* **170**, 4969–4971.

Design, Synthesis, and Biological Activity of a Potent Smac Mimetic That Sensitizes Cancer Cells to Apoptosis by Antagonizing IAPs

Kerry Zobel^{†,*}, Lan Wang[‡], Eugene Varfolomeev[†], Matthew C. Franklin[†], Linda O. Elliott[‡], Heidi J. A. Wallweber[†], David C. Okawa[‡], John A. Flygare[‡], Domagoj Vucic[†], Wayne J. Fairbrother[†], and Kurt Deshayes^{†,*}

[†]Departments of Protein Engineering and [‡]Medicinal Chemistry, Genentech, Inc., South San Francisco, California 94080

The proper regulation of apoptosis is crucial for development and sustained health (1). Many disease states result from over-sensitivity or resistance to apoptotic stimuli (2). Of special interest is the role of apoptotic resistance in aggressive cancers, since a lack of death response often permits cancers to tolerate conventional treatment (3). The inhibitor of apoptosis (IAP) proteins are key components of the apoptotic cascade (4) and are believed to prevent cell death through interactions between their baculoviral IAP repeat (BIR) domains and caspase-3, -7, or -9, which are critical for the initiation and execution phases of apoptosis (Figure 1). Overexpression of IAP proteins in human cancers has been shown to suppress apoptosis induced by a variety of stimuli (5–7). X-chromosome-linked IAP (XIAP) is a ubiquitously expressed IAP protein and a potent inhibitor of caspases that plays a critical role in resistance to chemotherapeutic agents and other pro-apoptotic stimuli (8–11). Melanoma inhibitor of apoptosis (ML-IAP) is upregulated in a number of melanomas but not expressed in most normal adult tissues (12, 13), and down regulation of ML-IAP by RNA interference leads to induction of apoptosis in tumor cells (14, 15). Cellular Inhibitors of apoptosis 1 and 2 (cIAP1 and cIAP2) are unique among IAP proteins for their ability to interact with tumor necrosis factor receptor-associated factors 1 and 2 (TRAF1 and 2) (16). In addition, cIAP1 and cIAP2 are targets of genetic amplification, which is potentially correlated with resistance to antitumor agents (17–19). Taken together, IAP proteins are attractive targets for anticancer therapeutic intervention (20).

One mechanism by which XIAP inhibits apoptosis is the interaction between its BIR3 domain and caspase-9, which prevents the enzyme from adopting the catalytic

ABSTRACT Designed second mitochondrial activator of caspases (Smac) mimetics based on an accessible [7,5]-bicyclic scaffold bind to and antagonize protein interactions involving the inhibitor of apoptosis (IAP) proteins, X-chromosome-linked IAP (XIAP), melanoma IAP (ML-IAP), and c-IAPs 1 and 2 (cIAP1 and cIAP2). The design rationale is based on a combination of phage-panning data, peptide binding studies, and a survey of potential isosteres. The synthesis of two scaffolds is described. These compounds bind the XIAP-baculoviral IAP repeat 3 (BIR3), cIAP1-BIR3, cIAP2-BIR3, and ML-IAP-BIR domains with submicromolar affinities. The most potent Smac mimetic binds the cIAP1-BIR3 and ML-IAP-BIR domains with a K_i of 50 nM. The X-ray crystal structure of this compound bound to an ML-IAP/XIAP chimeric BIR domain protein is compared with that of a complex with a phage-derived tetrapeptide, AVPW. The structures show that these compounds bind to the Smac-binding site on ML-IAP with identical hydrogen-bonding patterns and similar hydrophobic interactions. Consistent with the structural data, coimmunoprecipitation experiments demonstrate that the compounds can effectively block Smac interactions with ML-IAP. The compounds are further demonstrated to activate caspase-3 and -7, to reduce cell viability in assays using MDA-MB-231 breast cancer cells and A2058 melanoma cells, and to enhance doxorubicin-induced apoptosis in MDA-MB-231 cells.

*Corresponding author,
deshayes@gene.com.

Received for review June 27, 2006
and accepted August 25, 2006.

Published online September 15, 2006

10.1021/cb600276q CCC: \$33.50

© 2006 by American Chemical Society

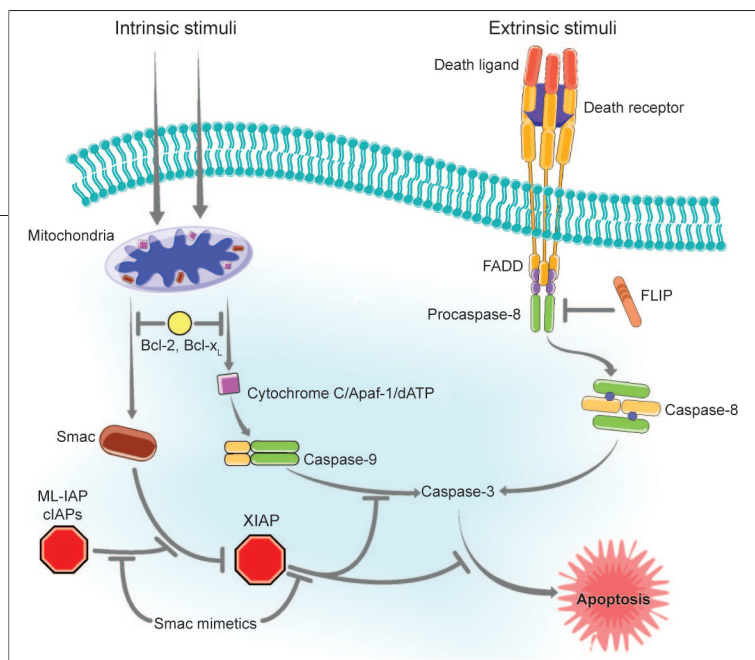


Figure 1. Apoptotic pathway. The extrinsic apoptotic pathway is triggered when death receptors, such as Fas, DR4, or TNF receptor 1, are engaged by their cognate ligands, resulting in recruitment of the adaptor protein Fas-Associated Death Domain (FADD) and the apical caspase, caspase-8. This leads to activation of caspase-8 and subsequent activation of the effector caspases, caspase-3 and -7. The intrinsic apoptotic pathway is triggered by stimuli such as irradiation, chemotherapeutic agents, or growth factor withdrawal. Activation of pro-apoptotic BH3-only members of the Bcl-2 family neutralizes the anti-apoptotic proteins Bcl-2, Bcl-x_L, and Mcl-1, leading to disruption of the mitochondrial membrane potential and the release of cytochrome c and Smac from the mitochondria into the cytoplasm. These events result in Apaf-1-mediated activation of caspase-9 and subsequent activation of the effector caspases, caspase-3 and -7, and culminate in dismantling of the cell. The IAP proteins are the last line of defense against cellular suicide and act by inhibiting caspases (XIAP) and by preventing Smac from blocking XIAP-mediated caspase inhibition (ML-IAP, cIAP1, and cIAP2). Smac mimetics bind to IAP proteins and block their inhibitory activity by antagonizing the critical IAP–caspase and IAP–Smac interactions.

cally active homodimer conformation (21). The four N-terminal residues of the small subunit of caspase-9, ATPF, make critical interactions with a peptide-binding groove on XIAP-BIR3 (21, 22). The endogenous IAP antagonist protein second mitochondrial activator of caspases (Smac) is released from the mitochondria in response to pro-apoptotic stimuli (23, 24). This mature, processed form of Smac has been demonstrated to bind to the same peptide-binding groove of XIAP-BIR3 *via* its four N-terminal residues, AVPI, thus releasing caspase-9 and promoting apoptosis (22, 25, 26).

Recent data suggest that ML-IAP, cIAP1, and cIAP2 may inhibit apoptosis primarily by binding Smac and thereby preventing it from antagonizing the ubiquitously expressed XIAP, rather than by directly inhibiting caspases (27–29). Regardless of the specific mode of action, it is reasonable to believe that binding a Smac-based peptidomimetic to the relevant BIR domains of these IAP proteins will have a pro-apoptotic effect. Indeed, Smac-derived peptides have been demonstrated to sensitize a number of different tumor cell lines

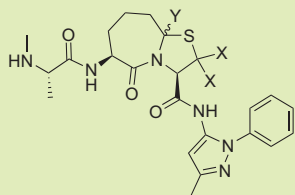
to apoptosis induced by a variety of pro-apoptotic drugs (20, 30–33). Pro-apoptotic activity has also been reported for peptidomimetics, both *in vitro* and *in vivo* (34–37). Such agents hold the promise of reducing therapeutic resistance in cancers where IAP expression precludes apoptosis. We report here the development of Smac mimetics based on an easily accessible [7,5]-bicyclic scaffold, including design rationale, synthesis, binding affinities, and an X-ray crystal structure of an isostere–IAP–BIR domain complex. Evidence for isostere-driven activation of caspase-3 and -7 and induction of apoptosis in cancer cells is also presented.

RESULTS AND DISCUSSION

Structural Analysis of Peptide–BIR Domain

Complexes. Structure/activity relationships of peptide binding to the BIR domain of ML-IAP and the BIR2 and BIR3 domains of XIAP were investigated previously using peptide-phage display technology and peptide library positional scanning (38). Most dramatically, for peptides selected to bind ML-IAP-BIR and XIAP-BIR3, alanine was found in the N-terminal (P1) position in 100% of the selectants. Other positions tolerated some variation, although preferences for valine at P2, proline at P3, and the aromatic residues phenylalanine or tryptophan at P4 were apparent. Of these, the P2 position could tolerate the most substitutions without significant losses in affinity. Examination of the crystal structure of the most potent peptide, AVPW (Table 1), in complex with MLXBIR3SG (an ML-IAP/XIAP chimeric BIR domain that preserves the native ML-IAP peptide-binding site) (27) reinforces the key points of the phage selection experiments (Figure 2, panel a).

The structure of the chimeric BIR domain in the 1.6 Å resolution structure of the MLXBIR3SG–AVPW complex is essentially identical to that reported previously for the MLXBIR3SG–Smac 9-mer peptide complex (27) with a C α root mean square deviation of 0.115 Å for 180 atom pairs. In the P1 position, the alanine methyl group is buried in a hydrophobic pocket formed by the side chains of Leu131, Trp134, and Glu143 (ML-IAP residue numbering) of the protein. Thus, substituting this methyl group with anything larger leads to a large decrease in binding affinity. The N-terminus of the peptide is in an acidic environment with the amino group of Ala1 hydrogen bonded to the side chain carboxylates of Asp138 and Glu143. Others have noted that N-methylation of

TABLE 1. Structure of isosteres 8–11 and binding affinities for XIAP-BIR3, ML-IAP-BIR, cIAP1-BIR3, and cIAP2-BIR3

Compound	X	Y	K_i (μM)			
			XIAP-BIR3	ML-IAP-BIR ^a	cIAP1-BIR3	cIAP2-BIR3
AVPIAQKSE (1)			0.67	0.5	<i>b</i>	<i>b</i>
2			39.8	15.3	<i>b</i>	<i>b</i>
AVPW			0.05	0.04	0.03	0.07
8	CH ₃	····H (<i>R</i>)	0.77	0.05	0.05	0.13
9	CH ₃	◀H (<i>S</i>)	2.34	0.19	<i>b</i>	<i>b</i>
10	H	····H (<i>R</i>)	0.27	0.28	0.11	0.49
11	H	◀H (<i>S</i>)	3.00	4.19	<i>b</i>	<i>b</i>

^aBinding affinity for ML-IAP-BIR was determined using the chimeric protein MLXBIR3SG; binding affinity for cIAP1-BIR3 was determined using the chimeric protein cIAP1XBIR3; binding affinity for cIAP2-BIR3 was determined using the chimeric protein cIAP2XBIR3.

^bNot determined.

the alanine is tolerated without loss of binding affinity (34, 37). The relative promiscuity at the P2 position is explained by the lack of substantial interaction between the Val2 side chain and the protein, although this residue does form two main chain–main chain hydrogen bonds with the protein. Pro3 makes extensive van der Waals contact with Trp147 and helps position the P1 and P4 residues for optimal binding. The P4-binding pockets of the BIR domains were shown to accommodate a number of large aromatic groups, as illustrated by the deep pocket binding the tryptophan side chain in the crystal structure of the MLXBIR3SG–peptide complex (Figure 2, panel a).

Despite its high binding affinity, AVPW has no measurable biological activity in cell-based assays (data not shown). Therefore, to obtain active agents, it is necessary to translate the key components of the AVPW structure onto a non-peptide scaffold with more druglike properties.

Isostere Design. The first step toward evolving the peptide into a more druglike molecule was to determine whether the bioactive conformation of the peptide

sequence overlapped with that of known dipeptide isosteres. A series of five dipeptide isosteres were substituted into positions P2 and P3 of the Smac-based peptide AVPIAQKSE, and K_i values were determined for binding to ML-IAP-BIR and XIAP-BIR3. Incorporation of conformationally rigid dipeptide isosteres such as benzodiazepine or 3-amino-1-carboxymethylvalerolactone in positions P2 and P3 resulted in complete loss of activity. In both cases, manual docking into the peptide-binding site of ML-IAP-BIR suggested unfavorable steric clashes with Trp147 as the cause of this loss in binding affinity. Molecular modeling suggested that a more flexible 7-membered ring might

adopt a conformation that allows favorable interactions with Trp147, while maintaining the P2 hydrogen bonding with the protein as well as optimal binding orientations at positions P1 and P4. Thus, substituting 3-amino-1-carboxymethylcaprolactam into positions P2

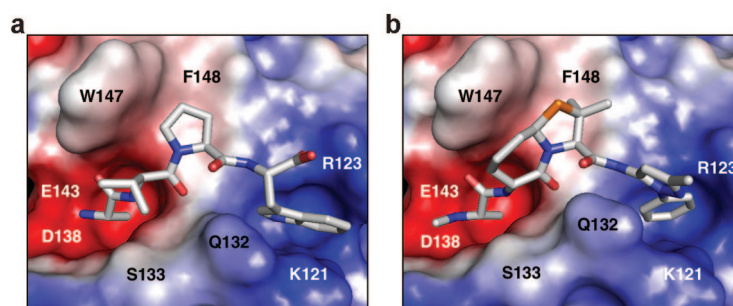


Figure 2. Structure of the AVPW peptide and isostere 8 bound to the Smac binding site. Solvent-accessible surface representation of the peptide-binding site of MLXBIR3SG in complex with a) the phage-derived peptide AVPW and b) peptide isostere 8. The protein surface is color-coded according to electrostatic surface potential: red is negatively charged; blue is positively charged. This figure was produced using the program PyMOL (www.pymol.org) (46).

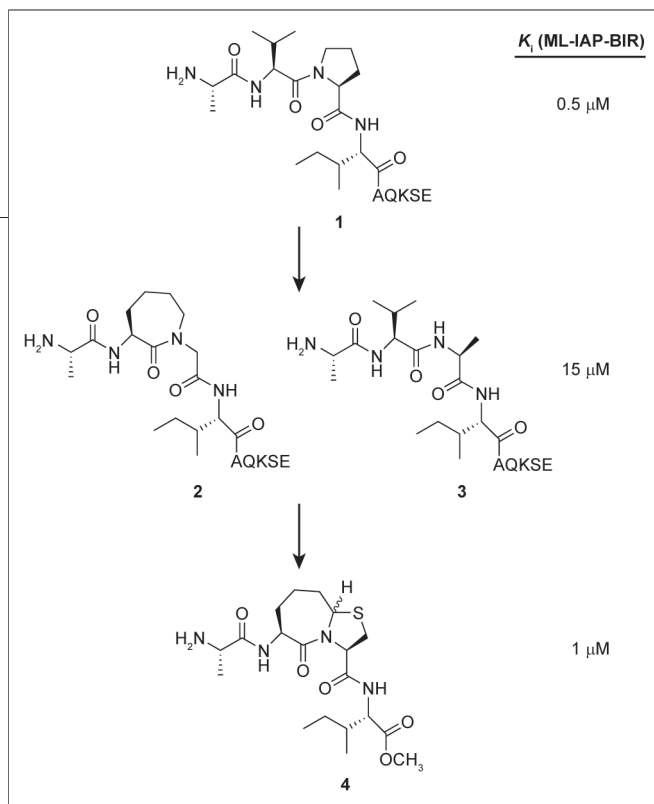


Figure 3. Conversion of peptide to isostere. Mutation of peptide 1 to give either the caprolactam 2 or the peptide 3, where proline is substituted with alanine, results in a 30-fold loss in binding affinity for the BIR domain of ML-IAP. Fusion of a five-membered ring to the seven-membered ring of 2 gives 4, which has a binding affinity comparable to the starting peptide 1.

and P3 of the Smac-derived 9-mer peptide resulted in compound **2** with binding affinity (K_i) to the ML-IAP-BIR and XIAP-BIR3 domains of 15 and 40 μM , respectively. While this represents a 30–60-fold loss in affinity relative to the starting peptide, prior peptide positional scanning data showed similar losses in affinity when proline was substituted with an alanine at position P3 (compound **3**) (38). We reasoned that introduction of a fused 5-membered ring within the context of the caprolactam to yield a [7,5]-bicyclic compound (**4**) should retain the high affinity of the endogenous peptide sequence, while altering the peptide character of the P2 and P3 positions (Figure 3).

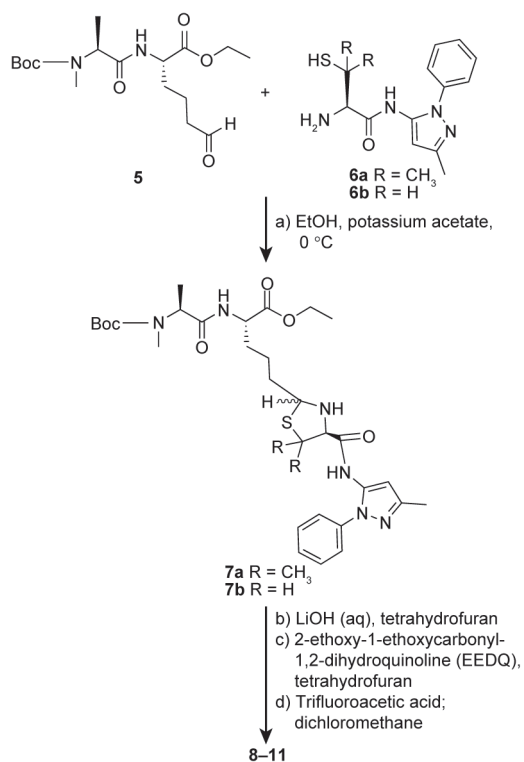
Similar bicyclic lactams have been used previously to conformationally constrain a peptide backbone (39). For synthetic ease, a sulfur was chosen for the 3 position of the 5-membered ring. Since we had discovered previously that substitution of proline in the tetrapeptide AVPI with γ -thioproline was tolerated without loss of affinity (data not shown), this scaffold was chosen as our initial target (Figure 3). The synthesis of Patchett *et al.* (40) was used to construct the bicyclic [7,5]-lactam **4** as a 1:1 mixture of the two ring-junction diastereomers (Figure 3). The diastereomeric mixture was determined to have a K_i value for binding to ML-IAP-BIR of 1 μM .

Depending on the potency of each diastereomer, the addition of the 5-membered ring to the caprolactam peptide isostere results in a 15–30-fold improvement in affinity for ML-IAP-BIR and thus gives a conformationally constrained Smac mimetic with an affinity comparable to that of the 9-mer peptide **1** (Table 1). This, in turn, indicates that the [7,5]-bicyclic lactam effectively constrains the middle portion of the molecule (P2–P3) and can be expected to provide a scaffold that can be further optimized at positions P1 and P4. *N*-Methyl alanine in the P1 position was found to impart proteolytic stability and, as noted previously, is tolerated without loss of binding affinity. Given the preference for aromatic groups in the P4 position, we focused on incorporating heterocyclic rings that were able to place an aromatic group in the correct orientation to interact with the P4-binding pocket of the protein. A variety of 5-membered ring heterocycles were synthesized, with pyrazole **6** (Scheme 1) used to construct isosteres providing optimal positioning of the aromatic ring in P4.

Synthesis. A convergent synthetic approach was required to allow access to a variety of analogues (Scheme 1). The aldehyde **5** was prepared in three steps from diphenylmethylene glycine ethyl ester (synthetic details are given in Supporting Information). The key step of the synthesis is the condensation of **5** with primary amines **6a,b**, derived from either cysteine or penicillamine, to give thiazolidines **7a,b**. Saponification followed by ring closure yielded peptide isosteres **8–11** (Table 1).

Discrimination between the BIR Domains. The K_i values for binding of the different Smac mimetics to the relevant BIR domains of XIAP, cIAP1, cIAP2, and ML-IAP were determined using a fluorescence polarization assay as described in Methods (Figure 4).

The isosteres with the preferred bridgehead stereochemistry bind to the BIR domains in the submicromolar range with the highest affinity observed for isostere **8** binding to MLXBIR3SG (which can be considered equivalent to ML-IAP-BIR, with regard to Smac peptide binding) (27) and cIAP1XBIR3 (which can be considered equivalent to cIAP1-BIR3, with regard to Smac peptide binding) (Table 1). In the case of the penicillamine derivatives, the nonpreferred ring-junction stereochemistry (isostere **9**) results in a 3–4-fold decrease in binding affinity for both XIAP-BIR3 and MLXBIR3SG relative to isostere **8**. Similarly, for the cysteine derivatives, reversing the ring-junction stereochemistry of the preferred isostere **10** to give the diastereomer **11** results in a greater than 10-fold decrease



Scheme 1. Synthesis of isosteres 8–11.

in affinity for both XIAP-BIR3 and MLXBIR3SG. As anticipated (38), the penicillamine derivatives bind more tightly to MLXBIR3SG than to XIAP-BIR3, while the cysteine-derived isosteres **10** and **11** bind with essentially equal affinity to both XIAP-BIR3 and MLXBIR3SG. These data agree with a model in which the hydroxyl group of Tyr324 of XIAP-BIR3 (the residue corresponding to Phe148 in ML-IAP, Phe330 in cIAP1, and Phe316 in cIAP2) has a steric clash with the (pro-*R*)-methyl groups in the P3 positions of isosteres **8** and **9**, while the aromatic side chains of the corresponding phenylalanine residues in the other BIR domains investigated have favorable interactions with the (pro-*R*)-methyl groups of these compounds. In these examples, the selectivity for ML-IAP-BIR over XIAP-BIR3 is on the order of 15-fold, compared with the ~100-fold discrimination observed previously when (3*S*)-methyl-proline was substituted into the AVPIAQKSE peptide (38).

Structure of MLXBIR3SG–Isostere 8 Complex. To gain a more detailed understanding of the interactions between the Smac mimetics and the BIR domain of

ML-IAP, the crystal structure of the complex between MLXBIR3SG and peptide isostere **8** was determined to a resolution of 2.3 Å and compared with that of the AVPW peptide complex (Figure 2, panel b). X-ray data collection and refinement statistics are listed in Supplementary Table 1.

The key contacts seen in the structure of the MLXBIR3SG–AVPW peptide complex (Figure 2, panel a) are conserved in the structure of the complex with peptide isostere **8** (Figure 2, panel b). For instance, the main chain–main chain hydrogen bonds noted in the AVPW complex structure are completely conserved in the structure of the complex with isostere **8**. The alanine groups are also superimposable in the two structures, with the side chain methyl groups being buried in a hydrophobic pocket and the *N*-methyl group of isostere **8** being exposed to the solvent. The [7,5]-bicyclic moiety of isostere **8** forms a slightly larger hydrophobic interface with the protein, through contacts with Trp147 and Phe148, than the corresponding valine and proline residues in the AVPW peptide complex. Overall, however, the two complexes bury approximately the same total surface area, with the AVPW peptide complex burying 722 Å² and the isostere **8** complex burying a total of 747 Å².

Mechanism of Action. To determine whether potent peptide isosteres **8** and **10** (and the weak binding **11** as a control) can prevent physical interaction between ML-IAP and Smac, 293T cells were transiently transfected with Flag-tagged ML-IAP or empty vector and myc-tagged Smac (Figure 5). Following transfection, Smac autoprocesses to generate the mature processed form that is capable of binding to IAP proteins (Figure 5). Incubation of lysates with isosteres **8**, **10**, and **11** reduced ML-IAP–Smac binding in a dose-dependent manner (Figure 5). Incubation of cellular lysates with isosteres **8** or **10** was more efficient at blocking the ML-IAP–Smac interaction than was addition of isostere **11**, in agreement with their respective binding affinities (Table 1). These results demonstrate that the peptide isosteres are capable of competitively blocking protein–protein interactions involving ML-IAP.

Caspase Activation and Cell Killing. Antagonism of IAP proteins has been shown to stimulate apoptosis in cancer cell lines. To determine whether the peptide isosteres can induce apoptosis in cancer cells, MDA-MB-231 breast cancer and A2058 melanoma cells were treated with isosteres **8**, **10**, and **11** (Figure 6, panels a

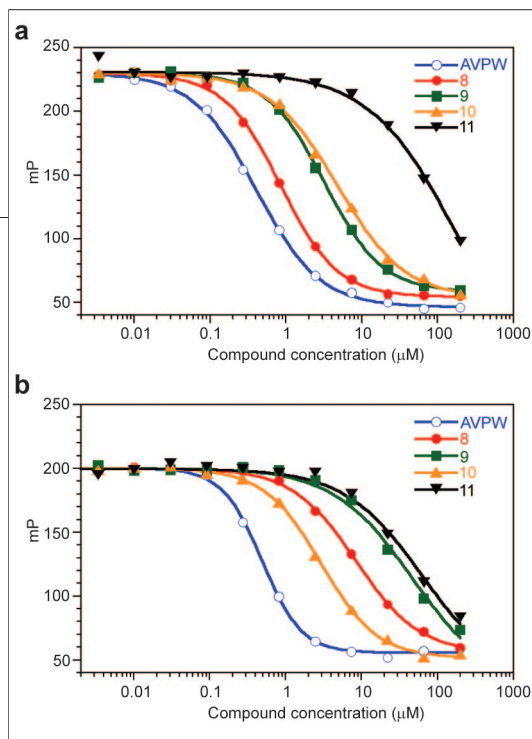


Figure 4. Competition binding assays. Competition binding curves for Smac mimetics binding to the a) MLXBIR3SG and b) XIAP-BIR3 domains. See Methods for details of the fluorescence polarization-based binding assays. The K_i values (Table 1) were calculated from IC_{50} values as described (43).

and b). Similar to other structurally distinct Smac mimetics (34, 37), isosteres **8** and **10** exhibit single-agent killing of the treated cells with approximate IC_{50} values of 0.1 and 2 μ M, respectively. Isostere **11** shows no detectable single-agent activity, in agreement with its lower affinity for the IAP-BIR domains. In addition, treatment of MDA-MB-231 cells with isostere **8** induced nuclear condensation and fragmentation further supporting its pro-apoptotic activity (data not shown).

Treatment of the susceptible A2058 melanoma cells with isosteres **8** and **10** also results in activation of caspase-3 and -7 (Figure 6, panel c), suggesting that the single-agent cell killing observed for these compounds is due to the induction of apoptosis. Isostere **8** induced significant caspase-3/7 activation in the submicromolar concentration range, consistent with its higher activity in the cell viability assay. Isostere **10** shows significant caspase-3/7 activation in the low micromolar concentration range, which is again consistent with their observed potency in the cell viability assay.

Smac-based peptides and Smac mimetics have been shown to sensitize cancer cells to a variety of chemotherapeutic agents. To determine whether the IAP-binding peptide isosteres might also collaborate with chemotherapeutic agents to stimulate cell death, cell viability of MDA-MB-231 cells treated with isosteres **8**, **10**, and **11** at 1 μ M concentrations in the presence of

increasing doses of doxorubicin was investigated (Figure 6, panel d). Isostere **8** shows potent single-agent activity and additivity with doxorubicin. Isostere **10** shows modest single-agent activity that also appears to be additive with the effect of doxorubicin. No significant activity or additivity with doxorubicin is seen with isostere **11**.

CONCLUSIONS

In this study, a discrete and tight-binding peptide motif has been translated into a peptide isostere that has comparable binding affinity for relevant BIR domains of IAP proteins. The peptide isostere contains a [7,5]-bicyclic lactam that conformationally constrains the P2–P3 portion of the molecule and a 5-membered heterocycle that optimally positions an aromatic ring in the P4 position. Whereas the tightest binding peptide discovered, AVPW, shows no cellular biological activity, the peptide isosteres with the preferred ring-junction stereochemistry disrupt key IAP protein–protein interactions, cause activation of effector caspases, and exhibit both single-agent cell killing and additivity with doxorubicin in cancer cell lines. Similar conformationally constrained Smac mimetics have been reported previously, although

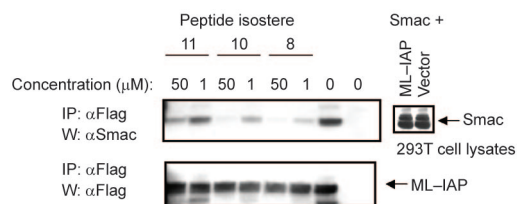


Figure 5. Immunoprecipitation shows isosteres compete with Smac for the ML-IAP binding site. HEK293T cells were transiently transfected with Flag-tagged ML-IAP or vector plasmid and myc-tagged Smac. After 40 h transfection, cells were lysed in NP40 lysis buffer. Lysates were incubated with 0, 1, or 50 μ M concentrations of the indicated peptide isosteres for 2 h, followed by immunoprecipitation with anti-Flag antibodies for 3 h, SDS-PAGE, and immunoblotting with anti-Flag (lower panel) and anti-Smac (upper panel) antibodies. The right-hand panel shows the input amount of Smac protein from HEK293T cells co-transfected with ML-IAP or the empty vector. The right-most lane of each panel represents cells co-transfected with myc-tagged Smac and the empty vector.

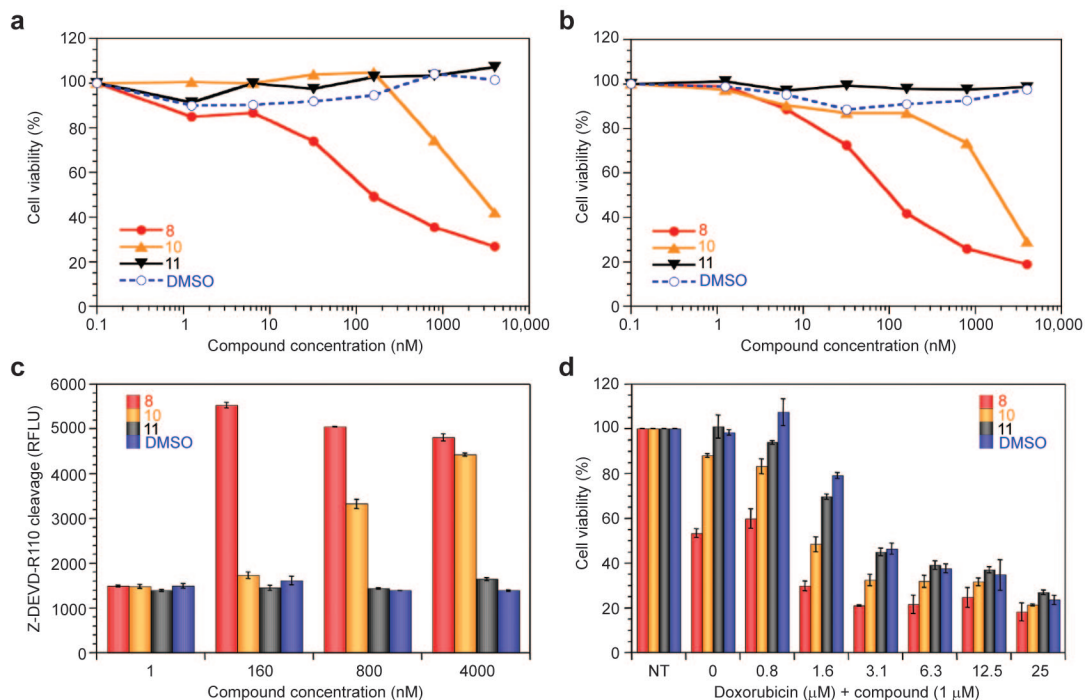


Figure 6. Cell killing and caspase activation experiments show single-agent activity and cooperation with doxorubicin. **a)** MDA-MB-231 breast carcinoma and **b)** A2058 melanoma cells were treated with indicated amounts of isosteres 8, 10, and 11 or DMSO control. Cell death was assessed by neutral red staining 30 h after the start of the treatment. **c)** A2058 cells were treated with indicated amounts of isosteres 8, 10, and 11 or DMSO control for 30 h, and caspase-3 and -7 activation was assessed using Z-DEVD-R110 as the fluorogenic substrate. **d)** MDA-MB-231 cells were treated with 1 μM of isosteres 8, 10, and 11 or DMSO control in the presence of indicated amounts of doxorubicin. Cell death was assessed as in panels **a** and **b**. NT = nontreated cells.

single-agent cell killing activity was not demonstrated for these compounds (36). In summary, these results demonstrate that small-molecule Smac mimetics can

overcome the inhibitory effects of the IAP proteins and that such compounds have promise as potential cancer therapeutics.

METHODS

Protein Production. The ML-IAP-BIR/XIAP-BIR3 chimeric protein, MLXBIR3SG, and the BIR3 domain of XIAP (XIAP-BIR3) were expressed and purified as described previously (27). cIAP1-BIR3/XIAP-BIR3 and cIAP2-BIR3/XIAP-BIR3 chimeric proteins, cIAP1XBIR3 and cIAP2XBIR3, respectively, were engineered, expressed, and purified in a similar fashion to MLXBIR3SG. The amino acid sequences of these BIR domains retain the native cIAP peptide-binding sites (Supplementary Figure 1). Purified proteins were concentrated and stored at $-80\text{ }^{\circ}\text{C}$.

Compound Preparation. Isosteres **8–11** were synthesized as described in Supporting Information text and stored as solids at $4\text{ }^{\circ}\text{C}$.

X-ray Crystallographic Analysis of MLXBIR3SG–Antagonist Complexes. The peptide was reconstituted from lyophilized powder in 10 mM 2-(*N*-morpholino)ethanesulfonic acid, pH 5.5; final peptide concentration (25 mg mL^{-1}) was verified by A_{280} .

Purified MLXBIR3SG protein ($20\text{--}25\text{ mg mL}^{-1}$) was mixed with peptide (AVPW) in a 1:2 protein/peptide ratio. The MLXBIR3SG–peptide complex was mixed with crystallization well solution (100 mM Bis-tris(hydroxymethyl)aminomethane (tris), pH 6; 200 mM lithium sulfate; 20–25% (w/v) poly(ethylene glycol) 3350) in a ratio of 1 mL of protein complex to 1 mL of well solution. Hanging or sitting drops of the mixed solutions were then allowed to equilibrate by vapor diffusion against a reservoir of the well solution. Tetragonal bipyramidal crystals appear in a few days and grow to full size (typically 0.1–0.2 mm long) in a week.

Crystals of MLXBIR3SG complexed with AVPW were removed from the crystallization drop and transferred to a stabilizer drop (5 μL) containing 100 mM Bis-tris, pH 6; 200 mM lithium sulfate; 30% (w/v) poly(ethylene glycol) 3350; and 0.5–1.0 mM of compound **8**. Crystals were left in the soaking solution overnight as a hanging drop over a reservoir of the same solution (to

prevent evaporation). Crystals were then transferred to a cryostabilizer containing 100 mM Bis-tris, pH 6, 200 mM lithium sulfate; 30% (w/v) poly(ethylene glycol) 3350; 15% (v/v) ethylene glycol; and 0.5–1.0 mM of the same compound used in the soaking stage. After 15–20 min in the cryostabilizer, the crystals were frozen in liquid nitrogen. Crystals of the parent complex (AVPW) were prepared in the same manner, except that the AVPW peptide was substituted for the inhibitor compound in the soaking and cryostabilizer solutions.

Data for the AVPW peptide complex were collected at the Advanced Photon Source (Argonne, IL), while data for the compound **8** complex were collected using an in-house X-ray source. Data statistics are listed in Supplementary Table 1.

The starting models for refinement of these structures were derived from a 1.3 Å resolution structure of a different peptidomimetic complex (details to be published elsewhere), which was stripped of the peptidomimetic antagonist molecule and all water molecules within 10 Å of it. After one round of refinement of the antagonist-free model, the new antagonists were built into clear difference electron density visible in $F_o - F_c$ maps. New water molecules were picked automatically using the program Arp/wArp (41), and the entire new complex models were subjected to several rounds of positional, anisotropic *B*-factor, and translation–libration–screw refinement using Refmac5 (42). Refinement statistics for the complex structures are in Supplementary Table 1.

Binding Assays. Initial polarization experiments were performed in order to determine dissociation constants (K_d) between IAP protein BIR domains and fluorescent probes. Samples for fluorescence polarization affinity measurements were prepared by addition of serial dilutions of MLXBIR3SG, XIAP-BIR3, cIAP1XBIR3, or cIAP2XBIR3 in polarization buffer (50 mM Tris [pH 7.2], 120 mM NaCl, 1% bovine globulins, 5 mM dithiothreitol, and 0.05% octylglucoside) to 5 nM 5-carboxy-fluorescein (5-FAM)-conjugated AVP-diphenylalanine-AKK (AVP-diPhe-FAM). The reactions were read after an incubation time of 30 min at RT with standard cut-off filters for the fluorescein fluorophore ($\lambda_{ex} = 485$ nm; $\lambda_{em} = 530$ nm) in 384-well black HE96 plates (Molecular Devices Corp.). Fluorescence polarization values were plotted as a function of the protein concentration, and the effective concentration 50 (EC_{50}) values were obtained by fitting the data to a four-parameter equation using KaleidaGraph software (Synergy software, Reading, PA). The apparent K_d values were determined from the EC_{50} values.

Inhibition constants (K_i) for the antagonists were determined by addition of 0.06 μ M MLXBIR3SG, 0.5 μ M XIAP-BIR3, 0.2 μ M cIAP1XBIR3, or 0.4 μ M cIAP2XBIR3 to wells containing 1:3 serial dilutions of the antagonists and 5 nM AVP-diPhe-FAM probe in the polarization buffer. Samples were read after a 30-min incubation. Fluorescence polarization values were plotted as a function of the antagonist concentration, and the IC_{50} values were obtained by fitting the data to a four-parameter equation using KaleidaGraph software. K_i values for the antagonists were determined from the IC_{50} values (43).

Cellular Coimmunoprecipitation. HEK293T cells were transiently transfected with Flag-tagged ML-IAP or vector plasmid and myc-tagged Smac. Forty hours after transfection, cells were lysed in NP40 lysis buffer (44). Lysates were incubated with 0, 1, or 50 μ M of the indicated compounds for 2 h, followed by immunoprecipitation with anti-Flag antibodies for 3 h, SDS-PAGE, and immunoblotting with anti-Flag and anti-Smac antibodies.

Analysis of Apoptosis. Human breast carcinoma MDA-MB-231 and human melanoma A2058 cells were obtained from ATCC. Cells were grown in 50:50 Dulbecco's modified Eagle's and FK12 medium supplemented with 10% fetal bovine serum (FBS), penicillin, and streptomycin. Doxorubicin was purchased from Sigma. Cells ($(1.5-2) \times 10^4$ cells per well) were seeded into

96-well dishes in media containing 5% heat-inactivated FBS. The medium was changed 8–12 h later, and cells were treated with different compounds, either alone or in combination with doxorubicin, for 30 h. Cell viability was measured by neutral red uptake as described previously (45).

Caspase Activation. Cells were seeded and treated with compounds, either alone or in combination with doxorubicin (as described above). Caspase-3/7 activity was measured 30 h later using the Apo-ONE™ Homogeneous Caspase-3/7 assay kit (Promega) according to the manufacturer's instructions.

Accession codes: The coordinates have been deposited with the Protein Data Bank and are listed as follows: The AVPW structure is listed as 2I3H and the isostere **8** structure as 2I3I.

Acknowledgments: We thank C. Quan, J. Tom, M. Stuble, and J. Dority and the Genentech DNA synthesis group for their valuable assistance.

Note added after print publication: Because of a production error, the following references were misformatted: 1–39, 41–44. These errors do not affect the scientific integrity of the article. This paper was originally posted September 15, 2006, and the electronic version was corrected and reposted to the web on October 20, 2006. An Addition and Correction may be found in *ACS Chem. Biol.* 1(9).

Supporting information available: This material is available free of charge via the Internet.

REFERENCES

- Aravind, L., Dixit, V. M., and Koonin, E. V. (1999) The domains of death: evolution of the apoptosis machinery, *Trends Biochem. Sci.* 24, 47–53.
- Thompson, C. B. (1995) Apoptosis in the pathogenesis and treatment of disease, *Science* 267, 1456–1462.
- Salvesen, G. S., and Dixit, V. M. (1997) Caspases: intracellular signaling by proteolysis, *Cell* 91, 443–446.
- Salvesen, G. S., and Duckett, C. S. (2002) IAP proteins: blocking the road to death's door, *Nat. Rev. Mol. Cell. Biol.* 3, 401–410.
- Tamm, I., Komblau, S. M., Segall, H., Krajewski, S., Welsh, K., Kitada, S., Scudiero, D. A., Tudor, G., Qui, Y. H., Monks, A., Andreeff, M., and Reed, J. C. (2000) Expression and prognostic significance of IAP-family genes in human cancers and myeloid leukemias, *Clin. Cancer Res.* 6, 1796–1803.
- Tamm, I., Richter, S., Scholz, F., Schmelz, K., Oltersdorf, D., Karawajew, L., Schoch, C., Haferlach, T., Ludwig, W. D., and Wuchter, C. (2004) XIAP expression correlates with monocytic differentiation in adult de novo AML: impact on prognosis, *Hematol. J.* 5, 489–495.
- LaCasse, E. C., Baird, S., Komeluk, R. G., and MacKenzie, A. E. (1998) The inhibitors of apoptosis (IAPs) and their emerging role in cancer, *Oncogene* 17, 3247–3259.
- Stennicke, H. R., Ryan, C. A., and Salvesen, G. S. (2002) Reprieve from execution: the molecular basis of caspase inhibition, *Trends Biochem. Sci.* 27, 94–101.
- Li, J., Feng, Q., Kim, J. M., Schneiderman, D., Liston, P., Li, M., Vanderyden, B., Faught, W., Fung, M. F., Senterman, M., Komeluk, R. G., and Tsang, B. K. (2001) Human ovarian cancer and cisplatin resistance: Possible role of inhibitor of apoptosis proteins, *Endocrinology* 142, 370–380.
- Fong, W. G., Liston, P., Rajcan-Separovic, E., St Jean, M., Craig, C., and Komeluk, R. G. (2000) Expression and genetic analysis of XIAP-associated factor 1 XAF1 in cancer cell lines, *Genomics* 70, 113–122.
- Cummins, J. M., Kohli, M., Rago, C., Kinzler, K. W., Vogelstein, B., and Bunz, F. (2004) X-linked inhibitor of apoptosis protein (XIAP) is a nonredundant modulator of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-mediated apoptosis in human cancer cells, *Cancer Res.* 64, 3006–3008.

12. Vucic, D., Stennicke, H. R., Pisabarro, M. T., Salvesen, G. S., and Dixit, V. M. (2000) ML-IAP, a novel inhibitor of apoptosis that is preferentially expressed in human melanomas, *Curr. Biol.* **10**, 1359–1366.
13. Gong, J., Chen, N., Zhou, Q., Yang, B., Wang, Y., and Wang, X. (2005) Melanoma inhibitor of apoptosis protein is expressed differentially in melanoma and melanocytic naevus, but similarly in primary and metastatic melanomas, *J. Clin. Pathol.* **58**, 1081–1085.
14. Crnkovic-Mertens, I., Hoppe-Seyler, F., and Butz, K. (2003) Induction of apoptosis in tumor cells by siRNA-mediated silencing of the livin/ML-IAP/KIAP gene, *Oncogene* **22**, 8330–8336.
15. Crnkovic-Mertens, I., Semzow, J., Hoppe-Seyler, F., and Butz, K. (2006) Isoform-specific silencing of the Livin gene by RNA interference defines Livin beta as key mediator of apoptosis inhibition in HeLa cells, *J. Mol. Med.* **84**, 232–240.
16. Rothe, M., Pan, M. G., Henzel, W. J., Ayres, T. M., and Goeddel, D. V. (1995) The TNFR2-TRAF signaling complex contains two novel proteins related to baculoviral inhibitor of apoptosis proteins, *Cell* **83**, 1243–1252.
17. Dai, Y., Rahmani, M., and Grant, S. (2003) Proteasome inhibitors potentiate leukemic cell apoptosis induced by the cyclin-dependent kinase inhibitor flavopiridol through a SAPK/JNK- and NF-kappaB-dependent process, *Oncogene* **22**, 7108–7122.
18. Imoto, I., Tsuda, H., Hirasawa, A., Miura, M., Sakamoto, M., Hirohashi, S., and Inazawa, J. (2002) Expression of cIAP1, a target for 11q22 amplification, correlates with resistance of cervical cancers to radiotherapy, *Cancer Res.* **62**, 4860–4866.
19. Imoto, I., Yang, Z. Q., Pimkhaokham, A., Tsuda, H., Shimada, Y., Imamura, M., Ohki, M., and Inazawa, J. (2001) Identification of cIAP1 as a candidate target gene within an amplicon at 11q22 in esophageal squamous cell carcinomas, *Cancer Res.* **61**, 6629–6634.
20. Yang, L., Cao, Z., Yan, H., and Wood, W. C. (2003) Coexistence of high levels of apoptotic signaling and inhibitor of apoptosis proteins in human tumor cells: implication for cancer specific therapy, *Cancer Res.* **63**, 6815–6824.
21. Shiozaki, E. N., Chai, J., Rigotti, D. J., Riedl, S. J., Li, P., Srinivasula, S. M., Alnemri, E. S., Fairman, R., and Shi, Y. (2003) Mechanism of XIAP-mediated inhibition of caspase-9, *Mol. Cell* **11**, 519–527.
22. Srinivasula, S. M., Hegde, R., Saleh, A., Datta, P., Shiozaki, E., Chai, J., Lee, R. A., Robbins, P. D., Fernandes-Alnemri, T., Shi, Y., and Alnemri, E. S. (2001) A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis, *Nature* **410**, 112–116.
23. Du, C., Fang, M., Li, Y., Li, L., and Wang, X. (2000) Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition, *Cell* **102**, 33–42.
24. Verhagen, A. M., Ekert, P. G., Pakusch, M., Silke, J., Connolly, L. M., Reid, G. E., Moritz, R. L., Simpson, R. J., and Vaux, D. L. (2000) Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins, *Cell* **102**, 43–53.
25. Liu, Z., Sun, C., Olejniczak, E. T., Meadows, R. P., Betz, S. F., Oost, T., Hermann, J., Wu, J. C., and Fesik, S. W. (2000) Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain, *Nature* **408**, 1004–1008.
26. Wu, G., Chai, J., Suber, T. L., Wu, J. W., Du, C., Wang, X., and Shi, Y. (2000) Structural basis of IAP recognition by Smac/DIABLO, *Nature* **408**, 1008–1012.
27. Vucic, D., Franklin, M. C., Wallweber, H. J., Das, K., Eckelman, B. P., Shin, H., Elliott, L. O., Kadkhodayan, S., Deshayes, K., Salvesen, G. S., and Fairbrother, W. J. (2005) Engineering ML-IAP to produce an extraordinarily potent caspase 9 inhibitor: implications for Smac-dependent anti-apoptotic activity of ML-IAP, *Biochem. J.* **385**, 11–20.
28. Wilkinson, J. C., Wilkinson, A. S., Scott, F. L., Csomos, R. A., Salvesen, G. S., and Duckett, C. S. (2004) Neutralization of Smac/Diablo by inhibitors of apoptosis (IAPs). A caspase-independent mechanism for apoptotic inhibition, *J. Biol. Chem.* **279**, 51082–51090.
29. Eckelman, B. P., and Salvesen, G. S. (2006) The human anti-apoptotic proteins cIAP1 and cIAP2 bind but do not inhibit caspases, *J. Biol. Chem.* **281**, 3254–3260.
30. Vucic, D., Deshayes, K., Ackerly, H., Pisabarro, M. T., Kadkhodayan, S., Fairbrother, W. J., and Dixit, V. M. (2002) SMAC negatively regulates the anti-apoptotic activity of melanoma inhibitor of apoptosis (ML-IAP), *J Biol Chem* **277**, 12275–12279.
31. Guo, F., Nimmanapalli, R., Paranawithana, S., Wittman, S., Griffin, D., Bali, P., O'Bryan, E., Fumero, C., Wang, H. G., and Bhalla, K. (2002) Ectopic overexpression of second mitochondria-derived activator of caspases (Smac/DIABLO) or cotreatment with N-terminus of Smac/DIABLO peptide potentiates epothilone B derivative-(BMS 247550) and Apo-2L/TRAIL-induced apoptosis, *Blood* **99**, 3419–3426.
32. Amt, C. R., Chiorean, M. V., Heldebrandt, M. P., Gores, G. J., and Kaufmann, S. H. (2002) Synthetic Smac/DIABLO peptides enhance the effects of chemotherapeutic agents by binding XIAP and cIAP1 in situ, *J Biol Chem* **277**, 44236–44243.
33. Fulda, S., Meyer, E., and Debatin, K. M. (2002) Inhibition of TRAIL-induced apoptosis by Bcl-2 overexpression, *Oncogene* **21**, 2283–2294.
34. Oost, T. K., Sun, C., Armstrong, R. C., Al-Assaad, A. S., Betz, S. F., Deckwerth, T. L., Ding, H., Elmore, S. W., Meadows, R. P., Olejniczak, E. T., Oleksijew, A., Oltersdorf, T., Rosenberg, S. H., Shoemaker, A. R., Tomaselli, K. J., Zou, H., and Fesik, S. W. (2004) Discovery of potent antagonists of the antiapoptotic protein XIAP for the treatment of cancer, *J. Med. Chem.* **47**, 4417–4426.
35. Sun, H., Nikolovska-Coleska, Z., Chen, J., Yang, C.-Y., Tomita, Y., Pan, H., Yoshioka, Y., Krajewski, K., Roller, P. P., and Wang, S. (2005) Structure-based design, synthesis and biochemical testing of novel and potent Smac peptido-mimetics, *Bioorg. Med. Chem. Lett.* **14**, 793–797.
36. Sun, H., Nikolovska-Coleska, Z., Yang, C.-Y., Xu, L., Liu, M., Tomita, Y., Pan, H., Yoshioka, Y., Krajewski, K., Roller, P. P., and Wang, S. (2004) Structure-based design of potent, conformationally constrained Smac mimetics, *J. Am. Chem. Soc.* **126**, 16686–16687.
37. Li, L., Thomas, R. M., Suzuki, H., De Brabander, J. K., Wang, X., and Harran, P. G. (2004) A small molecule Smac mimic potentiates TRAIL- and TNFalpha-mediated cell death, *Science* **305**, 1471–1474.
38. Franklin, M. C., Kadkhodayan, S., Ackerly, H., Alexandru, D., Distanco, M. D., Elliott, L. O., Flygare, J. A., Mausisa, G., Okawa, D. C., Ong, D., Vucic, D., Deshayes, K., and Fairbrother, W. J. (2003) Structure and function analysis of peptide antagonists of melanoma inhibitor of apoptosis (ML-IAP), *Biochemistry* **42**, 8223–8231.
39. Alberg, D. A., and Schreiber, S. L. (1993) Structure-based design of a cyclophilin-calcineurin bridging ligand, *Science* **262**, 248–250.
40. Patchett, A. A., and Wyvrat, M. J. (1986) U.S. Patent 4,617,301.
41. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) ARP/wARP and molecular replacement, *Acta Crystallogr. D57*, 1445–1450.
42. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240–255.
43. Keating, S. M., Marsters, J., Beresini, M., Ladner, C., Zioncheck, K., Clark, K., Arellano, F., and Bodary, S. (2000) Putting the pieces together: Contribution of fluorescence polarization assays to small-molecule lead optimization, *Proc. SPIE* **3913**, 128–137.
44. Vucic, D., Kaiser, W. J., Harvey, A. J., and Miller, L. K. (1997) Inhibition of Reaper-induced apoptosis by interaction with inhibitor of apoptosis proteins (IAPs), *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10183–10188.

45. Johnston, M. D., Finter, N. B., and Young, P. A. (1981) *Methods Enzymol.* 78, 394–399.
46. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stiltz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Congratulations, You've Won. . .

The beginning of October was rather uneventful for most of us. We woke to the sound of alarm clocks, spouses, or children and proceeded with our daily routines. A select few, including Roger Kornberg, Andrew Fire, and Craig Mello, took an unexpected path on the crisp October morning: the Nobel Foundation informed them that they had won a Nobel Prize. In this editorial, we celebrate their achievements. We speculate on whether we have turned a corner and narrowed the perceived cultural gap between chemists and biologists so that they can appreciate the value of combining their respective tools to dissect complex processes such as transcription and RNA interference (RNAi).

A picture is worth a thousand words, and this is evident in many of the images of the transcription machinery captured by Roger Kornberg (1–3), this year's winner of the Nobel Prize in Chemistry. Kornberg's lab, using X-ray crystallography coupled with biochemistry, systematically built molecular-level pictures of RNA polymerase, the machine responsible for copying the DNA into messenger RNA (mRNA). These images have led many scientists to dissect the details of the transcription process down to the atomic level, begun to explain how transcription factors modulate the activity of the polymerase, and opened the door for chemists to probe the mechanism by which this enzyme works.

mRNA, the product of the transcription process, led Andrew Fire and Craig Mello, who shared the Nobel Prize in Physiology or Medicine, to their discovery of the process of RNAi. They were trying to affect the levels of a muscle protein in worms by modulating the levels of mRNA transcribed from the relevant gene. Injecting more mRNA or antisense RNA did not modulate the concentration of protein, but injecting specific double-stranded RNA silenced the target gene (4). This discovery opened up a whole new area of RNA biology and chemistry.

It was Fire and Mello's original work that inspired Phil Zamore—one of the scientists featured on the Ask the Expert section of the *ACS Chemical Biology* web site—to enter this fast-paced field of research (5). Zamore notes that RNAi, which occurs naturally in plants and animals, is “one manifestation of a broader set of RNA silencing pathways that defend cells against external threats, such as viral infection, and internal threats, such as the ‘jumping’ of transposons, and that also regulate endogenous genes, shaping gene expression during development and in response to external environmental stimuli.” Today, this reverse-genetics tool is used to deliberately turn off target genes and learn more about human disease.

When a graduate student asked, “Speculate for a minute: where do you think the RNAi field will be in 5 years? What do you think are the most interesting questions to ask?” Zamore predicted that Fire and Mello, along with another colleague, would receive the prize soon. “In 5 years, the number of small RNA classes will likely double or triple from the three we now know. We will understand much more about how these small RNAs are made, sorted, and function. We will begin to understand their evolutionary origins, but we will still struggle with the important ‘why’ questions . . . Craig Mello, Andy Fire, and David Baulcombe will have justly won the Nobel Prize for the discovery of RNAi [Fire and Mello] and siRNAs [Baulcombe].” (For the complete answer, see <http://community.acs.org/journals/acbcct/cs/Ask-theExpert/ExpertResponse/tabid/72/Default.aspx?webEditionid=2&qid=5083>.) With the

Nobel Prize now awarded, we should expect an exciting 5 years in which we dissect the mechanism of RNAi, explore its potential and limits as a therapeutic (6, 7), and understand the machines that bind the RNA and use it to silence genes.

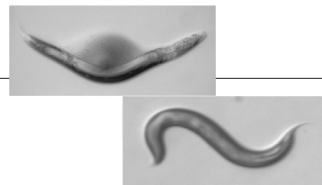
As we move forward, we look back over our shoulders to 1959, when Arthur Kornberg (Roger's father) received his Nobel Prize in Physiology or Medicine. How much have the disciplines of chemistry and biology changed in the last 47 years? Twenty years ago, Arthur Kornberg felt that chemists and biologists worked in two separate cultures (8). With the rise of chemical biology, has this cultural divide narrowed? Have chemists begun to appreciate the complexity of biological systems rather than fear it? Have they seen the need to apply their chemical tools to dissecting the complex biological machines at the molecular level? Are biologists appreciating the language and tools of chemistry and exploring how they can be used to elucidate molecular details of biological processes? When this year's Nobel Prize winners and other scientists were interviewed about the awards, they indicated that the advances made in the fields of transcription and RNAi have resulted from extensive collaborative work between chemists and biologists—perhaps we have turned a corner.



Evelyn Jabri
Executive Editor

REFERENCES

1. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution, *Science* 292, 1863–1876.
2. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution, *Science* 292, 1876–1882.
3. Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004) Structural basis of transcription: an RNA polymerase II–TFIIB cocystal at 4.5 angstroms, *Science* 303, 983–988.
4. Fire, A., Xu, S. Q., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature* 391, 806–811.
5. Brownlee, C. (2006) Discovering the building blocks of RNA interference, *ACS Chem. Biol* 1, 126–128.
6. Xia, J., Noronha, A., Toudjarska, I., Li, F., Akinc, A., Braich, R., Frank-Kamenetsky, M., Rajeev, K. G., Egli, M., and Manoharan, M. (2006) Gene silencing activity of siRNAs with a ribo-difluorotoluy nucleotide, *ACS Chem. Biol* 1, 176–183.
7. Snøve, O., Jr., and Rossi, J. J. (2006) Chemical modifications rescue off-target effects of RNAi, *ACS Chem. Biol.* 1, 274–276.
8. Kornberg, A. (1987) The two cultures: chemistry and biology, *Biochemistry* 26, 6888–6891.



Mutants by Design

Designing substrate specificity into enzyme active sites makes possible the generation of proteins with intriguing medical and industrial applications. However, the use of traditional methods for producing such altered proteins can be a formidable engineering feat. After failed attempts to employ conventional methods to tweak the substrate specificity of the DNA repair protein *O*⁶-alkylguanine-DNA alkyltransferase (AGT), Heinis *et al.* (p 575) develop a novel approach for mutant protein production that resulted in the identification of AGT mutants specific for *O*⁶-propargylguanine (PG).

Mutant proteins containing short amino acid sequences inserted into rationally chosen sites on the polypeptide surface were generated. The authors identified eight mutants that had acquired activity against PG and demonstrated several applications of such engineered proteins. For

example, fluorescent PG derivatives were used to examine the intracellular localizations of PG-binding AGT fusion proteins. In addition, two chimeric

proteins, one containing the PG-specific mutant and the other with AGT specific for *O*⁶-benzylguanine, were labeled with suitable fluorescent alkylguanine ligands. The interaction of the two fusion proteins upon addition of an appropriate ligand resulted in a FRET. This approach has significant potential for the creation of a variety of new enzymatic and fluorescent tools for biological discovery.

Published online October 20, 2006
10.1021/cb6004036 CCC: \$33.50

© 2006 by American Chemical Society

Pumping Up Antimicrobials

Microorganisms have the ability to pump out the very drugs designed to incapacitate them, rendering many potential antimicrobials essentially useless. Protein families termed multidrug-resistance pumps (MDRs) evict foreign molecules, often using a compound's polarity as the basis for recognition. The use of MDR inhibitors can prevent this process; however, harmonizing the pharmacokinetic properties of an MDR inhibitor with a structurally unrelated antimicrobial in the hopes of stumbling upon a favorable synergistic combination poses a significant challenge. Ball *et al.* (p 594) circumvent this hurdle *via* the synthesis and biological evaluation of a bifunctional-molecule-containing antibiotic and MDR efflux inhibitory properties.

The hydrophobic, cationic compound berberine has potent antibiotic activity but is also an excellent substrate for certain MDRs. Berberine was covalently linked to the MDR inhibitor INF₅₅ to produce a hybrid molecule called SS14. The antibiotic/efflux pump combination inhibitor was far more active than berberine against several bacterial strains, including *Staphylococcus aureus* and *Enterococcus faecalis*. In addition, SS14 rescued *Caenorhabditis elegans* infected with *En. faecalis*, demonstrating impressive *in vivo* efficacy. These results validate this strategy, which could be applied to other compounds susceptible to extrusion by MDRs.

Dissecting the Heterotrimeric G Proteins

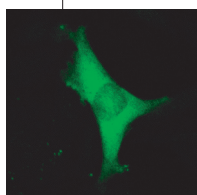
Heterotrimeric G proteins are a family of signaling molecules that interact with several receptors involved in cell signaling events. The α , β , and γ G protein subunits coexist in many different permutations, and this suggests that molecules that target specific subunits could be valuable molecular tools with which to probe G protein signaling pathways and could lead to the design of a multitude of therapeutic agents. Ja *et al.* (p 570) use messenger RNA display technology to discover and characterize the binding and biological activity of three classes of peptide-based modulators of G _{α i1}.

A library based on the peptide R6A, which is known to bind with high affinity to the GDP-state of G _{α} subunits, was designed in the search for peptides that bound to a conformation of G _{α i1} thought to be a transition-state mimic of GTP hydrolysis. Two peptides, AR6-04 and AR6-05, were identified and evaluated in cells expressing the G-protein-activated inwardly rectifying potassium channels, which become activated upon release of G _{β γ} subunits. Interestingly, the authors determined that the R6A peptide is unable to dissociate G _{α β γ} heterotrimers in these cells, whereas AR6-05 actively dissociates G _{β γ} from G _{α} and AR6-04 stabilizes the heterotrimeric complex. These new molecular scaffolds will facilitate exploration of the multifaceted world of G-protein-regulated signaling cascades.

Hormone Hopping

Thyroid hormones are critical for proper cellular development. Interaction of the active form of thyroid hormone, 3,5,3'-triiodo-L-thyronine, with three isoforms of the nuclear thyroid hormone receptor (TR α ₁, TR β ₁, and TR β ₂) triggers altered expression of various genes, stimulating a host of physiological responses. However, selective agonists and antagonists only exist for a subset of the receptors, and this makes systematic evaluation of the specific function of each receptor challenging. Now, Ocasio and Scanlan (p 585 and Point of View p 559) report the design, synthesis, and biological evaluation of the first selective agonist for TR α ₁.

Clues from TR β -selective agonists along with analysis of structure–activity data led to the creation of a novel thyromimetic containing an imidazolidinedione moiety. This compound, referred to as CO23, exhibited selective activation of TR α ₁ in cells, prompting exploration of its activity in live animals. The well-studied process of amphibian metamorphosis is known to correlate with TR-isoform expression. Notably, tadpoles treated with CO23 experienced more extensive hind-limb development but reduced resorption of gills, head, and tail than did controls. This work points to the role that TR α ₁ plays in tadpole development and provides new small-molecule tools for probing TR biology.



Introducing our AUTHORS



Christian Heinis

Current position: Medical Research Council, Cambridge, U.K., Centre of Protein Engineering, postdoctoral research associate in Sir Gregory Winter's group

Education: Swiss Federal Institute of Technology Zurich (ETHZ), M.S. in biochemistry, 2000; Ph.D. in biochemistry with Prof. Dr. Dario Neri, 2003

Postdoctoral work: Swiss Federal Institute of Technology Lausanne (EPFL) with Prof. Dr. Kai Johnsson, 2003–2005

Nonscientific interests: Mountain biking, traveling

I am interested in the evolution of proteins with desired qualities for industrial, biotechnological, or medicinal applications. In a recent project, I tested a novel strategy for altering protein substrate specificity by the insertion of amino acid loops into the protein backbone. The strategy was applied to reprogram the substrate specificity of *O*⁶-alkylguanine-DNA alkyltransferase, an enzyme that is used as a tag to label proteins in molecular imaging applications. (Read Heinis's article on p 575.)



Anthony R. Ball

Current position: Northeastern University, Department of Biology, Ph.D. candidate with Prof. Kim Lewis

Education: Northeastern University, Department of Biology, B.S. in biology, 2005

Nonscientific interests: Keeping abreast of current topics, reading fictional literature, chess, and the occasional poker game among friends

My research interests focus on circumventing bacterial mechanisms of antimicrobial resistance primarily through the inhibition of multidrug resistance (MDR) efflux transporters. In this issue, we describe the rational design of a berberine-MDR inhibitor hybrid molecule that, unlike berberine, penetrates well into bacterial cells. This approach suggests a new strategy to control drug-resistant pathogens. (Read Ball's article on p 594.)



William W. Ja

Current position: California Institute of Technology, Division of Biology, postdoctoral scholar with Prof. Seymour Benzer

Education: University of California, Berkeley, B.S. in chemistry with Prof. Alexander N. Glazer and Prof. Richard A. Mathies, high honors, 1998; California Institute of Technology, Ph.D. in chemistry with Prof. Richard W. Roberts, 2004

Nonscientific interests: Volleyball, softball, food, television, and piano

My work focuses on using messenger RNA display to develop ligands for biological targets of interest. This paper demonstrates that a diverse array of functions can be obtained from a naive library of molecules. The identified peptides can be used to study the biological mechanisms of G protein signaling and may serve as leads for therapeutic development. I am currently interested in applying these techniques to the study of aging and behavior in the fruit fly, *Drosophila melanogaster*. (Read Ja's article on p 570.)



Ofer Wiser

Current position: Hadassah University Hospital, Jerusalem, Goldyne Savad Institute of Gene Therapy, research officer with Prof. Benjamin Reubinoff

Education: Hebrew University of Jerusalem, B.S. in biology, 1993; M.S. in biochemistry, 1995; Ph.D. in molecular neurobiology with Prof. Daphne Atlas, 2001

Postdoctoral work: University of California, San Francisco, with Prof. Lily Jan, 2000–2005

Nonscientific interests: Reading books, watching movies, and spending time with my wife

What is the mechanism that enables biological processes to go on and respond to intra- and extracellular signals? I focus on monitoring interactions of the proteins participating in these processes within living cells. Here we utilize ion channels as a tool to characterize a signal transduction pathway within living cells involving heterotrimeric G proteins, and we screen peptide libraries for interactions with G proteins. Our results refine our knowledge about the critical structural determinants within G protein regulators and their physiological impact. (Read Wiser's article on p 570.)



Cory A. Ocasio

Current position: University of California, San Francisco, Chemistry and Chemical Biology Graduate Program, Ph.D. candidate with Prof. Thomas S. Scanlan

Education: University of Delaware, B.S. in biochemistry, Honors Degree with Distinction, 2001

Nonscientific interests: Traveling, running, swimming, outdoor activities, literature, spending time with animals, and San Francisco cafés

Combining synthetic organic chemistry with molecular modeling and structure–activity relationship analysis provides a means for rational drug design. I intend to identify selective ligands that elicit unique phenotypes from small panels of synthetic analogues using reverse chemical genetic screens. We utilized this approach in the discovery of a novel ligand that affects *Xenopus laevis* metamorphosis in a thyroid hormone receptor- α (TR α)-specific manner. This compound allows us to ask questions such as what role TR α plays in the mammalian heart. (Read Ocasio's article on p 585.)

Spotlight

Natural Ribozymes Get a Closer Look

Over the past two decades, a handful of catalytic RNAs, or ribozymes, have served as alternative models in the study of biological catalysis. Among the most battered and tested natural specimens are the self-cleaving RNAs, which function in the replication and maturation of viral and satellite RNAs. These ribozymes act as a built-in ruler and cutting tool to yield perfect RNA genome segments. Many labs have explored the therapeutic potential of this cutting-tool property to make custom ribozymes to stop the expression of several cellular or viral RNAs. A pair of recent studies shows that this idea was already sorted out in nature.

Recently, a ribozyme with an essential cofactor was discovered in bacteria. The GlmS ribozyme tweaks the gene expression of the nearby protein gene in a feedback loop reminiscent of the bacterial protein operons studied in basic biochemistry courses. The ribozyme resides in the untrans-

lated region upstream of the messenger RNA (mRNA), which encodes an enzyme, glucosamine-6 phosphate synthase. If the cell lacks the enzyme's product, glucosamine-6-phosphate, the mRNA remains intact and can be translated.

When the product builds

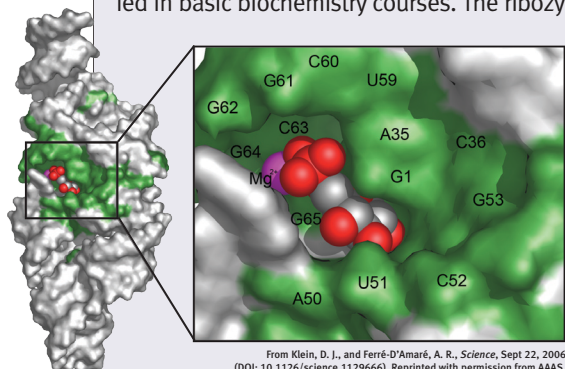
up, it binds to the ribozyme, and the self-cleavage occurs to inactivate the synthase mRNA. Now, a high-resolution look at the GlmS ribozyme demonstrates a likely mechanism that activates this cleavage event. Klein and Ferré-D'Amaré (*Science* 2006, 313, 1752–1756) determined the X-ray structure of the GlmS ribozyme with and without the cofactor mimic, glucose-6-phosphate, bound to the RNA. Whereas other metabolite-sensitive RNAs in bacteria, known as riboswitches, change conformation dramatically upon binding of their ligand, the GlmS ribozyme adopts a remarkably similar fold in both states. The RNA structure does not change in response to the mimic, and no RNA cleavage occurs. However, if the legitimate cofactor is added to the crystal, cleavage is detected. The authors conclude that the RNA structure acts as a platform to bind and position the glucosamine-6-phosphate amino group for general acid–base and electrostatic catalysis. Also, sequences forming the ligand pocket are

(continued on page 541)

Yin and Yang in p53

Interplay between *O*-GlcNAcylation and phosphorylation of proteins has been likened to a yin-yang concept, where an increase in one modification can correlate with a decrease in the other. With respect to the tumor suppressor protein p53, phosphorylation can stabilize the protein by decreasing its degradation via the ubiquitin–proteasome pathway. Although p53 is also *O*-GlcNAcylated, the sites of *O*-GlcNAcylation and its effects on p53 function are not well-defined. Now, Yang *et al.* (*Nat. Cell Biol.* 2006, 8, 1074–1083) use the small-molecule *O*-GlcNAcase inhibitor streptozotocin (STZ) to provide insight into the role of *O*-GlcNAcylation in p53 regulation and the codependent relationship of p53 *O*-GlcNAcylation and phosphorylation.

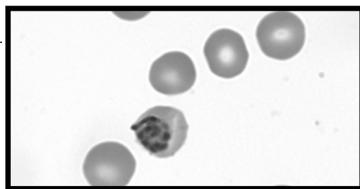
Studies using the human breast cancer cell line MCF-7 revealed that cells treated with STZ and doxorubicin (Dox), a DNA-damaging agent that stabilizes p53, retained increased levels of p53 and that the amount of *O*-GlcNAcylation specifically on p53 was also increased. In addition, the interaction between p53 and Mdm2, an E3 ubiquitin-ligase, was significantly reduced in the presence of STZ and Dox; this suggests that *O*-GlcNAcylation may lead to decreased p53 ubiquitination and, thus, increased stability. Although phosphorylation of p53 at certain sites does stabilize the protein by preventing Mdm2-mediated ubiquitination, paradoxically, phosphorylation of Thr155 actually facilitates p53 ubiquitination. Interestingly, the authors found that STZ greatly reduces Thr155 phosphorylation, but the link between the states of Thr155 phosphorylation and p53 *O*-GlcNAcylation was still ambiguous. Enter the yin-yang. Just down the polypeptide street of Thr155 resides Ser149, which the authors determined was indeed a site of *O*-GlcNAcylation. Further investigation revealed that the presence of an *O*-GlcNAc at Ser149 results in dramatically reduced phosphorylation of Thr155, which in turn suppresses ubiquitin-dependent proteolysis and stabilizes the protein. These findings provide intriguing clues into the complex contributions that post-translational modifications play in p53 regulation. **EG**



From Klein, D. J., and Ferré-D'Amaré, A. R., *Science*, Sept 22, 2006 (DOI: 10.1126/science.1129666). Reprinted with permission from AAAS.

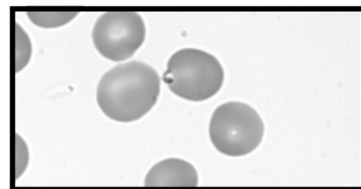
MetAP—A Match for Malaria?

Malaria, which is derived from medieval Italian meaning “bad air”, is the cause of >1 million deaths annually. Though the disease is not technically transmitted through “bad air” but through the parasite *Plasmodium falciparum* courtesy of its insect host, the human-health threat that it poses has certainly filled the airways around the globe. Building on recent findings that inhibitors of methionine aminopeptidases (MetAPs) inhibit the growth of *P. falciparum* in culture, Chen *et al.* (PNAS 2006, 103, 14,548–14,553)



now report the discovery of a new class of MetAP inhibitors that display exquisite selectivity for the malarial enzyme and show antimalarial activity in mouse malaria models.

MetAPs are highly evolutionarily conserved metalloproteases tasked with removing the N-terminal initiator methionine during protein synthesis, distinguishing them from other proteases and from other antimalarial drug targets. Examination of the *P. falciparum* genome



Chen, X., *et al.*, *Proc. Natl. Acad. Sci.* 2006, 103, 14,548–14,553. © 2006 National Academy of Sciences, U.S.A.

revealed four genes homologous to yeast and human *MetAP* genes, one of which was overexpressed in *Escherichia coli* and purified in quantities suitable for high-throughput screening for small-molecule inhibitors. One structural class containing a 2-(2-pyridinyl)pyrimidine core inherently displayed encouraging selectivity over the other *P. falciparum* and human *MetAP* enzymes and promising activity against *P. falciparum* in erythrocyte culture. Structure–activity analysis led to the identification of the compound XC11, which inhibited proliferation of both drug-sensitive and drug-resistant *P. falciparum* parasites in culture while maintaining low toxicity against human fibroblasts. In addition, when mice infected with both drug-sensitive and drug-resistant *Plasmodium* strains were treated with XC11, their average survival time was extended. The authors propose that the superior metal-chelating ability of the 2-(2-pyridinyl)pyrimidine functionality may be a key component in the mechanism of inhibition of XC11, which may incorporate a third metal ion into the enzyme active site similar to that observed with other *MetAP* inhibitors. This class of compounds is an exciting starting point for the development of potential new antimalarial agents. **EG**

Natural Ribozymes Get a Closer Look

continued from page 540

identical between many bacterial species; this shows that this unique mechanism spans a long evolutionary history.

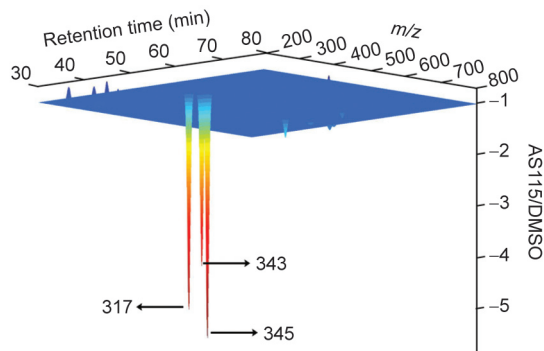
Do self-cleaving ribozymes exist in humans? K. Salehi-Ashtiani *et al.* have developed a clever new selection strategy for finding sequences in the human genome encode ribozymes (*Science* 2006, 313, 1788–1792). The scheme monitored trillions of genomic segments for instability under physiological salt conditions and the common ribozyme assistant, magnesium. Out of this complexity emerged four sequences that can self-cleave in a test tube. The authors chose to characterize one ribozyme that lies in the intron of the RNA-binding protein, CPEB3. The sequence showed high conservation among the sequenced mammalian genomes. Additionally, cross-species nucleotide changes often occurred in matched pairs, a hallmark of RNAs with functional secondary structures. Interestingly, this ribozyme appears to use a similar general acid–base catalytic mechanism and adopts a similar fold to the robust ribozymes from human hepatitis delta virus.

One lesson learned from the riboswitches and the *GlmS* ribozyme is that cofactor binding can affect RNA activity. Because the search for human ribozymes took place without cofactors, other ribozymes may be hidden among human mRNAs, awaiting the right partner to unlock a hidden talent. The RNA world hypothesis has pushed researchers to select new ribozymes to help along the chemical reactions that took place in a primordial soup; however, these two studies should prompt a new interest in looking at RNA sequences that nature already selected. **JU**

MAGE Magic

Though no scientific tie exists between MAGEs (monoalkyl-glycerol ethers) and the English term “mage” from which the word “magic” is derived, some sort of wizardry would definitely come in handy to elucidate the complex biochemical networks associated with MAGE biosynthesis and function. Fortunately, the use of a multidimensional profiling strategy combining activity-based proteomics and metabolomics by Chiang *et al.* (*Chem. Biol.* 2006, 13, 1041–1050) provided the magic touch for characterization of the enzyme KIAA1363, whose activity has been implicated in several types of cancer, as a 2-acetyl MAGE hydrolase.

The known reactivity of KIAA1363 with fluorophosphonate proteomics probes designated this otherwise uncharacterized enzyme as a serine hydrolase, and this facilitated the design of a selective, carbamate-based covalent inhibitor of the enzyme termed AS115. An LC/MS platform was used to evaluate the global metabolite profiles in cells treated with AS115, and the authors observed that the levels of certain MAGEs were dramatically reduced compared with untreated cells. Additional biochemical characterization of cancer cell lines and KIAA1363-transfected cells confirmed that the enzyme catalyzes the hydro-



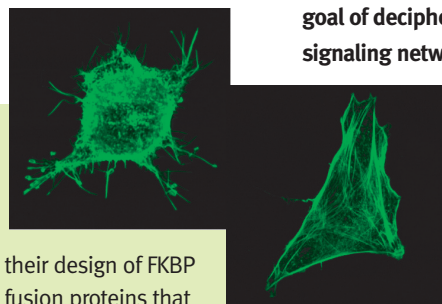
Reprinted from *Chem. Biol.*, 13, Chiang, K. P., *et al.*, An Enzyme that Regulates Ether Lipid Signaling Pathways in Cancer Annotated by Multidimensional Profiling, 1041–1050, copyright 2006, with permission from Elsevier.

lysis of 2-acetyl MAGEs to MAGEs. Efforts to place this enzyme within cancer cell lipid signaling networks led to the hypothesis that KIAA1363 links the platelet-activating factor (PAF) and lysophospholipid pathways. Indeed, treatment of cells with AS115 protected 2-acetyl MAGE from hydrolysis, and this led to accumulation of PAF (the phosphorylcholy derivative of 2-acetyl MAGE) and decreased levels of two lysophospholipids, alkyl-LPC and alkyl-LPA, which are derived from MAGE. Finally, short-hairpin-RNA-mediated interference was used to demonstrate that knockdown of KIAA1363 expression significantly reduced the ability of cancer cells to promote tumor growth in mice, an effect that was correlated with impairment of their migration capacity. Notably, cell migration was rescued by addition of LPA, pointing to a mechanism by which KIAA1363 could influence the aggressive properties of cancer cells. In addition to augmenting our understanding of lipid signaling in cancer cells, the general approach utilized in this work can be applied more broadly toward realizing the ultimate goal of deciphering the incredibly intricate metabolic and signaling networks in cells, magic notwithstanding. EG

Exploiting Instability

Manipulating protein function at the genetic or protein level is a powerful method for exploring biological pathways, but it is often easier said than done. Whereas strategies such as transcriptional regulation of protein expression, RNA interference, or small-molecule inhibition are frequently used to control protein function, inherent challenges unique to each method encumber their utility. Now, Banaszynski *et al.* (*Cell* 2006, 126, 995–1004) present a general strategy for controlling the stability, and consequently the function, of proteins. The approach involves small molecules and encompasses a level of versatility not available in other methods.

The authors exploit the well-studied interactions between the cytosolic protein FKBP and synthetic small-molecule FKBP ligands in



their design of FKBP fusion proteins that are stable only in the presence of a specific small-molecule ligand. To identify appropriately destabilized FKBP domains, the authors expressed a library of FKBP mutants fused to yellow fluorescent protein (YFP) in cells, and they selected and characterized variants that exhibited reduced fluorescence in the absence of an FKBP ligand but increased fluorescence in the presence of the ligand. Kinetic and mechanistic analyses using a morpholine-containing FKBP ligand termed Shield-1 (Shld1) revealed that FKBP-YFP protein levels were stabilized in a dose-dependent manner within 24 h after expo-

Reprinted from *Cell*, 126, Banaszynski, L. A., *et al.*, A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules, 995–1004, Copyright 2006, with permission from Elsevier.

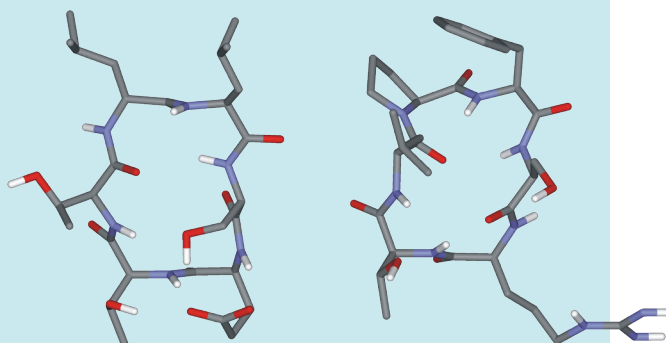
sure. In addition, rapid, proteasome-mediated degradation of the protein upon withdrawal of the ligand correlated with the extent of destabilization conferred by the mutation. Shld1-dependent control of FKBP fusion proteins was demonstrated in several cell lines and with a variety of proteins, substantiating the versatility of the method. Notably, when levels of proteins that cause known morphological changes in cells were manipulated in a Shld1-dependent manner, the cells adopted the predicted phenotypes. This compelling strategy can be applied in a variety of biological settings, including controlled investigation of proteins that exhibit dominant-negative phenotypes or the creation of knockin mice expressing Shld1-dependent genes. EG

Glycomimetics Come Full Circle

Carbohydrates are key recognition elements for a wide range of cellular communication processes, including development of the nervous system. Despite much recent progress in the field of oligosaccharide synthesis, generation of carbohydrate-containing compounds for biological exploration or therapeutic applications remains an arduous task. Bächle *et al.* (*Angew. Chem., Int. Ed.* 2006, 45, 6582–6585) now describe the design and synthesis of novel glycomimetic cyclic peptides that stimulate neurite outgrowth, pointing to a promising approach for developing novel therapeutics for neuron regeneration.

The human natural killer cell glycan HNK-1 is known to play a role in regulation of the nervous system, and the presence of this carbohydrate enhances neurite outgrowth. Two previously identified linear peptide sequences known to bind to an HNK-1 monoclonal antibody were used to design a library of 26 cyclic hexapeptides, each of which contained one D-amino acid. The cyclization of the backbone and the incorporation of the D-amino acid were purposefully included in

the design with the intent of reducing the conformational flexibility of the peptides, which could promote enhanced binding affinity and biological activity. Surface plasmon resonance was used to evaluate the binding properties of the peptides to the HNK-1 antibody, and selected peptides were investigated further for their ability to promote neurite outgrowth. Two cyclic hexapeptides were found to stimulate



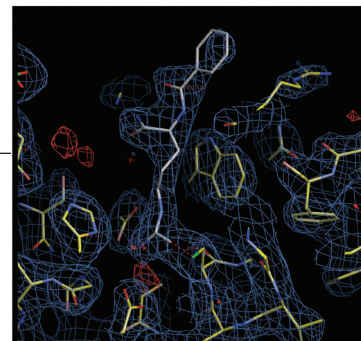
Reproduced with permission from *Angew. Chem., Int. Ed.* from Wiley-VCH, Bächle, D., *et al.*, 2006, 45, 6582.

neurite outgrowth and, notably, to enhance survival of motor neurons. The solution structures of the two peptides were characterized by NMR spectroscopy revealing intricate details of how their specific conformations may contribute to antibody binding. This study demonstrates the power of restricting peptide conformation to extrude enhanced biological activity. **EG**

Deciphering Deimination

Many proteins are known to contain the amino acid citrulline, which is generated upon the deimination of arginine residues by protein arginine deiminases (PADs). Protein citrullination appears to be involved in several important cellular processes, including apoptosis, gene transcription, and differentiation, but its precise functions are not well-defined. In addition, evidence shows that dysregulation of PAD4 may play a part in the development of rheumatoid arthritis (RA), and this makes PAD4 inhibitors attractive potential RA agents. In an effort to further elucidate PAD4 biology and create potential PAD4-based therapeutics, Luo *et al.* (*Biochemistry*, 2006, 45, 11,727–11,736) now report the functional and structural characterization of several analogues of the potent PAD4 inhibitor F-amidine.

F-amidine (*N*- α -benzoyl-*N*⁵-(2-fluoro-1-iminoethyl)-L-ornithine amide) irreversibly inhibits PAD4 in a calcium-dependent manner by reacting with an active site cysteine residue. Several additional inhibitors, modeled after the F-amidine structure, were synthesized that varied in the length of the side chain leading to the leaving group (containing either two or four methylene units) and in the identity of the leaving group (fluoride, chloride, or hydrogen). Structure–activity analysis of the compounds revealed that proper positioning of the reactive group in the active site and the presence of an effective leaving group are essential for potent, irreversible inhibition of the enzyme. The compound Cl-amidine was found to be the most potent inhibitor and



Reprinted with permission from Luo, Y., *et al.*, *Biochemistry*, DOI: 10.1021/bi061308i, Copyright 2006 American Chemical Society.

preferentially inactivated the calcium-bound form of the enzyme. In addition, a mammalian two-hybrid assay was used to confirm that Cl-amidine can specifically prevent the interaction between the nuclear receptor coactivator GRIP1 and the deiminated form of the GRIP1 binding domain of p300 in cells. Finally, the crystal structure of the PAD4–F-amidine–calcium complex was determined, and subtle conformational insights into the mechanism of PAD4 inactivation by F-amidine and related compounds were revealed. These detailed studies provide important structural information for the effective design of PAD4-binding compounds for future biological and therapeutic applications. EG

Slice, Dice, and Splice?

During the normal immune response, killer T lymphocytes recognize and kill extracellular antigens and selectively engulf cells that appear abnormal or infected. The 20S proteasome, responsible for routine turnover of cellular proteins, also liberates peptide fragments from foreign agents that then mark cells for destruction. A new study indicates that making these peptide flags can be much more complicated than a simple dicing of infectious proteins. Warren *et al.* (*Science* 2006, 313, 1444–1447) went looking for the peptide antigen causing T-cells to see other human cells as targets, an unfortunate phenomenon in organ or graft rejections. After a large gene screen, they zeroed in on a peptide of just 20 amino acids. Some individuals have a single amino acid change in this peptide, which results in an immune response. When this exact synthetic peptide caused no response, the authors tried putting the peptide inside the cell, and a potent response was restored. They

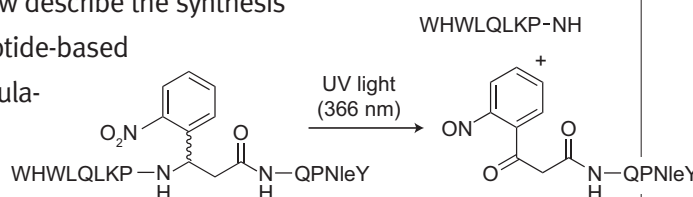
found that the peptide needs to encounter the proteasome first to generate the flag that sparks a T-cell attack. Using a series of peptides and immune assays, they found a surprising result. As shown previously, the proteasome can splice a longer peptide into a shorter one, removing several amino acids and joining two discontinuous fragments. But this time, the peptides were joined in an unconventional fashion. Two peptides were joined in a reverse fashion with respect to how they were coded and synthesized in their original protein. The carboxy end of the downstream peptide spliced with the amino end of the upstream peptide. The authors demonstrate that this strange transpeptidation reaction takes place after an acyl-enzyme intermediate is formed within the proteasome. This study brings to light a new function for the proteasome and dramatically expands the repertoire of short peptides that can be made in the cell. **JU**

Photocleavable Pheromones

Cell division is an extraordinarily complex process in all organisms, and yeast is an attractive model for probing the mechanisms involved in the cell cycle. The yeast peptide pheromone α -factor activates the mating pathway and arrests the cell cycle in the G1 phase, and its activity is regulated in part by secretion of a protease that degrades the pheromone. Use of α -factor to cause cell cycle arrest is often exploited to explore cell division pathways. However, treatment with the native pheromone does not provide practical temporal control over the signaling pathways that are activated upon exposure. Parker *et al.* (*Angew. Chem., Int. Ed.* 2006, 38, 6322–6325) now describe the synthesis and biological application of a photocleavable peptide-based pheromone analogue that enables external manipulation of the yeast mating signal.

Examination of the α -factor peptide sequence and structural information regarding binding

to its G-protein-coupled receptor revealed clues for generating a photocleavable α -factor analogue whose activity could be controlled by UV irradiation. Strategic insertion of the D-enantiomer of (2-nitro)- β -phenylalanine into the flexible region of the peptide, along with replacement of a methionine with norleucine to increase peptide stability, resulted in the generation of a photocleavable peptide capable of inducing the mating response. Treatment of yeast with the pheromone analogue resulted in characteristic morphological changes associated with activation of the mating pathway, growth inhibition related to G1 arrest, and complete G1 arrest of a yeast culture. Notably, UV irradiation of the culture at a wavelength that does not induce DNA damage resulted in significant degradation of the pheromone analogue and release of the cells into the cell cycle. This strategy provides an innovative tool for exploration of the yeast cell cycle, especially in the context of early gene expression analysis during cell cycle recovery and microscopy experiments on solid media that were not previously practical. **EG**



CONGRATULATIONS

Congratulations to Roger D. Kornberg for winning the 2006 Nobel Prize in Chemistry for his study of the molecular basis of eukaryotic transcription. Congratulations also to Andrew Z. Fire and Craig C. Mello for winning the 2006 Nobel Prize in Medicine for their discovery of RNA interference-gene silencing by double-stranded RNA.

Opportunities for Chemical Biologists: A View from the National Institutes of Health

Jeremy M. Berg*

National Institute of General Medical Sciences, 45 Center Drive, 2As-12, Bethesda, Maryland 20892

The most exciting and vibrant areas of research often lie at the interfaces between disciplines that have traditionally been separate. In the case of the chemical and biological sciences, such interfaces have been fruitfully explored for more than a century. The fields of pharmacology, biochemistry, and biophysical chemistry are relatively mature, yet they are still quite active and full of challenging problems and opportunities for new discoveries. These fields are integral to biomedical research and have had a tremendous impact on human health.

Chemical biology is a new variant at this interface. The definition of this interdisciplinary field is evolving as more researchers become actively engaged and more progress is made. Chemical biology is an outgrowth of the revolutionary advances in molecular biology that have taken place over the past two decades. The ability to manipulate DNA molecules and their expression at will, made possible by a combination of advances in our understanding of central biological pathways and enzymes as well as powerful and inexpensive methods of chemical synthesis, has linked biological sciences and chemistry in unprecedented ways. This marriage of chemical synthesis and molecular and cell biology is a major theme in chemical biology. Other chemical biology research involves the application of concepts from biology, such as the use of Darwinian selection and evolving systems, to chemistry. Biomedical science and public health can benefit

greatly from technical advances such as new synthetic methods and novel instrumentation as well as from completely new concepts that are developed as components of chemical biology research.

The National Institutes of Health (NIH), including its National Institute of General Medical Sciences (NIGMS), has been supporting research in chemical biology for many years. The primary mechanism for support of this research is the investigator-initiated grant, typically an R01 grant. More than 60% of the \$1.94 billion NIGMS budget this year is devoted to R01 grants or their equivalents. Approximately one-third of these grants have a significant chemical component. A more complete description of the NIGMS budget is available at <http://publications.nigms.nih.gov/loop/20060616.html#1>. Through investigator-initiated grant applications, researchers are free to submit ideas of their choosing relevant to the broad NIH mission and to compete for available funds through peer-review processes.

NIH highlights areas of particular interest by issuing program announcements. One announcement of relevance to chemical biology is called Metals in Medicine (<http://grants.nih.gov/grants/guide/pa-files/PA-05-001.html>). This announcement, from NIGMS and several other NIH components, solicits research in specific areas at the inorganic chemistry–medicine interface, including metal metabolism and regulation and the interactions of inorganic complexes with living systems. No funds are set aside for program announcements, and applications

Image courtesy of the National Institutes of Health



*Corresponding author,
bergj@mail.nih.gov.

Published online October 20, 2006
10.1021/cb6003993

This article not subject to U.S. Copyright.
Published 2006 by American Chemical Society

The greatest contributions to emerging fields such as chemical biology come from individual scientists.

compete directly with unsolicited applications for support.

Another NIGMS initiative relevant to chemical biology is a network of centers for Chemical Methodologies and Library Development (CMLD) (www.nigms.nih.gov/Initiatives/CMLD). Four active CMLD centers are developing new methods for chemical library synthesis and collaborating with other chemists and biologists on applications of these libraries.

NIGMS also supports two large-scale collaborative project awards ("glue grants") that have substantial chemical components. Glue grants promote integrative and team approaches to important, complex biological problems. The Consortium for Functional Glycomics (www.functionalglycomics.org/static/consortium) is directed toward understanding the role of carbohydrate-protein interactions at the cell surface in cell-cell communication, and the Lipid Metabolites and Pathways Strategy consortium (www.lipidmaps.org) addresses global changes in lipid metabolites that occur during biological processes such as signaling. These consortia are actively developing and disseminating important new tools and resources for broad scientific communities, including chemical biologists.

Research training programs also play a significant role in promoting the development of new fields like chemical biology. NIGMS has supported a chemistry-biology interface Ph.D. training program since 1993, with grants to 20 institutions at present (www.nigms.nih.gov/Training/Mechanisms/NRSA/InstPredoc/#chemistry). These training programs involve faculty and students from both chemistry-related and biological science departments and promote student exposure to scientific concepts and cultures from both sides of the chemistry-biology interface. Interactions among students from different backgrounds play important roles in these training programs. Individual fellowships are available to support postdoctoral training in chemical biology.

Chemical biology is well represented in a more recent set of initiatives collectively called the NIH Roadmap for Medical Research (<http://nihroadmap.nih.gov>). The Roadmap was developed with broad input from the scientific community to focus on areas of considerable relevance to the NIH mission that were not being adequately addressed by existing programs supported by single institutes or groups of institutes. The Roadmap is thus not a strategic plan, but rather an attempt to respond to perceived gaps in the NIH portfolio.

The Roadmap's Molecular Libraries and Imaging Initiative (MLII) (<http://nihroadmap.nih.gov/molecularlibraries>) has the largest chemical biology component. The backbone of this initiative is the Molecular Libraries Screening Center Network (MLSCN) and an associated NIH Molecular Libraries Small Molecule Repository. The goal of the MLSCN is to offer the public sector access to the large-scale screening capacity necessary to identify small molecules that can be optimized as chemical probes to study the functions of genes, cells, and biochemical pathways. Ten pilot-scale centers are currently in operation and are soliciting potential assays from the scientific community. The activities of the MLSCN and related initiatives are integrated through a publicly accessible database called PubChem (<http://pubchem.ncbi.nlm.nih.gov>). This database allows researchers to navigate parts of the universe of small molecules with links to relevant entries in other biomedical databases developed and maintained by NIH's National Center for Biotechnology Information.

Other components of the MLII include programs in New Methodologies for Natural Products Chemistry, Pilot-Scale Libraries for High-Throughput Screening, and Novel Preclinical Tools for Predictive ADME-Toxicology and the development of new tools and methods for high-resolution cellular imaging. See <http://nihroadmap.nih.gov/molecularlibraries> for more details about

funded research and potential funding opportunities.

Additional Roadmap activities that have significant relevance to chemical biology include Nanomedicine; High-Risk Research (through the NIH Director's Pioneer Award program); and Building Blocks, Biological Pathways, and Networks. Because the Roadmap programs cut across all of NIH, they are substantial in scope and depth, yet they have only a modest impact on the budgets of individual institutes. In fiscal year 2006, 1.2% of the overall NIH budget goes to support Roadmap activities, including 0.9% of the NIGMS budget.

Regardless of funding mechanisms, the greatest contributions to emerging fields such as chemical biology come from individual scientists. This past May marked the passing of R. Bruce Merrifield, a scientist whose seminal discoveries greatly influenced chemical biology and many other fields. Solid-phase synthesis, first applied to peptides, then to oligonucleotides, and now to a wide range of chemicals, transformed both chemical and biomedical research. Without this fundamental chemical insight, much of the now-routine operations in molecular biology, such as amplification of DNA by the polymerase chain reaction and DNA sequencing, would be vastly more difficult. In fact, many of the activities now associated with chemical biology would not only be much more difficult, they probably would never have even been conceived. As a science administrator and a scientist whose own research lies at the chemistry-biology interface, I look forward to many new discoveries and advances that can be expected as chemical biology continues to evolve.

19th International Symposium on Medicinal Chemistry

In late August, hundreds of chemists gathered on the eastern edge of Europe, some sipping raki while looking over the Bosphorus Strait to the Asian continent. More than 900 people attended the 19th International Symposium on Medicinal Chemistry in Istanbul from August 29 to September 2, 2006, hosted by the European Federation for Medicinal Chemistry (EFMC) and the Turkish Association of Pharmaceutical and Medicinal Chemistry. The motto for the conference, held for the first time in Turkey, was “where continents meet”. It could easily have been “where disciplines meet”. Roberto Pellicciari, head of the EFMC, observed an “unbelievable shift” in the kinds of topics that interest medicinal chemists. “There’s been a completely new understanding for medicinal chemistry coming from chemical biology.”

Describing even a modest fraction of the proceedings is impossible in this space. The symposium boasted 9 full-conference lectures, >100 talks in concurrent sessions, and a large poster session. As expected, classical medicinal chemistry talks were plentiful, particularly for the most alluring protein targets such as kinases, epigenetic regulators, and G-protein-coupled receptors. Researchers’ talks usually began with a small-molecule hit, which was then divided into pharmacophores for iterative rounds of modifications. Repeatedly, though, chemistry and biology bled into each other, with the traditions of one field transforming into powerful new tools in the context of the other.

Chemists’ Natural Tutor. Appropriately, this year’s conference marked the debut of a new prize: The GlaxoSmithKline Award for Outstanding Achievement in the Field of

Chemical Biology. Prize winner Herbert Waldmann, head of the chemical biology department at the Max Planck Institute in Dortmund, Germany, explained that medicinal chemists face a surfeit of choices in their quest for drug leads: an incomprehensibly large number of molecules are possible, but few can be synthesized and tested over the course of a project. He showed how to take nature’s lead in deciding which molecules to make. After all, nature itself has sampled only a tiny fraction of diversity space over billions of years of evolution, but it has done so more effectively than any scientist. A quick inspection of major classes of natural products shows that they bind to multiple proteins; that’s not surprising, said Waldmann, given that natural products are synthesized by proteins and so must be able to interact with them. Moreover, protein fold structures are often conserved across multiple proteins, even where amino acid sequences have diverged. Inspired by bioinformatics, Waldmann and colleagues turned to cheminformatics to identify common scaffolds in ~190,000 natural products, creating classification schemes similar to phylogenetic trees (1).

Their small-molecule analysis revealed three “kingdoms” of natural products: N-heterocycles, O-heterocycles, and carbocycles, more than half of them with only two to four rings. Waldmann and colleagues analyzed scaffolds within these groups and then asked whether these could inspire simpler, synthetically more accessible structures that retain biological activity. Although this is not a new concept, Waldmann and colleagues decided to pay strict attention to scaffolds seen in natural products rather

Published online October 20, 2006

10.1021/cb600409p CCC: \$33.50

© 2006 by American Chemical Society

“The first targeted inhibitor is really a dual inhibitor,”
concluded Shokat.



Image courtesy of Dan Etimison

than those that are most synthetically expedient.

A classification scheme in hand, Waldmann and colleagues turned to bioinformatics for information rather than inspiration (2). They identified a common binding core in three apparently disparate targets: *cdc25* (a phosphatase that regulates cell division), 11- β -hydroxysteroid dehydrogenase (a reductase that boosts levels of cortisol), and acetylcholinesterase (an enzyme that breaks down a neurotransmitter). Next, they made 500 compounds based on the scaffolds deemed most likely to bind these proteins. A functional screen of these revealed 9 targets with micromolar specificity, which inhibited one enzyme at least 100-fold more tightly than another. They achieved similar results with another set of proteins and compounds.

The molecules may not be impressive hits yet, says Waldmann, but the approach can now be used to identify tool compounds for exploring biology and starting points for medicinal chemistry programs, without the need to maintain huge screening libraries of compounds. The approach, called biology-oriented synthesis, has favorable hit rates compared with other library design approaches, such as those based on known ligands, homology modeling, and diversity-oriented synthesis (3).

Chemically Interrogating Kinases.

Inverting Waldmann's approach, Kevan Shokat of the University of California, San Francisco, used chemistry to probe biology. Inside cells, ~520 kinases regulate the activity of >10,000 proteins by using a common chemical reaction (phosphorylation). This sophisticated cellular “wiring” encompasses the equivalent of logic gates,

with several “inputs” (phosphorylation events) triggering single “outputs”, like cell division. Shokat used chemical tools to map this wiring both by inhibiting specific kinases and by identifying kinase substrates. In each case, he reengineered kinases and their common substrate, ATP. His lab mutated kinase genes to enlarge the ATP binding pocket and added a benzyl group to an ATP mimetic (4). The mimetic fits the engineered kinase but no other. Depending on the mimetic, these “bioorthogonal” molecules either inhibit the mutant kinase specifically or permit the mutant kinase to catalyze a non-natural reaction.

Shokat and collaborators have genetically engineered cells and even mice that express these mutant kinases. These function normally until a bioorthogonal inhibitor is introduced and allow subtle, precise probing of enzymes' functions. The effects of chemically knocking out a kinase can be observed within seconds to minutes, a much faster time scale than that of inducible genetic knockouts. Even better, the kinase of interest is still physically present in the cell and can still form protein complexes. This means researchers can home in on a kinase's phosphorylation activity separate from other functions that these proteins serve in the cell.

Shokat's team found results that defied conventional wisdom in both mice and cells. Using the bioorthogonal system to specifically inhibit the kinase BCR-ABL did not have the same effect as adding Gleevec, an approved cancer drug that inhibits this enzyme. However, when the bioorthogonal inhibitor was administered along with an antibody against another kinase, c-Kit, researchers observed the same results normally seen in unmodified mice treated only with Gleevec (5). “The first targeted inhibitor [Gleevec] is really a dual inhibitor,” concluded Shokat.

To identify kinases' phosphorylation substrates, Shokat started with the cyclin-dependent kinases (CDKs). These proteins regulate cell division and are targeted in multiple drug-development programs, but many substrates remain unidentified. First, Shokat and colleagues developed a bioorthogonal ATP analogue. In addition to the specificity-conferring benzyl group, the analogue contains a phosphorothioate in place of the terminal phosphate group. The researchers genetically engineered cells to produce CDKs with enlarged ATP pockets and added the bioorthogonal ATP analogue.

To find each kinase's substrate, Shokat's group looked for proteins that had been “phosphorylated” with the phosphorothioate group. An analysis of CDK1 and its binding partner, CyclinB, identified 70 substrates; 24% of these were known substrates, 23% were previously known to be phosphorylated by an unknown kinase, and 53% were completely novel (6).

In another kinase-centered talk in the Chemical Genomics section (cosponsored by ACS), Shripad Bhagwat of Ambit Biosciences demonstrated how chemical biologists can help medicinal chemists optimize compounds and make decisions about which targets to pursue. Ambit can quickly assess compounds' activities against hundreds of kinases with a screen that includes all branches of the kinome. KinomeScan uses phage display to express kinases,

which can bind to known ATP-site binding ligands that have been immobilized on a solid support. Compounds that bind the kinase being tested will compete for binding and prevent the kinase from binding to the immobilized ligand. The data are visualized as small-molecule kinase interaction maps that illustrate which kinases a compound is capable of binding. Bhagwat and colleagues assessed the promiscuity profiles of kinase inhibitors for which clinical data are available. Systematic profiling revealed that overall selectivity among clinical compounds varies widely and that toxicity cannot necessarily be directly predicted based on promiscuity. Ambit is using this technology to reverse the natural order of drug discovery and to guide medicinal chemists in the choice of compounds with selectivity appropriate for each development program. Instead of choosing a target and screening for hits, Ambit screens its compound library against its kinase panel and then picks its target, continuing to assess selectivity as compounds are advanced.

Michael Eck of Harvard Medical School and the Dana-Farber Cancer Institute used structural biology to argue that some individual proteins are better viewed as multiple targets. He presented the structure of focal adhesion kinase (FAK) in its inactive form, demonstrating that the 4.1, ezrin, radixin, and moesin (FERM) domain of FAK hugs the enzyme's active site, sterically preventing autophosphorylation and activation. Thus, medicinal chemists seeking to reduce the kinase's activity might do well to look for compounds that stabilize the inactive, closed state of the protein in addition to following the more obvious path of seeking inhibitors of the kinase's active site.

Besides movement of the protein domains themselves, said Eck, kinases present a moving target because they are frequently mutated in cancer. He used crystallography and modeling to dissect mutated and wild-type forms of epidermal

growth factor receptor (EGFR), implicated in several cancers, and the target of the small-molecule cancer drugs Tarceva and Iressa. At least two mutations found in lung cancers (L858R and G719S) activate the kinase by destabilizing the inactive state of the enzyme, shifting the equilibrium toward the active state. These mutations do not dramatically affect the conformation of the active site but can nevertheless affect the binding of inhibitors. Because of this, Eck suggested that the wild-type and some EGFR mutants should be considered different targets. In lung cancers in which EGFR is activated by mutation, compounds that specifically target the inactive conformation of the EGFR are likely to be ineffective. However, in other cancers in which the wild-type enzyme is overexpressed, targeting that state could prove fruitful.

Pin the Drug on the Target. Specificity versus promiscuity was a recurring theme for both target and disease categories. Finding the right balance for schizophrenia is particularly difficult: receptors for multiple neurotransmitters are implicated, multiple classes of receptor exist for each neurotransmitter, and desirable targets are a matter of debate. Andrew Payne of Glaxo-SmithKline (GSK) noted that marketed schizophrenia drugs target as many as 15 different receptors each. Not only do these molecules fail to treat several symptoms of schizophrenia, they have serious side effects. Payne's group felt that many of these activities are unnecessary and undesirable and set about discovering compounds that activate five, and only five, specific proteins: two dopamine receptors (D2 and D3) and three serotonin receptors (5-HT2A, 5-HT2C, and 5-HT6). The team found a compound that loosely fit these characteristics by interrogating GSK's database and then fine-tuned divergent structure-activity relationships to create an advanced lead that shows a promising *in vivo* profile.

Antonello Mai of the Università degli Studi di Roma in Italy described a similarly

complex picture for the much younger science of epigenetics, or the regulation of gene expression through modifications to DNA and to chromatin structure. DNA itself can be modified by DNA methyltransferases, whereas chromatin structure can be altered through phosphorylation, ubiquitination, ribosylation, methylation, and acetylation/deacetylation. Four classes of histone deacetylases (HDACs) each comprise several enzymes; inhibitors of several are in clinical trials against cancer. Mai reported several chemical series of HDAC inhibitors, including the first inhibitors specific for Class 2 HDACs, a group of five enzymes expressed in heart, muscle, and brain. To his surprise, these compounds showed no ability to inhibit the cell cycle, induce apoptosis, or cause differentiation in leukemia cells and thus may represent a new chemical tool to fine-tune the study of these proteins.

Even the most successful medicinal chemistry program will fail if carried out against an inappropriate target or without sufficient focus on the biological context of both the target and the disease to ensure that relevant mechanisms of action are pursued. Lamenting the fact that "large numbers of compounds lack contextual information," Giovanni Gaviraghi of Siena Biotech in Siena, Italy, discussed targets in Alzheimer's disease (AD). Three different secretases are involved in processing the amyloid precursor protein that forms the neural plaques characteristic of the disease, but all are problematic either chemically or biologically. Discovering appropriate compounds for α - and β -secretases is difficult because the activity of the former should be enhanced rather than inhibited and because the latter's flexible active site thwarts traditional medicinal chemical approaches. The challenges for γ -secretase are biological: compounds that inhibit this enzyme are likely to inhibit the NOTCH pathway, which is necessary for maintaining neural stem cells, proper cell differentiation,

Sometimes, it is not identifying targets but measuring compounds' effects against target families that is challenging.

and other functions. Tellingly, both Merck and Lilly have advanced γ -secretase inhibitors in the clinic; for the proteins posing chemistry challenges, β -secretase programs lag behind, and α -secretase is widely considered unenhanceable.

Several promising AD targets remain relatively unexplored. One example is the protein Dickkopf1 (DKK1), which is overexpressed in AD and acts to inhibit the Wnt signaling pathway. Cell studies show that inhibiting DKK1 or rescuing Wnt can prevent neural death and point to a potential therapeutic. DKK1 and importantly the Wnt pathway may be key areas for AD research. As Gaviraghi indicated, phenotypic screening will likely be most effective at untangling complex networks such as the Wnt signaling pathway and will require an integrated chemistry and biology approach.

Sometimes, it is not identifying targets but measuring compounds' effects against target families that is challenging. Andreas Termin, director of chemistry at Vertex Pharmaceuticals in San Diego, CA, described the creation of an assay to detect selective sodium channel modulators. The nine subtypes of voltage-gated sodium channels (NaVs) are expressed differentially in the central nervous system (CNS), peripheral nervous system, and skeletal muscle. Currently marketed NaV inhibitors, which are not selective, can relieve epilepsy, pain, and tinnitus but have limiting side effects. Unfortunately, attempts to find selective compounds have been hindered, because NaVs have fast kinetics, which make them intractable to most cell-based functional assays, and because the gold standard for measuring ion channel activity, patch clamping of a neuron, is slow and tedious. Termin estimated that using this method to assess the IC_{50} for all nine NaV subtypes for one compound would take two weeks.

Vertex researchers made this assessment tractable to high-throughput screening that could be completed for their entire compound library in a matter of months. They

filled multiwell plates with cell lines engineered to overexpress NaVs. Fluorescent dyes that change color according to voltage were embedded in the cells. They stimulated cells by placing them in an electric field and detected voltage changes by measuring fluorescence. The assay helped Vertex develop NaV-selective inhibitors just as effective as the neuralgia drug gabapentin in three animal models for pain. Moreover, CNS side effects measured by a rat balancing test were much less severe than those seen for gabapentin and a pan-NaV inhibitor.

But Termin expects the E-VIPR (electrical stimulation voltage/ion probe reader) technique to have wider applications. A surprisingly high number of marketed drugs block NaVs, a fact that might explain some mysterious side effects. E-VIPR could be used to identify potential off-target effects and thus eliminate problematic molecules earlier in the drug-development process.

Functionalizing Antibodies. Medicinal chemists have invaded traditional biology space with techniques to functionalize antibodies. Antibodies conjugated to radioactive isotopes and potent toxins came to the market as anticancer drugs in the 1990s. However, manufacturing and administering the drug remain cumbersome, largely because of problems attaching chemical groups to the complex proteins and ensuring that attached moieties are delivered appropriately. A key goal is to create conjugated antibodies that won't release their cargo until the antibodies are internalized into the cells they bind. This will deliver potent toxins to selected cells without poisoning nontargeted cells. Moreover, because the antibody actively carries the small molecules into the cell, medicinal chemists are freed from typical limitations of size and lipophilicity.

However, the chemistry of the linker between antibody and small-molecule cargo is essential. Because cytoplasm is more acidic than plasma, early versions of linkers

were designed to be cleaved by acid. In practice, however, these linkers frequently released their cargo before internalization, causing toxicity problems and reducing delivery to cancer cells. Peter Senter of Seattle Genetics and Vincent de Groot of Syntarga in the Netherlands described recent efforts to optimize this linker and even introduce new functionality to it.

Seattle Genetics has created a new linker that is cleaved by proteases within the cell and that yields more specific delivery of the potent antimetabolic agent monomethyl auristatin E (MMAE). The company filed an investigational new drug application for anticancer antibody SGN-35 in June 2006. Intriguingly, when an auristatin variant, MMAF, was attached *via* a noncleavable linker, it retained efficacy but showed less overall toxicity. This most likely occurred because of the fact that drug release required complete antibody degradation, a process less efficient in many target cells than in cancer cells. The released drug retained a modified form of the linker, which in the case of MMAF did not diminish its activity. This illustrates how the linker technology is important for controlling toxicity and efficacy, as well as how sensitive a drug is to the means by which it is attached to the carrier protein.

Syntarga is developing combinations of linkers with potent, DNA-damaging drugs from the class of duocarmycins. Company founder and CEO de Groot described a series of self-releasing linkers that could be used interchangeably to attach a wide variety of small molecules to antibodies. He also described a branched linker series designed to allow an antibody to release several-fold the number of drug molecules per linker with a single cleavage site.

Approaches that use antibodies to deliver small molecules to cells rely on antibodies' inherent selectivity. Ana-Isabel Hernández of Utrecht University in the Netherlands took a more radical approach, with a plan to mimic the selectivity regions of an antibody

with completely synthetic molecules. Such synthetic “antibodies” could be used to better understand and perhaps to enhance or alter an antibody’s interaction with its antigen. Hernández and colleagues created a triazacyclophane scaffold capable of holding three of the six hypervariable peptide loops through which antibodies recognize antigens. She used click chemistry to successfully graft the 7- or 8-mer peptides onto this scaffold. They are currently working out purification techniques necessary before further characterization.

The conference ended with an emphasis on the growing connections between biology and chemistry. In the concluding talk, Jan van der Greef of TNO Systems Biology in the Netherlands urged attendees to think less about individual targets and more about biological systems, echoing Pellicciari’s earlier observations that this was already happening. At the conference two years ago, no events were explicitly devoted to chemical biology. This year’s conference had a designated prize, plenary speakers, and sessions. In other words, chemistry and biology are meeting not just where Europe and Asia meet, but in countless labs throughout the world.

—Monya Baker, Science Writer

4. Bishop, A. C., Ubersax, J. A., Petsch, D. C., Matheos, D., Blethrow, J. A., Gray, N. S., Schultz, P. G., Shimizu, E., Tsien, J. Z., Rose, M. D., Wood, J. L., Morgan, D. O., and Shokat, K. M. (2000) A chemical switch for inhibitor sensitive alleles of any protein kinase, *Nature* 407, 395–401.
5. Wong, S., McLaughlin, J., Cheng, D., Zhang, C., Shokat, K. M., and Witte, O. N. (2004) Sole BCR-ABL inhibition is insufficient to eliminate all myeloproliferative disorder cell populations, *Proc. Natl. Acad. Sci. U.S.A.* 101, 17456–17461.
6. Allen, J., Lazerwith, S. E., and Shokat, K. M. (2005) Bio-orthogonal affinity purification of direct kinase substrates, *J. Am. Chem. Soc.* 127, 5288–5289.

REFERENCES

1. Koch, M. A., Schuffenhauer, A., Scheck, M., Wetzels, S., Casaulta, M., Odermatt, A., Ertl, P., and Waldmann, H. (2005) Charting biologically relevant chemical space: A structural classification of natural products (SCONP), *Proc. Natl. Acad. Sci. U.S.A.* 102, 17272–17277.
2. Koch, M. A., Wittenberg, L.-O., Basu, S., Jeyaraj, D. A., Gourzoulidou, E., Reinecke, K., Odermatt, A., and Waldmann, H. (2004) Compound library development guided by protein structure similarity clustering and natural product structure, *Proc. Natl. Acad. Sci. U.S.A.* 101, 16721–17626.
3. Nören-Müller, A., Reis-Corrêa, Jr, I., Prinz, H., Rosenbaum, C., Saxena, K., Schwalbe, H. J., Vestweber, D., Cagna, G., Schunk, S., Schwarz, O., Schiewe, H., and Waldmann, H. (2006) Discovery of protein phosphatase inhibitor classes by biology-oriented synthesis, *Proc. Natl. Acad. Sci. U.S.A.* 103, 10606–10611.

Divide and Conquer: Investigating the Mechanisms behind Mitosis

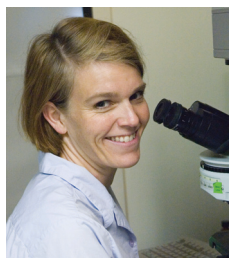


Image courtesy of Deborah Stallford

Soon after a sperm meets an egg, the single fertilized cell splits into two cells, then four, and then eight. Cell division is responsible for producing each of the trillions of cells present in every human body. During adulthood, division supplies replacements for cells lost to age, injury, and disease, but it can also form the basis for illnesses such as cancer. Despite the importance of mitosis in development and medicine, researchers have much to learn about the molecular mechanisms that regulate it. Cell biologist Rebecca Heald of the University of California, Berkeley, is striving to iron out these details. Heald's work concentrates on the mitotic spindle, a structure that is essential for correctly distributing copied chromosomes to daughter cells. Using techniques that blend biology and chemistry, she and her colleagues are identifying molecules and proteins that play major roles in directing this dynamic cell process.

Discovering Division

Heald was born in 1963 in Bellefonte, PA, a town near Pennsylvania State University, where her father Emerson was a postdoctoral fellow. About a year later, she and her family moved 2 hours west to Greenville, PA, home to a small liberal arts school, Thiel College. Heald's father, a physical chemist, taught there for more than 40 years. Some of Heald's earliest memories are of visiting her father and his colleagues in the chemistry building. "The whole building smelled like the lab, and it's a smell I really liked. I liked seeing all the glassware and watching people doing experiments—it seemed like a lot of fun," she recalls. "To this day, I really love a lab. I feel really comfortable there."

Though both of her parents were trained chemists, Heald notes that they rarely pushed her to follow in their footsteps. Rather, her mother and father encouraged Heald and her two older sisters to pursue a variety of interests in addition to science, including math, reading, and writing.

By the time she graduated from high school in 1981, Heald considered math and literature to be two of her strongest subjects. However, soon after she entered Hamilton College, in Clinton, NY, she chose chemistry as her major. "I thought, I can always read in my spare time, but I can't work in a lab in my spare time," she says.

Describing herself as "squeamish," Heald says that she chose chemistry over biology to avoid dissecting a cat, a requirement for all biology majors at Hamilton. However, she notes that she enjoyed her biology classes, filling her schedule with selections such as microbiology and biochemistry.

By the time she received her bachelor's degree from Hamilton, Heald was almost certain that she wanted to pursue science in graduate school, although she had not decided on a field. To gain more experience to guide her decision, she wrote letters to several investigators whose work intrigued her, hoping that one might invite her to work as an assistant. Eventually, she was hired by Sarah Hitchcock-DeGregori, a muscle protein biochemist and cell biologist at Rutgers Medical School in Piscataway, NJ. For the next 2 years, Heald worked with Hitchcock-DeGregori on a project to generate mutant versions of a muscle protein called tropomyosin. The researchers expressed the proteins in bacteria, purified them, and then characterized these mutants. They observed

Profiles provide insights into the lives, backgrounds, career paths, and futures of scientists who serve as Experts on *ACS Chemical Biology's* online Ask the Expert feature. Readers are encouraged to submit questions to the Experts at www.acschemicalbiology.org. The editors will post the most interesting exchanges on the web site.

Published online October 20, 2006

10.1021/cb600414p CCC: \$33.50

© 2006 by American Chemical Society

these proteins' ability to bind to actin and determined whether they promoted ATPase activity in conjunction with other muscle proteins.

By the time Heald left Hitchcock-DeGregori's laboratory, she had two published papers (1, 2) and a sense of the direction she wanted her career to take.

With Hitchcock-DeGregori's helpful mentorship, Heald decided to study cell biology at Harvard Medical School's doctoral program.

When she entered the program in 1987, Heald says that she was unsure of what focus her studies would ultimately take. However, the school used a series of laboratory rotations to provide multiple chances for students to select a mentor. After just two rotations, Heald chose to work in the laboratory of Frank McKeon. "I would choose the same lab all over again," she says.

McKeon's work concentrated on proteins in the nuclear lamina, a network of proteins that line the nuclear envelope. One of his projects involved making mutant versions of lamin genes, transfecting these genes into cells, and then observing the effect of the mutant proteins with a microscope. The visual aspect of McKeon's work was what drew her to join the project, says Heald. "I found looking through the microscope to be really exciting," she adds. "A lot of my science is visually inspired, and just being able to see what something is doing is the most meaningful thing to me."

Generating new lamins harboring mutant phosphorylation sites, Heald eventually found one that wouldn't allow the nuclear lamina to break down during mitosis, in turn preventing the nucleus from dividing once the mitotic spindle had formed (3). An image of this event, which accompanied the paper that Heald and McKeon published describing this work, was featured on the cover of the journal *Cell*. The image, showing a normal cell undergoing mitosis beside one with an altered lamina, "made it very clear that these sites were important

for regulation of the assembly and disassembly of the lamina," says Heald.

Positive Visualization

Heald completed her doctoral degree after working with McKeon for 5 years. Encouraged by her success in studying cell division, she decided to make this process the focus of her work during her postdoctoral fellowship. Seeking a change of scenery, Heald says that she looked for postdoctoral opportunities far from home. Her first choice for continuing her studies was at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany.

"People warned me that it could be bad for my career, that Europe was a backwater that I'd never emerge from again, and that I'd have trouble getting a job afterwards," says Heald. Respectfully ignoring this advice, she joined the laboratory of Eric Karsenti, a cell biologist who studies the regulation of microtubule dynamics. Heald joined Karsenti on a project to determine what factors regulate catastrophe, a process in which microtubules go from a growing to a shrinking state, and an important parameter of microtubule dynamics that is altered during spindle assembly.

Her original plan, says Heald, was to fractionate *Xenopus* egg extracts, using this material to investigate catastrophe. Though other researchers used yeast as a model organism to identify microtubule regulatory factors, Heald appreciated the frog eggs' robust size, which made it possible to generate extracts that could reconstitute mitotic processes and visualize them outside the cell.

Heald notes that many of her early experiments in Karsenti's laboratory ended without success, but in 1995 her work took a positive turn. At that time, she decided to pursue a new angle to investigate mitotic spindle assembly.

She explains that as the mitotic spindle forms, microtubules are stabilized by interacting with chromosomes. Heald wondered

whether an enzyme on chromatin might ultimately be responsible for this activity and sought to isolate it.

To more easily isolate chromatin from the egg extract, she worked with colleagues at EMBL who were coating metal beads with DNA. She reasoned that chromatin would form on beads, which would be much easier to isolate and characterize biochemically. However, she realized that it wasn't yet known whether chromatin itself was the stabilizer or whether other material in chromosomes was responsible for this effect. After adding chromatin-coated beads to the extract, she saw that mitotic spindles could form around the beads. An image of this phenomenon earned her a second journal cover, this time in *Nature* (4). "This image said that you don't need real chromosomes to build the bipolar mitotic spindle—all you need is chromatin," she says.

Heald calls this discovery a "turning point" that opened up numerous questions to steer her future research: Which motor proteins are responsible for generating the mitotic spindle? What in chromatin is responsible for allowing the mitotic spindle to grow? Heald accepted an assistant professorship in 1997 at the University of California, Berkeley, to pursue the answers to these questions.

Fruitful Collaborations

For almost a decade, Heald's laboratory at Berkeley has investigated mitotic spindle regulation. As a nod to her doctoral and postdoctoral work, she and her colleagues have focused on crafting visual experiments that allow them to quickly see their results.

One of Heald's most fruitful projects at her Berkeley laboratory has been in collaboration with her Berkeley neighbor, cell biologist Karsten Weis, who studies nucleocytoplasmic transport. Previous studies had suggested that the machinery that regulates this process is also involved in controlling mitosis. Working together, their

“We could see that there was a physical gradient of this protein even without the nuclear envelope or anything holding it in.”

laboratories used FRET probes to visualize proteins around mitotic chromatin in *Xenopus* eggs. This technique revealed that RanGTP forms a gradient that's enriched near chromatin (5).

That view challenged a general concept of how some proteins exist within cells. “In general, people think that different things are compartmentalized in cells by membranes and other structures,” Heald says. “We could see that there was a physical gradient of this protein even without the nuclear envelope or anything holding it in.”

More recently, Heald and Weis collaborated with Berkeley researcher Ehud Isacoff. Using a technique called fluorescence lifetime image microscopy, Petr Kalab, a postdoctoral researcher working with all three groups, found that a similar RanGTP gradient exists in human somatic cells, a system that's significantly smaller and more complex than *Xenopus* eggs (6).

Heald worked on a different kind of collaborative project with the laboratory of chemist Peter Schultz, a former Berkeley researcher who now works at the Scripps Research Institute in La Jolla, CA. Heald and Schultz joined forces to screen chemical libraries for compounds that inhibit mitotic spindle assembly but don't target microtubules.

In an approach Heald describes as “low throughput,” graduate student Sarah Wignall added ~1500 individual chemicals to *Xenopus* egg extracts to test whether each one inhibited spindle assembly. Chemicals that passed the first test went through a second assay to determine whether they inhibited tubulin polymerization. Finally, compounds that made it through both assays were coupled to an affinity matrix and mixed with *Xenopus* egg extract to determine which egg proteins bound to the compound. The researchers ultimately identified one compound, called diminutol, that bound to an NADP-dependent oxidoreductase. Such a compound

might eventually be used as a basis for anticancer therapy or as part of a genetic screen to determine what other factors are functioning in a redox pathway that could regulate spindle assembly (7).

This September, Heald received the National Institutes of Health Director's Pioneer Award, a grant that will guide more than half her laboratory's efforts over the next 5 years. Heald says that she plans to use these funds to investigate the “really fascinating question” of how cells appropriately scale the sizes of their intracellular components. She and her students will use their favorite model, *Xenopus laevis*, along with a related species, *Xenopus tropicalis*, to explore this subject.

X. tropicalis is about one-fifth the size of *X. laevis*, notes Heald. She explains that egg volume and mitotic spindle size are similarly scaled in the two species. Her laboratory has found that mixing extracts from the two species' eggs changes the mitotic spindle's size. “In the cytoplasm, there seems to be some readout of the size of the cell,” says Heald. “We'd like to find out how that works and what other structures in the cell are scaled like that.”

Investigating this question will keep Heald occupied for years to come, but she anticipates that other questions will continually arise to pique her interest. Pursuing new lines of research with her colleagues and collaborators is one of the best parts of her job, she says.

There are “still really tremendous unanswered questions, fundamental things about biology that we haven't been able to investigate very well in the past,” she says. “There's just so much to learn.”

—Christen Brownlee, Science Writer

REFERENCES

1. Hitchcock-DeGregori, S. E., and Heald, R. W. (1987) Altered actin and troponin binding of amino-terminal variants of chicken striated muscle alpha-tropomyosin expressed in *Escherichia coli*, *J. Biol. Chem.* 262, 9730–9735.
2. Heald, R. W., and Hitchcock-DeGregori, S. E. (1988) The structure of the amino terminus of tropomyosin is critical for binding to actin in the absence and presence of troponin, *J. Biol. Chem.* 263, 5254–5259.
3. Heald, R., and McKeon, F. (1990) Mutations of phosphorylation sites in lamin A that prevent nuclear lamina disassembly in mitosis, *Cell* 61, 579–589.
4. Heald, R., Tournebise, R., Blank, T., Sandaltzopoulos, R., Becker, P., Hyman, A., and Karsenti, E. (1996) Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts, *Nature* 382, 420–425.
5. Kalab, P., Weis, K., and Heald, R. (2002) Visualization of a Ran-GTP gradient in interphase and mitotic *Xenopus* egg extracts, *Science* 295, 2452–2456.
6. Kalab, P., Pralle, A., Isacoff, E. Y., Heald, R., and Weis, K. (2006) Analysis of a RanGTP-regulated gradient in mitotic somatic cells, *Nature* 440, 697–701.
7. Wignall, S. M., Gray, N. S., Chang, Y. T., Juarez, L., Jacob, R., Burlingame, A., Schultz, P. G., and Heald, R. (2004) Identification of a novel protein regulating microtubule stability through a chemical approach, *Chem. Biol.* 11, 135–146.

Induction of Pluripotency in Fibroblasts through the Expression of Only Four Nuclear Proteins

Yuequin Fang and Brendan P. Orner*

Division of Chemistry and Biological Chemistry, Nanyang Technological University, 1 Nanyang Walk, Block 5 Level 3, Singapore 637616

Cells that have the power to develop into any differentiated precursor or into a fully differentiated cell type are said to be “pluripotent”. It has been speculated that these cells could be incorporated as part of therapies to reverse cellular degeneration in diseases such as Alzheimer’s and Parkinson’s or used in tissue engineering and *ex vivo* organ growth applications. These cells could also provide illumination into human development by opening the door to *in vitro* experimentation. The most established method to derive human pluripotent cells is to isolate the inner cell mass from embryonic blastocysts to create human embryonic stem cell (hESC) lines (1). Many challenges must be overcome before these lines can be used as the basis of therapeutic technologies (2). Currently, most of the conditions to propagate the cells in a pluripotent state are complex, expensive, and ill-defined. Methods used to differentiate hESCs also tend to be low yielding and nonspecific and require multiple steps. Recently, combinatorial techniques have been explored with the intent to rapidly solve these problems (3). An additional challenge is perhaps less easily overcome. All of the early hESC lines were derived in a similar manner and were propagated on a feeder layer of mouse embryonic fibroblasts (MEFs). They all have been shown to express murine-specific antigens (4). Because of the current political climate,

all federally funded labs in the U.S. are restricted to only these cell lines. Although hESC lines grown in humanized conditions have been established (5), they are not available to federally funded researchers in the U.S. In order to work around these restrictions, scientists have sought to more fully understand pluripotency. Researchers have already determined that cells can be de-differentiated by transferring their nuclei to unfertilized eggs (6) and that pluripotency can be transferred by fusing a stem cell with a differentiated cell (7). These results suggest that pluripotent cell lines could be established by expressing certain unknown key factors in differentiated cells of a non-embryonic origin. An additional benefit would be that transplantation rejection could be prevented when pluripotent cells are derived from the patient to be treated.

A recent paper by Takahashi and Yamanaka (8) published in *Cell* suggests that only a small number of factors expressed in concert are necessary to induce pluripotency in differentiated cells. The article describes a screen in which the authors select, from a pool of 24 embryonic nuclear proteins, factors pivotal in the maintenance of ESC-like properties. The screen links G418-resistance in MEFs to induction of pluripotency and thereby establishes lines of induced pluripotent stem (iPS) cells. The results of the study demonstrate that iPS cells expressing all 24 factors have mor-

ABSTRACT Pluripotent cell lines have the potential to provide an unlimited supply of cells to therapeutically replace those damaged by various diseases. Understanding the nature of the pluripotency of these cells could result in more controlled methods for their propagation and differentiation and could help work around the politically based policy restrictions currently dogging the field. A recent paper describes how it is possible to generate pluripotent cell lines from differentiated adult murine fibroblast cells. Establishing a similar method for human cells would offer tissues for transplantation that elicit no rejection response and would provide an embryo-independent and patient-specific source of therapeutic cells that would quell ethical and political issues.

*Corresponding author,
orner@ntu.edu.sg.

Published online October 20, 2006

10.1021/cb600402w CCC: \$33.50

© 2006 by American Chemical Society

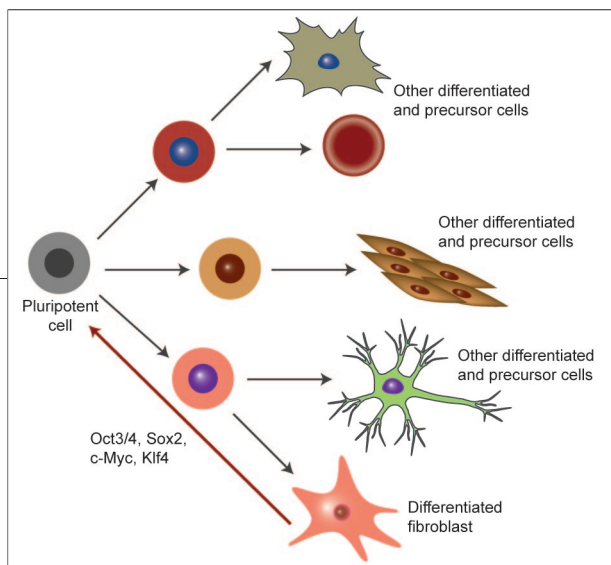


Figure 1. Differentiated mouse fibroblasts can be induced into a pluripotent state through the expression of only four nuclear factors.

phologies and growth rates and express many key markers similar to those of pluripotent stem cells. Interestingly, methylation patterns of key genes associated with pluripotency were similar but not identical to those of ESCs. By winnowing down the factors, the authors demonstrate that it is possible to maintain the pluripotency properties even if only 4 of the 24 are expressed (Figure 1). The established iPS cell lines could form embryoid bodies similar to those of ESCs. Additionally, the cells were injected under the skin of nude mice to induce teratoma tumor formation. These teratomas differentiated into all three developmental germ layers, demonstrating the pluripotency of the iPS cells. These four factors also could induce pluripotent properties in fully differentiated adult fibroblasts. Interestingly, these cells were injected into mouse blastocysts, but no chimeric mice were produced; this suggests that these cells cannot contribute to embryo development.

The four factors that were determined to induce pluripotency were Oct3/4, Sox2, c-Myc, and Klf4. It is not surprising that Oct3/4 and Sox1 are involved because they have been shown to be key for maintaining pluripotency (9). It is surprising, however, that the factor Nanog, which has also been shown to be important, is non-essential for the establishment of iPS cells in this study. The factors c-Myc and Klf4 are oncoproteins. The authors suggest that together their expression might balance the levels of p53 or p21^{CIP1} and thereby regulate apoptosis or cell proliferation; however, c-Myc in particu-

lar is rather nonspecifically global in its expression of downstream proteins.

The results described in this paper are intriguing. This study implies that the total deprogramming of differentiated cells

might be much easier than was previously thought. If true, then a stripped-down, rational method could be envisioned to supply cells for therapies in human patients. This method would circumvent government restrictions and make available tissues that would be safe from immune rejection upon transplantation. However, as promising as these results are, a great deal more research is required to determine whether similar approaches can be used for human therapies. Pluripotency in murine models seems to be regulated differently than in humans. For example, proliferating mouse ESCs is possible with the addition of a single growth factor. For hESCs, no magic factor has been discovered, and a complex cocktail of compounds must be added to maintain pluripotency and promote proliferation. Therefore, reprogramming human somatic cells would probably not be as simple as for their murine counterparts. In addition, the iPS cells, unlike ESCs, seem to be at an intermediate stage of pluripotency. The DNA methylation of key pluripotency genes resembles a half-way position between that of ESCs and that of the fibroblasts from which the iPS cells were derived (10). Furthermore, the failure of chimeric mouse development suggests that the cells may have the appearance of pluripotent cells in some ways but may not be “pluripotent enough”. Interestingly, this suggests that different degrees of pluripotency exist and that it may prove wise for the field to revisit its criteria for the determination of this important cellular property. To understand how these results can contribute to stem cell science and develop-

mental biology requires much more experimentation. It is unclear whether dedifferentiating human somatic cells can really be done so simply or even at all.

REFERENCES

1. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998) Embryonic stem cell lines derived from human blastocysts, *Science* 282, 1145–1147.
2. Oh, S. K. W., and Choo, A. B. H. (2006) Human embryonic stem cells: technological challenges towards therapy, *Clin. Exp. Pharmacol. Physiol.* 33, 489–495.
3. Orner, B. P., Derda, R., Lewis, R. L., Thomson, J. A., and Kiessling, L. L. (2004) Arrays for the combinatorial exploration of cell adhesion, *J. Am. Chem. Soc.* 126, 10808–10809.
4. Martin, M. J., Muotri, A., Gage, F., and Varki, A. (2005) Human embryonic stem cells express an immunogenic nonhuman sialic acid, *Nat. Med.* 11, 228–232.
5. Richards, M., Fong, C. Y., Chan, W. K., Wong, P. C., and Bongso, A. (2002) Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells, *Nat. Biotechnol.* 20, 933–936.
6. Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. S. (1997) Viable offspring derived from fetal and adult mammalian cells, *Nature* 385, 810–813.
7. Cowan, C. A., Atienza, J., Melton, D. A., and Eggan, K. (2005) Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells, *Science* 309, 1369–1373.
8. Takahashi, K., and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, *Cell* 126, 663–676.
9. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells, *Cell* 122, 947–956.
10. Rodolfa, K. T., and Eggan, K. (2006) A transcriptional logic for nuclear reprogramming, *Cell* 126, 652–655.

Small-Molecule Triggers of Tadpole Metamorphosis

Blake R. Peterson*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

The ability to achieve a high level of selectivity for a specific biological target is a key requirement for the generation of effective therapeutic agents and molecular probes. This endeavor can be quite challenging. Pockets that bind enzyme cofactors, hormones, and other key regulatory molecules are often highly conserved among protein family members that control diverse biological processes. However, our increasingly better understanding of the molecular recognition of protein–small-molecule interactions is beginning to allow the design of selective compounds that can target a single subtype among highly structurally similar members of protein families (1, 2). In this issue of *ACS Chemical Biology*, Ocasio and Scanlan (3) report on the use of a rational design approach to identify the first specific agonist of thyroid hormone receptor (TR)-alpha (TR α). This agonist has profound effects on TR α -mediated metamorphosis of *Xenopus laevis*, and it represents an important new molecular probe for dissecting signaling pathways controlled by this nuclear receptor (NR).

Members of the NR superfamily of transcription factors share conserved ligand binding domains (LBDs). By binding small molecules, NRs play key roles in embryonic development, cellular differentiation, metabolism, and cell death. Many of these proteins represent major drug targets, and selective therapeutic agents targeting specific NR subtypes have the potential to revolutionize the treatment of diseases of the endocrine system (4). One such NR is the TR,

which comprises two related subtypes, TR α and TR β , encoded by two different genes (5). The LBDs of these subtypes are ~75% identical in amino acid sequence. The thyroid hormone 3,5,3'-triiodo-L-thyronine (Figure 1, T₃, 1) binds the LBDs of both subtypes with high affinity ($K_d = 0.06$ nM) (6). Consequently, T₃ influences numerous physiological parameters, including growth, development, homeostasis, metabolism, heart rate, lipid levels, and mood (7). Pharmacological treatment with thyroid hormone has the potential to control body weight and lower cholesterol and triglycerides (8). However, this therapeutic approach has been plagued by side effects of hyperthyroidism, such as elevated heart rate and arrhythmia (9). Studies of patients resistant to thyroid hormone and knockout of the TR subtypes in mice have revealed that TR α mediates the effect of thyroid hormone on heart rate, whereas TR β affects other responses to this hormone (10). Because of the potential of TR β as a target of obesity, hyperlipidemia, depression, and osteoporosis, selective TR β agonists are of significant interest as therapeutic agents (11). However, the design of compounds selective for only one of the two TR subtypes is difficult. The residues that define the hydrophobic ligand binding pockets (LBPs) of TR α and TR β differ by only one amino acid; Ser277 in TR α is replaced by Asn331 in TR β . Despite these similarities, TR β -selective ligands such as GC-1 (2) (12) and compound 3 (Figure 1) (13) have been reported. A structural comparison of the LBDs of TR α bound to

ABSTRACT Small molecules that function as highly selective agonists and antagonists of cellular receptors comprise some of the most valuable therapeutic agents and molecular probes. A recent paper describes the design, synthesis, and evaluation of CO23, the first potent and specific agonist of thyroid hormone receptor-alpha (TR α), a member of the nuclear receptor (NR) superfamily of transcription factors. Together with previously reported TR β -selective agonists such as GC-1, these compounds represent powerful new tools for studying gene expression, signaling, differentiation, and development controlled by this important NR.

*Corresponding author,
brpeters@chem.psu.edu.

Published online October 20, 2006

10.1021/cb600398a CCC: \$33.50

© 2006 by American Chemical Society

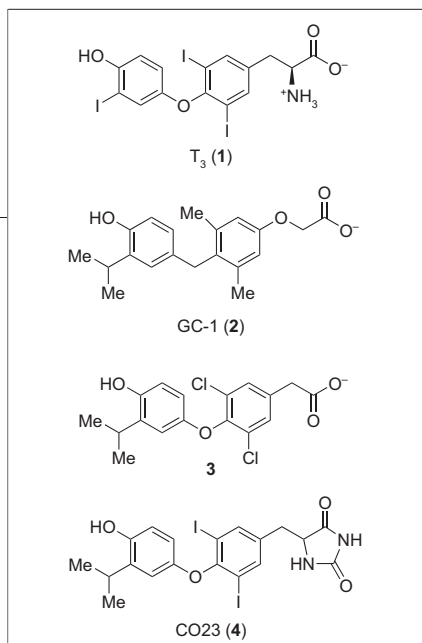


Figure 1. Structures of TR agonists. The thyroid hormone T₃ (1), which activates both TR α and TR β , is shown on the top. The TR β -selective thyromimetics GC-1 (2) and compound 3 are shown in the middle. The TR α -selective thyromimetic CO23 (4) is shown on the bottom.

compound **3** (13) and TR β bound to T₃ (1) (14) is shown (Figure 2).

Selective TR α agonists are of significant interest as molecular probes of TR signaling pathways. On page 585 of this issue of *ACS Chemical Biology*, Ocasio and Scanlan (3) describe the first such TR α -selective thyromimetic, termed CO23 (Figure 1, 4). CO23 (4) is a selective TR α agonist despite binding to both of the purified TR subtypes with equal affinity. This similarity in affinity is perhaps not surprising given that the LBPs are identical except for the single Ser277 to Asn331 substitution. In contrast, GC-1 (12) and compound **3** (13) are selective agonists of TR β because they bind substantially more tightly to this protein subtype. Several available X-ray structures of TRs bound to ligands (13–16) show that the selective binding of GC-1 (2) and compound **3** to TR β appears to result from participation of the carboxylate moiety in a hydrogen bonding network that includes the TR β Asn331 residue.

When evaluated against TRs expressed in mammalian cells, CO23 (4) selectively activates gene expression controlled by TR α . Previous studies of related estrogen receptors (17, 18) show that this selectivity may arise from subtle differences in the induced

conformations of amino acid side chains that define the LBP. Minor effects on amino acid side chains can influence the dynamics of the conformationally mobile helix-12, the most carboxy-terminal helix of NR LBDs (19). Perturbation of helix-12 can affect the recruitment of coactivator proteins to the LBD, potentiate interactions with the transcriptional machinery, and regulate gene expression.

Because of its selectivity for TR α , CO23 (4) has unique effects on TR signaling *in vivo*. When tadpoles of the frog *X. laevis* are treated with CO23 (4), this compound triggers only part of the developmental program that regulates metamorphosis; hind limbs of tadpoles grow, but other developmental changes are not activated. These changes include resorption of the head and tail observed upon treatment with the pan-agonist T₃ (1) or the TR β -selective agonist GC-1 (2) (20). Moreover, sequential treatment of tadpoles with CO23 (4) followed by GC-1 (2) revealed that these isoform-selective agonists can be used for temporal control of tadpole morphogenesis; the proper sequence of activation of the two TR subtypes is critical for correct execution of the developmental program. Furthermore, it was also confirmed by real-time PCR analysis of gene expression that selective activation of TR α *in vivo* by CO23 (4) is controlled by this TR subtype.

GC-1 (2) and CO23 (4) represent members of a growing family of small molecules that affect development or cellular differentiation by interacting with defined molecular targets. Among the compounds known to affect NRs in this way, BMS493, a pan-antagonist of the retinoic acid receptors alpha, beta, and gamma, affects hindbrain patterning in chick embryos (21). Rosi-

glitazone and other members of the thiazolidinedione family of antidiabetic drugs can control the differentiation of adipocytes by functioning as selective agonists of the peroxisome proliferator-activated receptor-gamma (PPAR γ) (22). Conversely, the selective PPAR γ antagonist T0070907 inhibits differentiation of adipocytes (23). Other compounds that affect development and differentiation through well-characterized mechanisms include pumorphamine (24, 25) and cyclopamine (26). These compounds function as agonists and antagonists, respectively, of Smoothened, a key regulator of the Hedgehog signaling pathway. Other examples of the use of chemical genetics to probe developmental biology have also been reviewed (27).

Activation of gene expression mediated by the genomic effects of TR α and TR β can now be dissociated *in vivo* through the use of specific small molecules. This chemical-genetics approach provides new tools to elucidate the complex biology regulated by NRs. However, in addition to their genomic effects, TRs and many other NRs activate distinct nongenomic signaling pathways (28). Given the previous achievements of using small molecules to separate genomic from nongenomic actions of estrogen receptors (29, 30), identifying compounds that isolate these specific functions of TR subtypes to provide even finer resolution

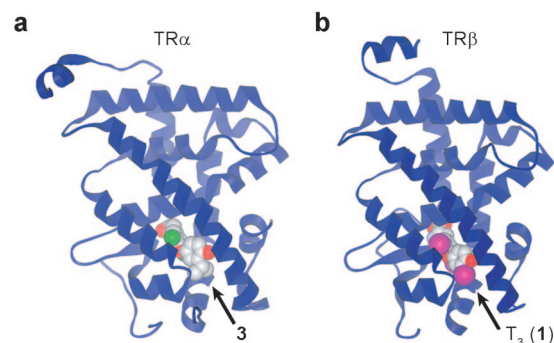


Figure 2. X-ray structures of TR LBDs. a) TR α bound to compound 3 (PDB code 1NAV). b) TR β bound to T₃ (1, PDB code 1XZX).

should be possible. Further development of this strategy has the potential to lead to highly selective therapeutic agents for various life-threatening diseases and to powerful molecular probes of signaling pathways yet to be fully elucidated.

REFERENCES

- Cohen, M. S., Zhang, C., Shokat, K. M., and Taunton, J. (2005) Structural bioinformatics-based design of selective, irreversible kinase inhibitors, *Science* 308, 1318–1321.
- Deininger, M., Buchdunger, E., and Druker, B. J. (2005) The development of imatinib as a therapeutic agent for chronic myeloid leukemia, *Blood* 105, 2640–2653.
- Ocasio, C. A., and Scanlan, T. S. (2006) Design and characterization of a thyroid hormone receptor-alpha (TRalpha)-specific agonist, *ACS Chem. Biol.* 1, 585–593
- Gronemeyer, H., Gustafsson, J. A., and Laudet, V. (2004) Principles for modulation of the nuclear receptor superfamily, *Nat. Rev. Drug Discovery* 3, 950–964.
- Thompson, C. C., Weinberger, C., Lebo, R., and Evans, R. M. (1987) Identification of a novel thyroid hormone receptor expressed in the mammalian central nervous system, *Science* 237, 1610–1614.
- Apriletti, J. W., Eberhardt, N. L., Latham, K. R., and Baxter, J. D. (1981) Affinity chromatography of thyroid hormone receptors. Biospecific elution from support matrices, characterization of the partially purified receptor, *J. Biol. Chem.* 256, 12094–12101.
- Yen, P. M. (2001) Physiological and molecular basis of thyroid hormone action, *Physiol. Rev.* 81, 1097–1142.
- Yoshihara, H. A., and Scanlan, T. S. (2003) Selective thyroid hormone receptor modulators, *Curr. Top. Med. Chem.* 3, 1601–1616.
- Osman, F., Gammage, M. D., and Franklyn, J. A. (2002) Hyperthyroidism and cardiovascular morbidity and mortality, *Thyroid* 12, 483–487.
- Grover, G. J., Mellstrom, K., Ye, L., Malm, J., Li, Y. L., Bladh, L. G., Sleph, P. G., Smith, M. A., George, R., Vennstrom, B., Mookhtiar, K., Horvath, R., Speelman, J., Egan, D., and Baxter, J. D. (2003) Selective thyroid hormone receptor-beta activation: a strategy for reduction of weight, cholesterol, and lipoprotein (a) with reduced cardiovascular liability, *Proc. Natl. Acad. Sci. U.S.A.* 100, 10067–10072.
- Baxter, J. D., Webb, P., Grover, G., and Scanlan, T. S. (2004) Selective activation of thyroid hormone signaling pathways by GC-1: a new approach to controlling cholesterol and body weight, *Trends. Endocrinol. Metab.* 15, 154–157.
- Chiellini, G., Apriletti, J. W., Yoshihara, H. A., Baxter, J. D., Ribeiro, R. C., and Scanlan, T. S. (1998) A high-affinity subtype-selective agonist ligand for the thyroid hormone receptor, *Chem. Biol.* 5, 299–306.
- Ye, L., Li, Y. L., Mellstrom, K., Mellin, C., Bladh, L. G., Koehler, K., Garg, N., Garcia Collazo, A. M., Litten, C., Husman, B., Persson, K., Ljunggren, J., Grover, G., Sleph, P. G., George, R., and Malm, J. (2003) Thyroid receptor ligands. 1. Agonist ligands selective for the thyroid receptor beta1, *J. Med. Chem.* 46, 1580–1588.
- Sandler, B., Webb, P., Apriletti, J. W., Huber, B. R., Togashi, M., Cunha Lima, S. T., Juric, S., Nilsson, S., Wagner, R., Fletterick, R. J., and Baxter, J. D. (2004) Thyroxine-thyroid hormone receptor interactions, *J. Biol. Chem.* 279, 55801–55808.
- Bomgraaber, S., Budny, M. J., Chiellini, G., Cunha-Lima, S. T., Togashi, M., Webb, P., Baxter, J. D., Scanlan, T. S., and Fletterick, R. J. (2003) Ligand selectivity by seeking hydrophobicity in thyroid hormone receptor, *Proc. Natl. Acad. Sci. U.S.A.* 100, 15358–15363.
- Wagner, R. L., Huber, B. R., Shiau, A. K., Kelly, A., Cunha Lima, S. T., Scanlan, T. S., Apriletti, J. W., Baxter, J. D., West, B. L., and Fletterick, R. J. (2001) Hormone selectivity in thyroid hormone receptors, *Mol. Endocrinol.* 15, 398–410.
- Shiau, A. K., Barstad, D., Radek, J. T., Meyers, M. J., Nettles, K. W., Katzenellenbogen, B. S., Katzenellenbogen, J. A., Agard, D. A., and Greene, G. L. (2002) Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism, *Nat. Struct. Biol.* 9, 359–364.
- Shang, Y., and Brown, M. (2002) Molecular determinants for the tissue specificity of SERMs, *Science* 295, 2465–2468.
- Steinmetz, A. C., Renaud, J. P., and Moras, D. (2001) Binding of ligands and activation of transcription by nuclear receptors, *Annu. Rev. Biophys. Biomol. Struct.* 30, 329–359.
- Furlow, J. D., Yang, H. Y., Hsu, M., Lim, W., Ermio, D. J., Chiellini, G., and Scanlan, T. S. (2004) Induction of larval tissue resorption in *Xenopus laevis* tadpoles by the thyroid hormone receptor agonist GC-1, *J. Biol. Chem.* 279, 26555–26562.
- Dupe, V., and Lumsden, A. (2001) Hindbrain patterning involves graded responses to retinoic acid signalling, *Development* 128, 2199–2208.
- Lehrke, M., and Lazar, M. A. (2005) The many faces of PPARgamma, *Cell* 123, 993–999.
- Lee, G., Elwood, F., McNally, J., Weiszmann, J., Lindstrom, M., Amaral, K., Nakamura, M., Miao, S., Cao, P., Learned, R. M., Chen, J. L., and Li, Y. (2002) T0070907, a selective ligand for peroxisome proliferator-activated receptor gamma, functions as an antagonist of biochemical and cellular activities, *J. Biol. Chem.* 277, 19649–19657.
- Sinha, S., and Chen, J. K. (2006) Purmorphamine activates the Hedgehog pathway by targeting Smoothened, *Nat. Chem. Biol.* 2, 29–30.
- Wu, X., Ding, S., Ding, Q., Gray, N. S., and Schultz, P. G. (2002) A small molecule with osteogenesis-inducing activity in multipotent mesenchymal progenitor cells, *J. Am. Chem. Soc.* 124, 14520–14521.
- Chen, J. K., Taipale, J., Young, K. E., Maiti, T., and Beachy, P. A. (2002) Small molecule modulation of Smoothened activity, *Proc. Natl. Acad. Sci. U.S.A.* 99, 14071–14076.
- Yeh, J. R., and Crews, C. M. (2003) Chemical genetics: adding to the developmental biology toolbox, *Dev. Cell* 5, 11–19.
- Hiroi, H., Kim, H. H., Ying, Y., Furuya, F., Huang, Z., Simoncini, T., Noma, K., Ueki, K., Nguyen, N. H., Scanlan, T. S., Moskowitz, M. A., Cheng, S. Y., and Liao, J. K. (2006) Rapid nongenomic actions of thyroid hormone, *Proc. Natl. Acad. Sci. U.S.A.* 103, 14104–14109.
- Harrington, W. R., Kim, S. H., Funk, C. C., Madak-Erdogan, Z., Schiff, R., Katzenellenbogen, J. A., Katzenellenbogen, B. S. (2006) Estrogen dendrimer conjugates that preferentially activate extranuclear, nongenomic versus genomic pathways of estrogen action, *Mol. Endocrinol.* 20, 491–502.
- Tobias, S. C., Qiu, J., Kelly, M. J., and Scanlan, T. S. (2006) Synthesis and biological evaluation of SERMs with potent nongenomic estrogenic activity, *ChemMedChem* 1, 565–571.

Errors in Translation Cause Selective Neurodegeneration

Jean-Christophe Rochet*

Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, Heine Pharmacy Building, 575 Stadium Mall Drive, West Lafayette, Indiana 47907-2091

ABSTRACT The 3D structure of a protein is determined by the unique sequence of amino acid residues comprising the polypeptide chain. Sequence changes can cause protein misfolding, a potentially toxic phenomenon implicated in various neurodegenerative disorders. In a recent paper, translational misincorporation is proposed to be a new biochemical mechanism for generating mutant proteins that misfold and kill neurons.

A typical globular protein adopts a compact structure with buried hydrophobic residues (1). This protein fold is stabilized by a network of intermolecular contacts (including hydrophobic, electrostatic, and hydrogen-bonding interactions) involving residues throughout the polypeptide chain. Modifications to the amino acid sequence can disrupt these stabilizing interactions and result in protein misfolding. In turn, misfolded polypeptides have a high propensity to form aggregates *via* interactions among exposed hydrophobic domains. Several proteins have been found to misfold and aggregate in neurodegenerative disorders (NDDs) (2). In general, these proteins undergo misfolding because of sequence changes originating from gene mutations or post-translational modifications (PTMs). For example, mutant forms of $\text{Cu}^{2+}/\text{Zn}^{2+}$ superoxide dismutase produce aggregates in familial amyotrophic lateral sclerosis (ALS) (3), and oxidative modifications stabilize aggregated forms of α -synuclein in Parkinson's disease (PD) (4, 5). Now in a recent paper published in *Nature*, Susan Ackerman, Paul Schimmel, and colleagues (6) report that translational misincorporation, a process in which the wrong amino acid is incorporated into the growing polypeptide chain, is another mechanism by which misfolded, neurotoxic proteins are generated.

Normally, misfolded proteins in the cytosol are refolded by chaperones (including heat shock proteins (Hsps) upregulated

in response to protein unfolding) or degraded by the ubiquitin proteasome pathway (UPP) (7). However, in cells with a high degree of protein unfolding, the capacity of these "quality-control" systems is exceeded, and misfolded proteins form aggregates. These aggregates are then recruited to perinuclear inclusion bodies named "aggresomes", where they are targeted for destruction *via* autophagy (8–10). Autophagy is a cellular process involving sequestration of cellular material into a vesicular structure termed an "autophagosome" (9, 10). The autophagosome ultimately delivers its contents to the lysosome for degradation. In addition to triggering aggresome formation, a buildup of cytosolic protein aggregates elicits protein misfolding in the endoplasmic reticulum (ER); the result is ER stress and induction of the unfolded protein response (UPR) (11). The UPR is a cellular program involving upregulation of ER chaperones, increased degradation of misfolded polypeptides, translational suppression, and under severe conditions of extreme ER stress, apoptotic signaling (11, 12).

From the above discussion, it is clear that amino acid substitutions resulting in protein misfolding can have profoundly disruptive effects on cellular homeostasis. One mechanism by which substitutions are avoided is *via* faithful transmission of the genetic code from DNA to protein. Aminoacyl-transfer RNA (tRNA) synthetases (aaRSs) play a critical role in this process by ensuring that each

*Corresponding author,
rochet@pharmacy.purdue.edu.

Published online October 20, 2006

10.1012/cb6004068 CCC: \$33.50

© 2006 by American Chemical Society

tRNA is only charged with its cognate amino acid (13). Each of the 20 aaRS catalyzes a two-part reaction: activation of the amino acid with ATP, yielding enzyme-bound aminoacyl-AMP with release of PP_i; and ligation of the amino acid to the 3' end of the tRNA, yielding aminoacyl-tRNA (aa-tRNA) with release of AMP. The specificity of the reaction for the cognate amino acid is largely determined by the exclusion of amino acids that are too bulky or lack key functional groups to interact properly with the active site. However, because smaller amino acids can escape these constraints, aaRSs have evolved a second "editing" domain that catalyzes hydrolysis of incorrect aa-tRNAs (the correct aa-tRNA is sterically excluded from the editing site).

Ackerman, Schimmel, and colleagues (6) discovered a link between errors in translation and protein misfolding in their studies of mice with the "sticky" (*sti*) mutation. These mice (so named because of the unkempt appearance of their fur) were found to have motor deficits, including trembling and ataxia. The mutant mice also exhibited age-dependent, apoptotic neuronal death in the Purkinje cell layer of the cerebellum. The authors mapped the *sti* mutation to the gene encoding alanyl-tRNA synthetase (AlaRS) and found that it resulted in the substitution of Ala734 with a glutamate residue (A734E). Expression of a transgene encoding wild-type AlaRS suppressed Purkinje cell loss and ataxia in the *sti* mutant mice, confirming that the A734E substitution is responsible for the neurodegenerative phenotype.

Ala734 is a conserved residue in the putative editing domain of AlaRS. Mutations in this domain of the *Escherichia coli* enzyme lead to increased levels of misacylated Ser- or Gly-tRNA^{Ala}, and cells expressing mutant AlaRS have decreased viability in media supplemented with high amounts of serine or glycine (14). Presumably, these elevated noncognate amino acids act as an environmental "stressor" that elicits toxicity

via increased tRNA^{Ala} misacylation and, therefore, decreased translational fidelity. To address whether misacylation is involved in the *sti* phenotype, Ackerman, Schimmel, and colleagues (6) compared wild-type and mutant mouse embryonic fibroblasts (MEFs) in terms of their sensitivity to elevated amino acids in the cell-culture medium. Homozygous (*sti/sti*) and heterozygous (*sti/+*) mutant MEFs were markedly more sensitive to elevated serine than were wild-type fibroblasts, and this suggests that inefficient editing of Ser-tRNA^{Ala} by mutant AlaRS may contribute to neurodegeneration in the mutant mice.

To characterize the editing function of AlaRS-A734E more directly, Ackerman, Schimmel, and colleagues (6) conducted acylation and deacylation assays with the recombinant enzyme. Wild-type and mutant AlaRS exhibited similar kinetics of tRNA^{Ala} acylation with alanine, and neither enzyme catalyzed Ala-tRNA^{Ala} deacylation above background levels. In contrast, A734E catalyzed deacylation of Ser-tRNA^{Ala} less efficiently than wild-type AlaRS (although the rate of deacylation by the mutant enzyme was still well above background). The A734E mutant also catalyzed more rapid misacylation of tRNA^{Ala} with serine compared with the wild-type enzyme. These biochemical data provided strong evidence that the *sti* mutation causes a specific defect in AlaRS editing of Ser-tRNA^{Ala}.

The authors predicted that the editing defect of mutant AlaRS would lead to errors in translation and, therefore, a buildup of misfolded proteins with amino acid substitutions (Figure 1, panel a). Many of these misfolded polypeptides should be ubiquitinated, reflecting abortive attempts by the UPP to eliminate these potentially toxic species (see above). In support of this idea, the authors found that ubiquitylated proteins were more abundant in *sti/sti* MEFs and in mutant Purkinje cells than in the corresponding wild-type cells. Strikingly, perinuclear inclusions and autophagosome-

like structures were detected in cerebellar neurons of *sti/sti* mice. The mutant Purkinje neurons also contained increased levels of cytosolic chaperones, including Hsp72, Hsc70, and Hsp40, and components of the UPR, including the ER chaperone immunoglobulin heavy-chain binding protein (BiP) and the pro-apoptotic transcription factor C/EBP homologous protein (CHOP). These data are consistent with a model in which mutant AlaRS causes increased protein misfolding and aggregation in *sti* mutant mice and results in cerebellar neurodegeneration (Figure 1, panel b).

Defects in tRNA acylation have been shown to play a role in some human NDDs. Two groups reported that overaccumulation of p38, a cofactor involved in the assembly of the multi-aaRS complex, contributes to the death of dopaminergic neurons in PD (15, 16). In addition, mutations in the gene encoding glycyl- or tyrosyl-tRNA synthetase were linked to motor neuropathy in Charcot-Marie-Tooth disease (17, 18). The study by Ackerman, Schimmel, and colleagues is the first to show that impaired aaRS proofreading causes neurodegeneration. The finding that the *sti* mutation targets the editing domain of AlaRS is highly significant because it suggests that amino acid substitutions will be introduced throughout the proteome. Accordingly, this mutation is predicted to cause more widespread misfolding and aggregation than do monogenic lesions that elicit destabilization of a single protein. Protein aggregation in *sti* mutant mice may trigger cell death via a gain of function involving toxic effects of oligomeric species (e.g., membrane permeabilization) (19, 20) or via a loss of function involving the inactivation of proteins that are recruited into the aggregates.

Although the data argue convincingly that Purkinje cell death in *sti* mutant mice involves a buildup of toxic protein aggregates, the cerebellar phenotype may also result from the disruption of protein activities required for neuronal survival, indepen-

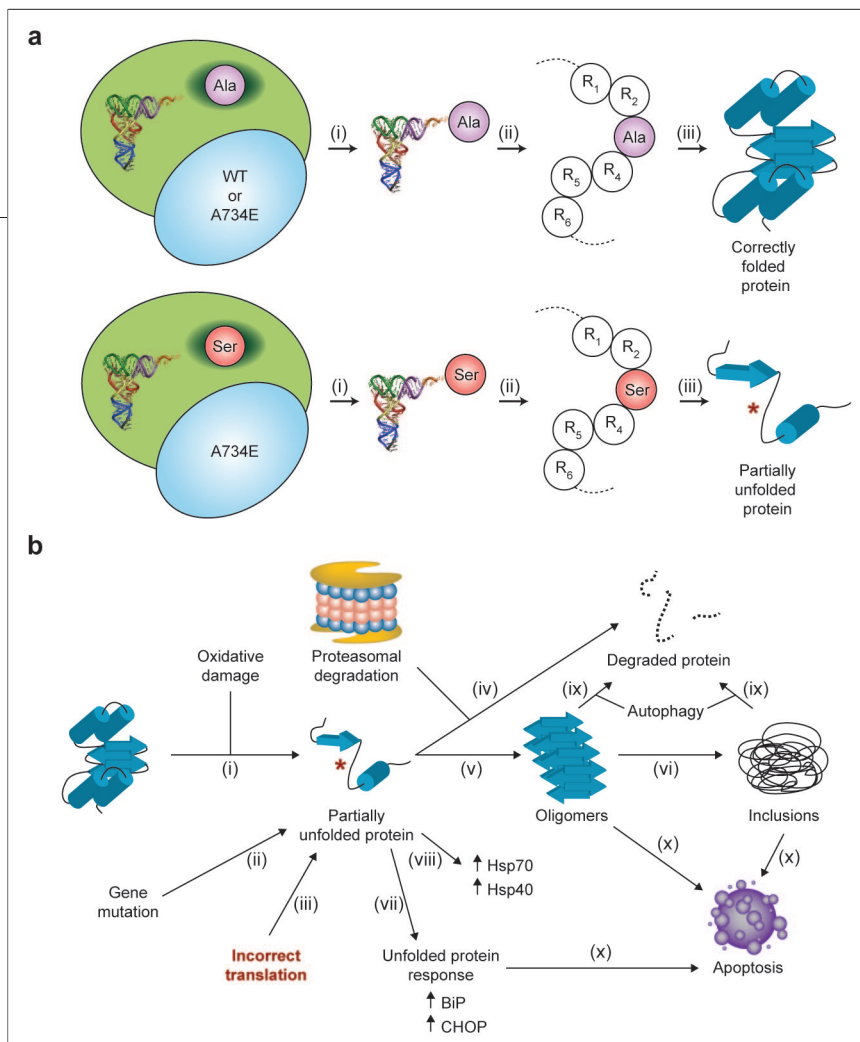


Figure 1. Defects in the editing function of aaRSs cause protein misfolding and aggregation. **a)** Wild-type AlaRS and the A734E mutant catalyze the acylation of tRNA^{Ala} with the cognate amino acid, alanine, resulting in error-free translation and correct protein folding (top). A734E also catalyzes the acylation of tRNA^{Ala} with the noncognate amino acid, serine, resulting in translational misincorporation and protein misfolding (bottom). The large green oval represents the AlaRS aminoacylation domain, and the smaller blue oval corresponds to the editing domain. **b)** Protein misfolding is induced by (i) PTMs, for example, oxidative damage, (ii) gene mutations, or (iii) errors in translation due to defective AlaRS editing. (iv) Under normal conditions, misfolded proteins are eliminated by the UPP. However, at sufficiently high concentrations, misfolded polypeptides form aggregates, shown here as (v) early oligomers and (vi) mature inclusions. Protein misfolding induces (vii) the UPR, including upregulation of BiP and CHOP, and (viii) increased expression of cytosolic chaperones, including Hsp70 and Hsp40. (ix) The cell attempts to eliminate aggregated proteins *via* autophagy. (x) Failure to eliminate toxic protein aggregates ultimately leads to apoptosis, in part regulated by UPR signaling. tRNA image courtesy of Neil Voss, Wikipedia.

dent of aggregation. It is also possible that the *sti* mutation triggers cell death by eliminating noncanonical functions of AlaRS distinct from aa-tRNA synthesis or editing (21). Additional evidence in support of a role for protein aggregation rather than these alternative mechanisms might be obtained by testing whether the overexpression of chaperones such as Hsp70 suppresses inclusion formation and neurodegeneration in *sti* mutant mice.

The extent to which the *sti* mutation disrupts translational fidelity *in vivo* has yet to be determined. Because recombinant A734E has significant (albeit reduced) Ser-tRNA^{Ala} deacylation activity, only a fraction of alanine residues in the proteome are likely to be replaced with serine. The data obtained by Ackerman, Schimmel, and colleagues suggest that this fraction may vary from tissue to tissue because of differences in serine content. Even a low overall percent-

age of alanine-to-serine substitutions would be expected to have a pronounced impact on cellular function, given that all of the proteins in the cell would be targeted. Moreover, small amounts of substituted, misfolded polypeptides may increase their toxic effect by “seeding” the aggregation of more abundant wild-type isoforms (22, 23).

Because translational proofreading is an essential part of protein synthesis in all cell types, it is remarkable that impaired AlaRS editing results in a specific neurodegenerative phenotype in *sti* mutant mice. The observation that protein misfolding specifically targets the central nervous system in these mice is consistent with the involvement of protein aggregation in other NDDs, including Alzheimer’s disease, PD, ALS, Huntington’s disease, and the spinocerebellar ataxias (2, 24). One reason for the sensitivity of terminally differentiated neurons to protein misfolding is that these cells are unable to dilute out toxic protein aggregates *via* mitosis (6). In addition, postmitotic neurons are highly dependent on the UPP to avoid cell-cycle re-entry, a phenomenon that can trigger neuronal apoptosis (25). Accordingly, these cells may be more vulnerable than other cell types to mechanisms that cause impairment of the UPP, including protein misfolding and aggregation.

Although one can rationalize why terminally differentiated neurons are sensitive to protein misfolding, it is unclear why Purkinje cells are especially vulnerable in *sti* mutant mice. Purkinje neurons are selectively targeted in other protein aggregation disorders, including spinocerebellar ataxias (26); this suggests that these cells have inherently inefficient quality-control systems (6). Nevertheless, brain regions other than the Purkinje cell layer (*e.g.*, the *substantia nigra*) are sensitive to conditions that induce protein misfolding (27–29), and these would also be expected to undergo neurodegeneration as a consequence of impaired translational editing in *sti* mutant mice. A rationale for the selective degenerative phe-

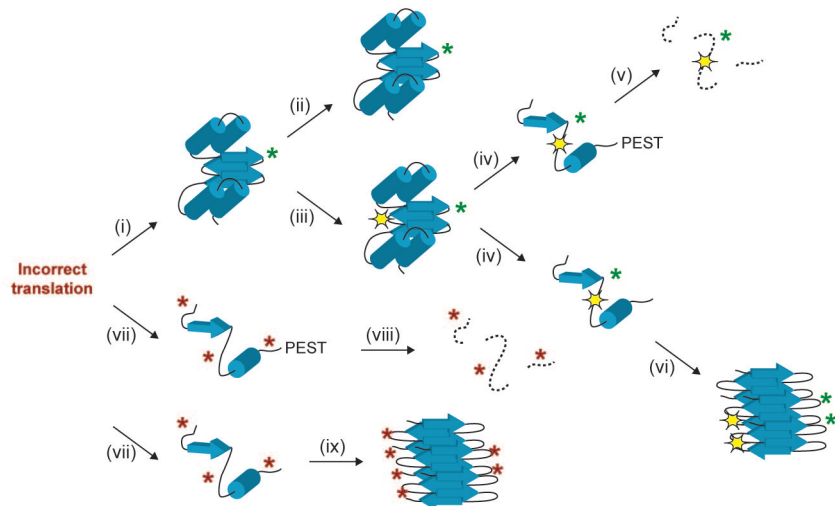


Figure 2. Model illustrating various possible outcomes of the AlaRS proofreading defect. (i) Alanine-to-serine substitutions may be introduced in proteins with relatively few buried alanine residues at sites that cannot accommodate a serine residue. (ii) This pattern of replacement (represented by a green asterisk) causes little disruption of the protein fold. (iii) Some proteins with these substitutions may undergo additional “hits” in the form of PTMs, represented by a yellow “burst”. (iv) The combined effects of the alanine-to-serine replacements and PTMs may weaken the protein structure and thus trigger misfolding. (v) Some misfolded polypeptides may be inherently unstable (e.g., because of the presence of a PEST sequence) and therefore will be rapidly degraded. (vi) Longer-lived misfolded proteins may accumulate and form aggregates, depicted as β -sheet-rich oligomers. (vii) Alanine-to-serine substitutions may also be introduced in proteins with a relatively large number of buried alanine residues. This pattern of substitution (represented by red asterisks) leads to protein misfolding. (viii) Some of the misfolded polypeptides may be rapidly degraded, whereas others (ix) with longer half-lives may accumulate to form aggregates.

notype of these mice is implied by the predicted biochemical consequences of the AlaRS editing defect. Although impaired proofreading should trigger translational misincorporation throughout the proteome, only a fraction of substituted polypeptides are expected to undergo misfolding and aggregation (Figure 2). Some proteins may be resistant to the potentially destabilizing effects of alanine-to-serine substitutions because they have a relatively low proportion of buried alanine residues at sites that cannot accommodate a serine residue (30). Other proteins that unfold in response to alanine-to-serine replacements may be targeted for rapid degradation before they can accumulate (e.g., proteins with a proline,

glutamic acid, serine, and threonine (PEST) sequence (31)). Conversely, proteins with long half-lives and large numbers of buried alanine residues are more likely to form aggregates in response to impaired AlaRS editing. An abundance of these proteins in the Purkinje cell layer could explain why this region of the brain is specifically targeted in *sti* mutant mice. Moreover, some proteins with alanine-to-serine substitutions may only misfold and aggregate after undergoing PTMs that further destabilize the protein fold (Figure 2). If these PTMs occur frequently in Purkinje neurons, then this “multi-hit” phenomenon could also account for the selective nature of the *sti* phenotype.

In summary, the findings reported by Ackerman, Schimmel, and colleagues suggest a novel mechanism for protein misfolding and aggregation in postmitotic neurons and provide intriguing insight into potential biochemical changes underlying selective neurodegeneration. To better understand the molecular basis for the *sti* phenotype, it will be important to determine which proteins misfold and aggregate in response to defective translational editing. This problem can be addressed *via* comparative proteomics, to identify proteins aggregated in mutant but not wild-type cerebellar tissue, and *via* computational modeling, to identify proteins with a high likelihood of misfolding in response to alanine-to-serine substitutions. It would also be of interest to determine whether defects in the editing function of aaRSs other than AlaRS cause neurodegeneration and whether the patterns of cell loss in these cases differ from the selective pathology in *sti* mutant mice. Finally, a high priority will be to determine whether mutations in aaRS genes are involved in human NDDs. Presumably, such mutations could only cause a moderate impairment of translational proofreading, with neurodegeneration occurring after the onset of reproductive age; otherwise, they would not be retained in the human population. Evidence of a role for defective editing in human NDDs could have important implications for drug discovery: it would suggest that improving translational fidelity may be a useful therapeutic strategy in the treatment of these disorders.

REFERENCES

1. Dobson, C. M. (2003) Protein folding and misfolding, *Nature* 426, 884–890.
2. Taylor, J. P., Hardy, J., and Fischbeck, K. H. (2002) Toxic proteins in neurodegenerative disease, *Science* 296, 1991–1995.
3. Valentine, J. S., and Hart, P. J. (2003) Misfolded CuZnSOD and amyotrophic lateral sclerosis, *Proc. Natl. Acad. Sci. U.S.A.* 100, 3617–3622.
4. Conway, K. A., Rochet, J.-C., Bieganski, R. M., and Lansbury, P. T., Jr., (2001) Kinetic stabilization of the α -synuclein protofibril by a dopamine- α -synuclein adduct, *Science* 294, 1346–1349.

5. Cole, N. B., Murphy, D. D., Lebowitz, J., Di Noto, L., Levine, R. L., and Nussbaum, R. L. (2005) Metal-catalyzed oxidation of alpha synuclein: helping to define the relationship between oligomers, protofilaments and filaments, *J. Biol. Chem.* **280**, 9678–9690.
6. Lee, J. W., Beebe, K., Nangle, L. A., Jang, J., Longo-Guess, C. M., Cook, S. A., Davisson, M. T., Sundberg, J. P., Schimmel, P., and Ackerman, S. L. (2006) Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration, *Nature* **443**, 50–55.
7. Berke, S. J., and Paulson, H. L. (2003) Protein aggregation and the ubiquitin proteasome pathway: gaining the UPper hand on neurodegeneration, *Curr. Opin. Genet. Dev.* **13**, 253–261.
8. Johnston, J. A., Ward, C. L., and Kopito, R. R. (1998) Aggresomes: a cellular response to misfolded proteins, *J. Cell Biol.* **143**, 1883–1898.
9. Iwata, A., Riley, B. E., Johnston, J. A., and Kopito, R. R. (2005) HDAC6 and microtubules are required for autophagic degradation of aggregated huntingtin, *J. Biol. Chem.* **280**, 40282–40292.
10. Ravikumar, B., Acevedo-Arozena, A., Imarisio, S., Berger, Z., Vacher, C., O’Kane, C. J., Brown, S. D., and Rubinsztein, D. C. (2005) Dynein mutations impair autophagic clearance of aggregate-prone proteins, *Nat. Genet.* **37**, 771–776.
11. Rutkowski, D. T., and Kaufman, R. J. (2004) A trip to the ER: coping with stress, *Trends Cell. Biol.* **14**, 20–28.
12. Oyadomari, S., and Mori, M. (2004) Roles of CHOP/GADD153 in endoplasmic reticulum stress, *Cell Death Differ.* **11**, 381–389.
13. Gestain, R., and Ribas de Pouplana, L. (2004) Regulation of RNA function by aminoacylation and editing? *Trends Genet.* **20**, 604–610.
14. Beebe, K., Ribas De Pouplana, L., and Schimmel, P. (2003) Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability, *EMBO J.* **22**, 668–675.
15. Corti, O., Hampe, C., Koutnikova, H., Darios, F., Jacquier, S., Prigent, A., Robinson, J. C., Pradier, L., Ruberg, M., Mirande, M., Hirsch, E., Rooney, T., Fournier, A., and Brice, A. (2003) The p38 subunit of the aminoacyl-tRNA synthetase complex is a Parkin substrate: linking protein biosynthesis and neurodegeneration, *Hum. Mol. Genet.* **12**, 1427–1437.
16. Ko, H. S., von Coelln, R., Sriram, S. R., Kim, S. W., Chung, K. K., Pletnikova, O., Troncoso, J., Johnson, B., Saffary, R., Goh, E. L., Song, H., Park, B. J., Kim, M. J., Kim, S., Dawson, V. L., Dawson, T. M. (2005) Accumulation of the authentic parkin substrate aminoacyl-tRNA synthetase cofactor, p38/ITV-1, leads to catecholaminergic cell death, *J. Neurosci.* **25**, 7968–7978.
17. Antonellis, A., Ellsworth, R. E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S. Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L. T., Sivakumar, K., Ionasescu, V., Funalot, B., Vance, J. M., Goldfarb, L. G., Fischbeck, K. H., and Green, E. D. (2003) Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V, *Am. J. Hum. Genet.* **72**, 1293–1299.
18. Jordanova, A., Irobi, J., Thomas, F. P., Van Dijk, P., Meerschaert, K., Dewil, M., Dierick, I., Jacobs, A., De Vriendt, E., Guergueltcheva, V., Rao, C. V., Tournev, I., Gondim, F. A., D’Hooghe, M., Van Gerwen, V., Callaerts, P., Van Den Bosch, L., Timmermans, J. P., Robberecht, W., Gettemans, J., Thevelein, J. M., De Jonghe, P., Kremensky, I., and Timmerman, V. (2006) Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy, *Nat. Genet.* **38**, 197–202.
19. Glabe, C. G., and Kaye, R. (2006) Common structure and toxic function of amyloid oligomers implies a common mechanism of pathogenesis, *Neurology* **66**, S74–S78.
20. Lashuel, H. A., and Lansbury, P. T. (2006) Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q. Rev. Biophys.* 1–35, published online Sept. 18, 2006. DOI: 10.1017/S0033583506004422.
21. Park, S. G., Ewalt, K. L., and Kim, S. (2005) Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers, *Trends Biochem. Sci.* **30**, 569–574.
22. Krebs, M. R., Morozova-Roche, L. A., Daniel, K., Robinson, C. V., and Dobson, C. M. (2004) Observation of sequence specificity in the seeding of protein amyloid fibrils, *Protein. Sci.* **13**, 1933–1938.
23. Jones, E. M., and Surewicz, W. K. (2005) Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids, *Cell* **121**, 63–72.
24. Gatchel, J. R., and Zoghbi, H. Y. (2005) Diseases of unstable repeat expansion: mechanisms and common principles, *Nat. Rev. Genet.* **6**, 743–755.
25. Staropoli, J. F., and Abeliovich, A. (2005) The ubiquitin-proteasome pathway is necessary for maintenance of the postmitotic status of neurons, *J. Mol. Neurosci.* **27**, 175–183.
26. Koeppen, A. H. (2005) The pathogenesis of spinocerebellar ataxia, *Cerebellum* **4**, 62–73.
27. Betarbet, R., Sherer, T. B., MacKenzie, G., Garcia-Osuna, M., Panov, A. V., and Greenamyre, J. T. (2000) Chronic systemic pesticide exposure reproduces features of Parkinson’s disease, *Nat. Neurosci.* **3**, 1301–1306.
28. McNaught, K. S., Perl, D. P., Brownell, A. L., and Olanow, C. W. (2004) Systemic exposure to proteasome inhibitors causes a progressive model of Parkinson’s disease, *Ann. Neurol.* **56**, 149–162.
29. Fornai, F., Schluter, O. M., Lenzi, P., Gesi, M., Ruffoli, R., Ferrucci, M., Lazzeri, G., Busceti, C. L., Pontarelli, F., Battaglia, G., Pellegrini, A., Nicoletti, F., Ruggieri, S., Paparelli, A., and Sudhof, T. C. (2005) Parkinson-like syndrome induced by continuous MPTP infusion: convergent roles of the ubiquitin-proteasome system and α -synuclein, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3413–3418.
30. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W., and Matthews, B. W. (1993) Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala→Ser and Val→Thr substitutions in T4 lysozyme, *Biochemistry* **32**, 11363–11373.
31. Rechsteiner, M., and Rogers, S. W. (1996) PEST sequences and regulation by proteolysis, *Trends Biochem. Sci.* **21**, 267–271.

Deepening Ribosomal Insights

Anders Liljas*

Department of Molecular Biophysics, Center for Chemistry and Chemical Engineering, Lund University, Box 124, SE-221 00, Lund, Sweden

Translation of messenger RNA (mRNA) to protein occurs on the ribosome.

The mRNA binds around the neck of the small (30S) subunit, between the head and the body, whereas the peptidyl transfer occurs at the peptidyl transfer center (PTC) on the large (50S) subunit (Figure 1). The ribosome has three major sites for transfer RNA (tRNA) molecules, the A-, P-, and E-sites. The A-site is where the codon–anticodon interactions are scrutinized to ensure that the correct amino acid will be incorporated. The acceptor end of the tRNA is composed of three conserved single-stranded nucleotides, C74, C75, and A76. The aminoacyl moiety attached to A76 of the tRNA in the A-site accepts the growing polypeptide from the tRNA in the P-site during peptidyl transfer. The new peptidyl-tRNA is subsequently translocated from the A-site to the P-site. The deacylated tRNA is translocated from the P-site to the exit (E)-site, where it finally dissociates. Translation is catalyzed by several GTPases, primarily elongation factors Tu (EF-Tu) and G (EF-G). Whereas EF-Tu with GTP delivers the charged tRNA to the ribosome and dissociates after GTP hydrolysis, EF-G with GTP catalyzes the translocation of mRNA and tRNAs (see below).

Two new structures of 70S ribosomes (1, 2) add significantly to the information gained by the high-resolution structures of separate 30S subunits from *Thermus thermophilus* and 50S subunits (3, 4) from *Haloarcula marismortui* (5) and *Deinococcus radiodurans* (6), the 5.5-Å structure of 70S *T. thermophilus* ribosomes (7), and the two

complete *Escherichia coli* ribosomes analyzed at 3.5-Å resolution (8). For a full understanding of ribosomal function, we need to study complete ribosomes and as many of their complexes as possible at high resolution. Even though the ribosomes in the most recent studies come from the same species, the varying crystallization conditions have led to different crystal packing and diffraction power (1, 2). With improved resolution, new details can be identified and old observations can be confirmed or contradicted and are of great general interest.

With regard to the general structure, certain features remain too flexible to be seen. In both structures, proteins L10, L7/L12, and L11 at the GTPase center are not seen, and a few additional poorly visible proteins differ between the two structures. In the new 2.8-Å resolution structure from Selmer *et al.* (1), the authors fitted protein L28 into the density that previously had been interpreted as L31. This gave a better agreement with both the electron density and biochemical observations. Protein L31 could instead be placed in a density close to protein L5 in the 50S subunit. Furthermore, they could see no density for protein L36 where it had been placed previously in *Deinococcus* (1), whereas Korostelev *et al.* (2) seem to have seen it. L36 has been found to organize a conserved region of the 50S subunit (9). These observations do not lead to any alteration of the functional mechanism.

The mRNA and the Sites for tRNAs. The two structures provide snapshots of the ribosome bound to different components. In

ABSTRACT The remarkable progress of cryo-electron microscopy and crystallography in elucidating ribosomal structure and function continues. Most recently, two papers about complete 70S ribosomes from *Thermus thermophilus* at 2.8- and 3.7-Å resolution give us more details about the conformations of bound transfer RNA (tRNA) molecules; the bridges between subunits; the locations and roles of proteins, magnesium ions, and water molecules; and the dynamics of ribosomes. Very significant new insights have been gained, particularly for the tRNAs, which can only be studied in their entirety in full ribosomes.

*Corresponding author,
anders.liljas@mbfys.lu.se.

Published online October 20, 2006

10.1021/cb600407u CCC: \$33.50

© 2006 by American Chemical Society

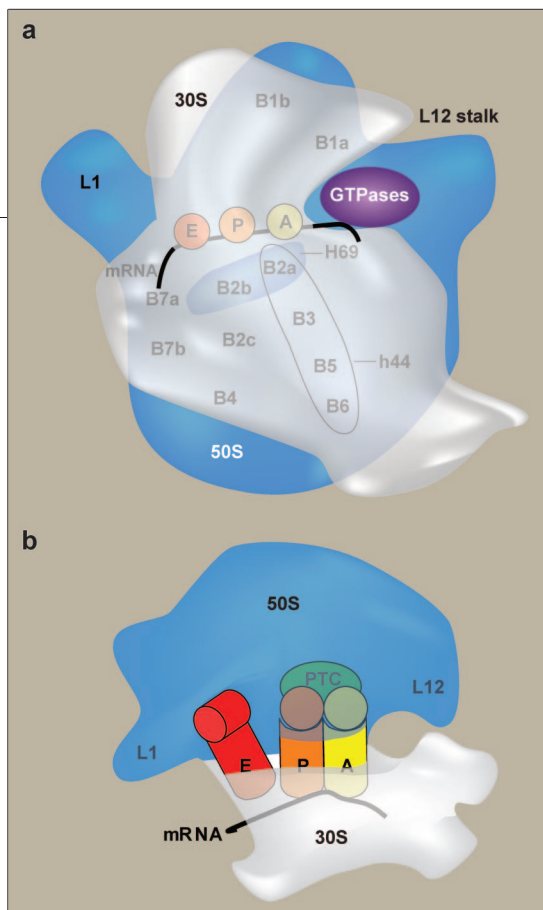


Figure 1. a) A schematic illustration of the organization of the ribosome. The large subunit (50S) is seen behind, and the small subunit (30S) is in the foreground. The ribosomal functional sites are between the ribosomal subunits. The mRNA is bound around the neck of the small subunit between the head and the body. Three sites for tRNA are shown, the A-, P-, and E-sites. The binding sites for the catalyzing translational factors, the GTPases, are also shown. Several bridges (B1–B7) between the subunits are also illustrated. Two RNA helices are of special interest, h44 (30S) and H69 (50S). They make a functionally important interaction at bridge B2a, which is at the decoding site (A-site) for tRNA. b) This top view of the ribosome shows the mRNA, the sites for tRNA, and the PTC.

the 3.7-Å ribosomal structure from Korostelev *et al.* (2), a 10-nucleotide mRNA and 2 tRNA molecules bound: a tRNA^{Phe} in the P-site and a mixture of deacylated tRNAs in the E-site. The 2.8-Å structure contains an mRNA, an aminoacyl-tRNA^{Phe} in the A-site, a deacylated initiator tRNA^{fMet} in the P-site, and a noncognate tRNA in the E-site (1). The antibiotic paromomycin was bound to stabilize the interactions at the A-site and to inhibit translocation.

As seen before, the mRNA makes a sharp kink between the codons in the A- and P-sites (1, 2). This bend is now seen to be stabilized by a Mg²⁺ ion between the closest phosphates of the two codons (1).

is distorted in both new structures compared with isolated tRNA. This seems to be due to opposing interactions by the head of the small subunit and helix H69 of the large subunit on the ASL. An interesting observation is the fact that releasing the distortion drives the tRNA toward the E-site (1).

The question of whether the E-site tRNA contacts its codon on the 30S subunit is close to an answer. In the 2.8-Å structure, this anticodon is closer to the rRNA than to the mRNA (1). Even though the codon–anticodon pair is noncognate, it would not seem possible to form a cognate codon–anticodon interaction without significant movements (Figure 1, panel b). The opposite, or

This magnesium ion is also held in place by the 30S subunit.

In the 2.8-Å structure, only the anticodon stem and loop (ASL) of the A-site tRNA was visible. Its conformation would allow the acceptor arm to be located in the PTC (1). The details of the interaction support the mechanism for decoding established earlier from studies of 30S subunits alone (10). In this mechanism, the aminoacyl-tRNA is severely bent between the ASL and the D-stem during the initial binding of aminoacyl-tRNA in complex with EF-Tu (11, 12). The fact that only the ASL is seen suggests that the tRNA retains its flexibility at this region.

Compared with the tRNA in the A-site, the P-site tRNA is bound firmly and surrounded by ribosomal RNA (rRNA) and ribosomal proteins from both subunits (1, 2). This prevents loss of the peptidyl-tRNA and also helps to maintain the reading frame. The P-site tRNA

acceptor, end of the tRNA is firmly held against the large subunit. The terminal A76 is intercalated between G2421 and A2422, and the 3'-OH is surrounded by elements of the 50S subunit. It is obvious from the narrow space that only a deacylated tRNA can be bound at this site. Also, the E-site tRNA is distorted, but less so than what is seen for the A- and P-sites. The strain placed on the tRNA from the initial binding to the ribosome in complex with EF-Tu and GTP could thus be gradually decreased in the subsequent sites.

The PTC. Peptidyl transfer occurs on the 50S subunit (Figure 1, panel b). The acceptor ends of the two tRNAs in the A- and P-sites with their aminoacyl and peptidyl moieties are closely bound to promote the transfer of the peptide on the tRNA in the P-site to the aminoacyl residue of the tRNA in the A-site. Unlike other polymerases, RNA forms vital parts of the functional sites (2). However, the claim that the ribosome is a ribozyme is now less certain for several reasons. First, no group from the rRNA seems to be directly involved with the central activity of the ribosome, peptidyl transfer (13, 14). Second, Maguire *et al.* (15) found that the three N-terminal residues of L27 were important for full peptidyl transferase activity. In the 3.7-Å structure, the nine N-terminal residues of L27 could not be seen (2), but in the 2.8-Å structure, the L27 N-terminus was close enough to interact with A76 of the P-site tRNA (1). Archaea are different from bacteria in that they do not have protein L27 (5, 6). The rRNA performs no major catalytic role, but with the aid of the ribosomal proteins, it provides selective binding sites for mRNA, tRNAs, and protein factors (13, 14).

The Bridges between the Subunits. A primary interaction between the subunits is the three tRNAs, in particular the P-site tRNA. In addition, 12 subunit bridges exist between different ribosomal components (7). These bridges, initially identified from cryo-electron microscopy studies (16), are now seen with improved clarity. One additional bridge is seen in the 2.8-Å structure

toward the bottom of helix h44 (1). Helix h44 (30S) is part of the decoding site, and H69 (50S) is part of the peptidyl transfer site. They are also part of the ribosomal interface and the areas where the changes of subunit orientation take place during the translocation of mRNA and tRNAs. One important path for communicating correct codon–anticodon interactions in the decoding site to the GTPase binding site on the 50S subunit has been identified. In bridge B2a (Figure 1, panel a), A1913 of H69 is inserted into a pocket between h44 and the A-site tRNA and makes a hydrogen bond to the 2'OH of nucleotide 37 next to the two ribosomal bases A1492 and A1493, which participate in discriminating noncognate from cognate tRNAs. Two magnesium ions further stabilize the contacts between H69, h44, and A-site tRNA in this bridge. This interaction is not present in the empty *E. coli* ribosomes and requires a significant conformational change. Three other bridges (B5, B6, and B8) are also stabilized by magnesium ions. Bridge B2c is purely magnesium-mediated. Bridge B6 is mediated by a single solvent molecule.

Ribosomal Dynamics. Scientists have long known that translation is a dynamic process. The codons of the mRNA and the tRNA molecules are translocated successively from the A- to the P- to the E-sites. In the process, they change conformation. The neck of the small subunit allows movements of the head in different states of the functional cycle. Individual ribosomal components also undergo dynamic structural changes that can be inferred from the lack of density for several components. It is now clear that nucleotides A1339 and G1338 in the head of the small subunit, as well as nucleotide 790 on the platform of the small subunit, interact with the ASL of the P-site tRNA and prevent it from moving into the E-site (1, 2, 8). Obviously, this gate needs to open during translocation. Protein L1 interacts with the E-site tRNA (1, 2, 5–7). To release this deacylated tRNA, L1 undergoes

a large conformational change to the one observed in empty ribosomes (1, 2, 8). This is seen in greater detail in the new structures.

Moving forward. Researchers in the field of ribosomes have reached yet another goal: to see, at high resolution, the complete ribosome with bound mRNA and tRNAs. Many of the previous findings are now more firmly established. The improved resolution also leads to greater detail, corrections of details, and identification of distorted or strained structures. Thus, the tRNAs do not conform to the structure seen in isolation but deviate to a decreasing extent from the A-site to the E-site. The E-site tRNA does not seem to interact with its codon. Crystallographic data shows that in the PTC, a protein, L27, interacts with A76 of the P-site tRNA. The improved detail of the 70S structures will provide a richer source for precise biochemical experimentation. Much is known, but much remains to be explored by crystallography. Several translation factors have been studied on the ribosome, but key factors remain to be analyzed, primarily the GTPases. So far, no useful crystals with any of the translational GTPases have been reported. Their binding site seems to be an important site for crystal contacts with 70S ribosomes. A crystal structure with one of these GTPases would most likely also elucidate some of the now invisible proteins at the L12 stalk in functional interactions.

REFERENCES

- Selmer, M., Dunham, C. M., Murphy, F. V., IV, Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA, *Science* 313, 1935–1942.
- Korostelev, A., Trakhanov, S., Laurberg, M., and Noller, H. (2006) Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements, *Cell* 126, 1065–1077.
- Wimberly, B. T., Brodersen, D. F., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vornrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature* 407, 327–339.
- Schlünzen, F., Tocilj, A., Zariwach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution, *Cell* 102, 615–623.
- Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution, *Science* 289, 905–920.
- Harms, J., Schlunzen, F., Zariwach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium, *Cell* 107, 679–688.
- Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., and Noller, H. F. (2001) Crystal structure of the ribosome at 5.5 Å resolution, *Science* 292, 883–896.
- Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structure of the bacterial ribosome at 3.5 Å resolution, *Science* 310, 827–834.
- Maeder, C., and Draper, D. E. (2005) A small protein unique to bacteria organizes rRNA tertiary structure over an extensive region of the 50 S ribosomal subunit, *J. Mol. Biol.* 354, 436–446.
- Ogle, J. M., Murphy, F. V., Tarry, M. J., and Ramakrishnan, V. (2002) Selection of tRNA by the ribosome requires a transition from an open to a closed form, *Cell* 111, 721–732.
- Valle, M., Sengupta, J., Swami, N. K., Grassucci, R. A., Burkhardt, N., Nierhaus, K. H., Agrawal, R. K., and Frank, J. (2002) Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process, *EMBO J.* 21, 3557–3567.
- Valle, M., Zavialov, A., Li, W., Staggs, S. M., Sengupta, J., Nielsen, R. C., Nissen, P., Harvey, S. C., Ehrenberg, M., and Frank, J. (2003) Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy, *Nat. Struct. Biol.* 10, 899–906.
- Schmeing, T. M., Huang, K. S., Strobel, S. A., and Steitz, T. A. (2005) An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA, *Nature* 438, 520–523.
- Trobro, S., and Åquist, J. (2005) Mechanism of peptide bond synthesis on the ribosome, *Proc. Natl. Acad. Sci. U.S.A.* 102, 12395–12400.
- Maguire, B. A., Beniaminov, A. D., Ramu, H., Mankin, A. S., and Zimmermann, R. A. (2005) A protein component at the heart of an RNA machine: the importance of protein L27 for the function of the bacterial ribosome, *Mol. Cell* 20, 427–435.
- Frank, J., Verschoor, A., Li, Y., Zhu, J., Lata, R. K., Rademacher, M., Penczek, P., Grassucci, R., Agrawal, R. K., and Srivastava, S. (1995) A model of the translational apparatus based on a three-dimensional reconstruction of the *Escherichia coli* ribosome, *Biochem. Cell Biol.* 73, 757–765.

Turning G Proteins On and Off Using Peptide Ligands

William W. Ja^{†,||}, Ofer Wiser^{*,||,**}, Ryan J. Austin[†], Lily Y. Jan[†], and Richard W. Roberts^{§,*}

[†]Division of Biology, California Institute of Technology, Pasadena, California 91125, [‡]Howard Hughes Medical Institute, Departments of Physiology and Biochemistry, University of California, San Francisco, 1550 Fourth Street, San Francisco, California 94143 and

[§]Department of Chemistry and Mork Family Department of Chemical Engineering, University of Southern California, Los Angeles, California 90089, ^{**}Current address: Goldyne Savad Institute of Gene Therapy, Hadassah University Hospital, P.O. Box 12000, Jerusalem 91120, Israel, ^{||}These authors contributed equally to this work

ABSTRACT Intracellular G α subunits represent potential therapeutic targets for a number of diseases. Here we describe three classes of new molecules that modulate G protein signaling by direct targeting of G α . Using messenger RNA display, we have identified unique peptide sequences that bind G α_{i1} . Functionally, individual peptides were found that either enhance or repress basal levels of G protein-activated inwardly rectifying potassium (GIRK) channel signaling, a downstream effector of G protein activation, indicating that the peptides directly turn G proteins on or off *in vivo*. A third functional class acts as a signaling attenuator; basal GIRK channel activity is unaffected but responses to repeated G protein activation are reduced. These data demonstrate that G protein-directed ligands can achieve physiological effects similar to those resulting from classical receptor targeting and may serve as leads for developing new classes of therapeutics.

Heterotrimeric guanine nucleotide-binding proteins (G proteins), composed of α , β , and γ subunits, play a critical role in communicating extracellular signals to intracellular signal transduction pathways through membrane-spanning G protein-coupled receptors (GPCRs) (1, 2). Activation of GPCRs by extracellular agonists triggers the exchange of GDP with GTP in the G α subunit and dissociation of G $\beta\gamma$ heterodimers from G α -GTP, which both regulate multiple effectors. G $\beta\gamma$ subunits, for example, can directly regulate adenylyl cyclase, phospholipase C β isozymes, and G protein-activated inwardly rectifying potassium (GIRK) channels (3). GTP hydrolysis by the inherent G α GTPase activity, a reaction catalyzed by various GTPase-activating proteins (GAPs), returns G α to the GDP-bound state and results in reassociation with G $\beta\gamma$ and termination of signaling.

Intracellular G proteins have potential as drug targets for a number of diseases (4–7). The large number of possible combinations of α , β , and γ subunits suggests that direct G protein ligands could affect individual effector pathways and/or modify signaling kinetics with great specificity (5, 8, 9). The G protein regulatory (GPR) or GoLoco motif, for example, is a peptide guanine nucleotide dissociation inhibitor (GDI) that is implicated in receptor-independent signaling (10, 11). Other recent advances include the identification of ligands for G $\beta\gamma$ that affect downstream signaling pathways using peptide (12) or small molecule (13) libraries.

In vitro peptide selection methods have been widely successful in isolating ligands for biological targets (14, 15). Various proteins in the G protein signaling pathway have been targeted by selection libraries, including receptors and G α and G $\beta\gamma$ subunits (8). Messenger RNA (mRNA) display is a selection technique where each peptide in a library is covalently coupled with its encoding mRNA (16, 17). Previously, we used mRNA display selection to identify a peptide (R6A) and its core motif (R6A-1) that bind with high affinity and specificity to the GDP-state of G α subunits (18, 19). R6A and R6A-1 act as GDIs and compete with G $\beta\gamma$ for binding to G α_{i1} (18, 19). We hypothesized that the 9-residue R6A-1 sequence could be used as a scaffold for developing new peptide ligands with different activities and/or specificities for G α subunits. Here we design an mRNA display library based on the R6A-1 core motif and use *in vitro* selection to identify unique peptides that differentially modulate G protein signaling.

A DNA template was constructed to encode the R6A-1 peptide (DQLYWWEYL) flanked by random hexamers on each end (see Methods). Nucleotide incorporation was controlled such that each wild-type residue in the core motif was ~40–50% conserved (20). mRNA display selection was performed on N-terminally biotinylated G α_{i1} (Nb-G α_{i1}) due to the previous finding that R6A-derived peptides bind preferentially to Nb-G α_{i1} over the C-terminally biotinylated Cb-G α_{i1} (18). Aluminum fluoride (AlF) was

*Corresponding author,
richrob@usc.edu.

Received for review September 8, 2006
and accepted September 18, 2006.

Published online October 20, 2006
10.1021/cb600345k CCC: \$33.50

© 2006 by American Chemical Society

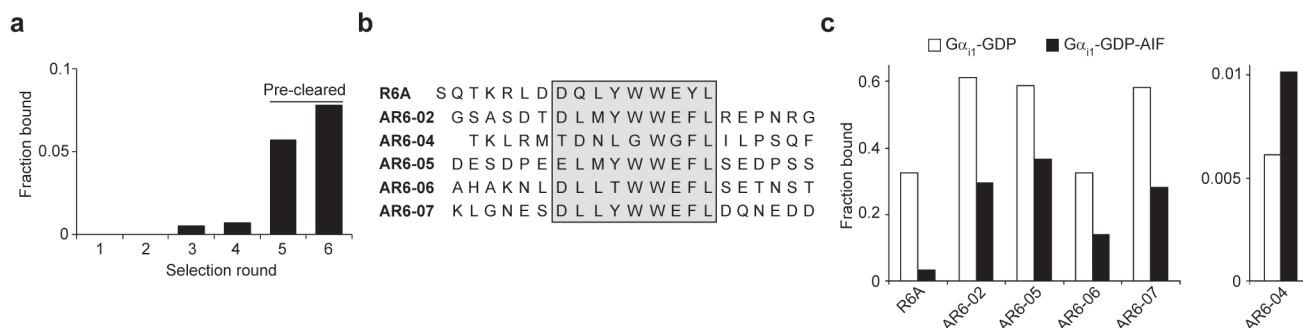


Figure 1. *In vitro* selection targeting $G\alpha_{11}$ -GDP-AIF. **a)** Fraction of ^{35}S -Met-labeled mRNA display pools from each round of selection bound to immobilized $G\alpha_{11}$ -GDP-AIF and recovered by elution with SDS. The inputs for the fifth and sixth rounds were precleared against $G\alpha_{11}$ -GDP prior to selection. **b)** Sequences of peptides used in *in vitro* studies. The region corresponding to the R6A-1 core motif is boxed (gray). The C-terminal constant region is not shown. **c)** Binding of individual, RNase-treated, ^{35}S -Met-labeled mRNA display fusions to $G\alpha_{11}$ -GDP or $G\alpha_{11}$ -GDP-AIF. Except for AR6-04, all tested peptides have a preference for binding to $G\alpha_{11}$ -GDP.

supplemented into the selection buffer to attempt to select for peptides specific for the GDP-AIF state of $G\alpha_{11}$, a transition state mimic of GTP hydrolysis (21, 22). AIF (either as AIF₃ or AIF₄⁻) has been shown to activate $G\alpha$ subunits, preventing association with $G\beta\gamma$ heterodimers, and GAPs have been shown to bind exclusively to this transition state mimic. Six rounds of selection were performed, and significant binding was observed by the third round (Figure 1, panel a). On the basis of the starting library complexity of $\sim 2 \times 10^{13}$ and a maximum enrichment of 10,000-fold per round, we estimate that the third-round input pool contained >100,000 unique, $G\alpha_{11}$ -binding peptide sequences. To enrich for peptides specific for the AIF-bound state of $G\alpha_{11}$, the fifth- and sixth-round pools were precleared against $G\alpha_{11}$ -GDP prior to selection against $G\alpha_{11}$ -GDP-AIF.

DNA sequencing of clones from the sixth-round pool showed that the core 9-mer was primarily conserved, except for a preference for leucine instead of glutamine in the second position (Figure 1, panel b and Supplementary Table 1). The random hexamer regions showed no obvious sequence conservation, although the residues directly flanking the core motif favored several amino acids, including leucine,

aspartate, and glutamate. *In vitro* binding assays with individual clones revealed that the peptides bind ~ 1 –40% to immobilized $G\alpha_{11}$ -GDP-AIF (Figure 1, panel c). The wide range of binding may suggest that the selection was not complete or that specificity to the AIF-bound state of $G\alpha_{11}$ produces a trade-off in overall binding.

Binding assays of individual peptides to $G\alpha_{11}$ -GDP in the presence or absence of AIF

show that most peptides favor the GDP-bound state (Figure 1, panel c). Hence, the selection identified peptides with a loss of specificity compared with the original R6A sequence. Only one peptide, AR6-04, exhibited better binding in the presence of AIF, but this peptide appears to have an affinity for $G\alpha_{11}$ significantly lower than that of other peptides. Because the selected peptides bind both states, preclearing the fifth- and

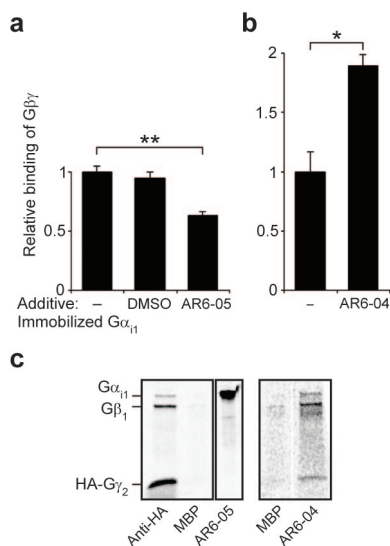


Figure 2. AR6-04 and AR6-05 $G\alpha_{11}$ -binding peptides differentially affect $G\beta\gamma$ association. **a, b)** Binding of ^{35}S -Met-labeled $G\beta_1\gamma_2$ to immobilized $G\alpha_{11}$ in the presence or absence of AR6-05 (20 μM) or AR6-04 (33 μM). AR6-05 competes with $G\beta\gamma$ for association to $G\alpha_{11}$ ($n = 4$, $p = 0.0050$), whereas AR6-04 increases $G\beta\gamma$ binding ($n = 3$, $p = 0.041$). DMSO ($\sim 1\%$, v/v) had no effect on $G\beta\gamma$ binding ($n = 3$, $p = 0.50$). **c)** Binding of ^{35}S -Met-labeled $G\alpha_{11}\beta_1\gamma_2$ to immobilized peptides. Anti-hemagglutinin (HA) antibody immunoprecipitates the HA-tagged $G\gamma_2$ subunit and confirms the presence of reconstituted heterotrimers. Immobilized maltose-binding protein (MBP) fails to pull down G proteins, while binding of $G\alpha_{11}$ to immobilized AR6-05-MBP completely precludes $G\beta_1\gamma_2$ association. AR6-04-MBP, however, pulls down the intact heterotrimer. The control MBP lane is shown again at the same contrast as the AR6-04 lane for comparison.

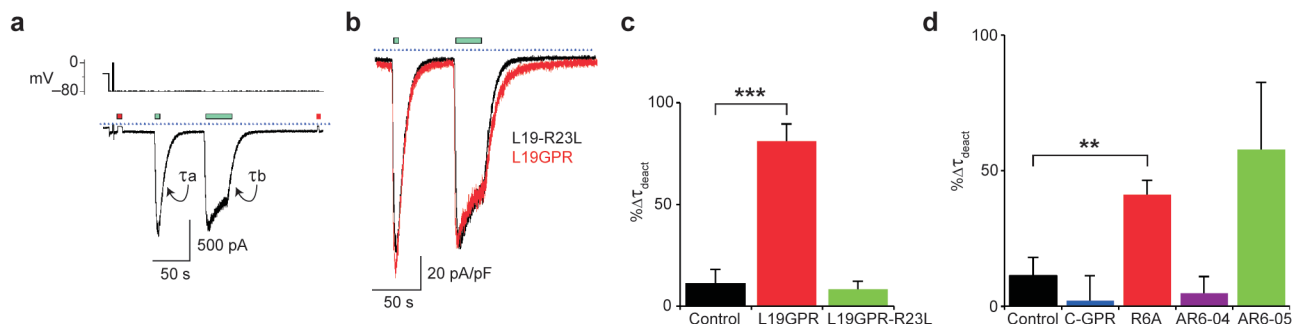


Figure 3. Effect of intracellular application of peptides on GIRK deactivation kinetics. **a)** HEK293 cells stably expressing GIRK1 and 2 and the dopamine receptor D_{2s} were recorded by whole-cell patch-clamp (see Methods). Zero- K^+ buffer (red bar) was perfused for 4 s to determine GIRK basal activity. Application of dopamine for 4 and 30 s (green bars) activated GIRK currents. Dopamine washout was followed by GIRK channel deactivation. τ_a and τ_b are the GIRK deactivation time constants following the short and long dopamine applications, respectively. The dotted line represents 0 pA. **b)** Superposition of representative current traces of cells recorded in the presence of 2 μM of the control peptide L19GPR-R23L (black) or the L19GPR peptide (red). L19GPR-R23L is a negative control peptide that contains a mutation to a critical arginine residue (26). Current traces were normalized to cell membrane capacitance and current amplitude in Zero- K^+ buffer was subtracted from current traces in High- K^+ buffer. **c)** L19GPR (2 μM) increases τ_b after prolonged dopamine application ($n = 7$, $p = 3.9 \times 10^{-5}$), whereas the control L19GPR-R23L peptide (2 μM) has no effect ($n = 2$, $p = 0.71$). $\% \Delta \tau_{\text{deact}}$ is the percentage change of τ_b from τ_a . **d)** R6A (100 μM) moderately increases τ_b ($n = 4$, $p = 0.0065$), whereas AR6-04 (40 μM , $n = 5$, $p = 0.49$) and the control C-GPR peptide (100 μM , $n = 5$, $p = 0.44$) have no effect. AR6-05 (40 μM) appears to increase τ_b ($n = 5$, $p = 0.13$), but there is significantly increased error in the kinetics measurements likely due to the effect that AR6-05 has on basal GIRK activity. In panels c and d, the control contains <0.5% (v/v) DMSO.

sixth-round pools on $G\alpha_{i1}$ -GDP may have removed the highest affinity peptides while only marginally enriching for specificity to $G\alpha_{i1}$ -GDP-AIF.

AR6-04 and AR6-05 exhibited the highest AIF/GDP-state binding ratios for $G\alpha_{i1}$ and were synthesized for further characterization. Their affinities to immobilized $G\alpha_{i1}$ -GDP were determined by surface plasmon resonance (SPR). Corresponding with the lower binding seen in the *in vitro* assays, the K_D of AR6-04 for $G\alpha_{i1}$ -GDP appears to be >10 μM . Conversely, AR6-05, with an apparent K_D of ~ 10 nM, is the highest affinity $G\alpha$ -directed peptide that we have tested, binding >6-fold better to $G\alpha_{i1}$ -GDP than our previously described R6A peptide and >20-fold better than the R6A-1 core motif (18). Whereas R6A and R6A-1 show clear 1:1 bimolecular binding kinetics, AR6-05 binding data were well fit only with a more complex kinetics model (Supplementary Figure 1).

R6A and R6A-1 were previously shown to compete with $G\beta\gamma$ heterodimers for binding

to $G\alpha_{i1}$ *in vitro* (18, 19). Binding of radiolabeled $G\beta_1\gamma_2$ to immobilized $G\alpha_{i1}$ in the presence or absence of peptides was performed to determine the peptide effects on $G\beta\gamma$ association. AR6-05 competes with $G\beta\gamma$ for binding to $G\alpha_{i1}$ (Figure 2, panel a). Like the R6A peptide, binding of $G\alpha_{i1}$ to immobilized AR6-05 precludes $G\beta\gamma$ association (Figure 2, panel c).

Surprisingly, AR6-04 appears to enhance $G\beta_1\gamma_2$ binding to $G\alpha_{i1}$. Several *in vitro* assays were performed that support this observation: (i) labeled $G\beta\gamma$ shows higher binding to immobilized $G\alpha_{i1}$ in the presence of free AR6-04 peptide (Figure 2, panel b), (ii) labeled AR6-04 peptide shows 66% higher binding to immobilized $G\alpha_{i1}$ in the presence of $G\beta_1\gamma_2$, and (iii) experiments with labeled $G\alpha_{i1}\beta_1\gamma_2$ show that immobilized AR6-04 is able to pull down all three subunits (Figure 2, panel c).

To test the activity of the peptides in a cellular context, we used a HEK293 cell line expressing GIRK1 and 2 and the dopamine D_{2s} GPCR. Previous cell culture studies have

shown that, similar to the G protein specificity observed *in vivo*, only $G_{i/o}$ -coupled receptors activate GIRK channels (23, 24). In these cells, GIRK channels are the dominant downstream effectors of released $G\beta\gamma$ subunits. The GPR consensus peptide (10) was previously shown to attenuate signaling events after an initial agonist application, without affecting basal GIRK activity (25). The authors hypothesized that the GPR peptide is able to interact with $G\alpha$ subunits only after an initial activation, which frees $G\alpha$ for peptide binding. We confirmed these results with the L19GPR peptide, which differs from the GPR consensus at a redundant residue (10, 18, 26). In the absence of peptide, the kinetics of channel deactivation (τ , deactivation time constant) are similar after short (τ_a) followed by long (τ_b) dopamine applications (10.6 ± 1.9 s, $n = 10$ and 13.7 ± 3.4 s, $n = 7$, respectively, $p = 0.68$; Figure 3, panel a). In contrast, L19GPR increased τ_b significantly compared with controls (Figure 3, panels b and c). GIRK basal activity returned to its initial values

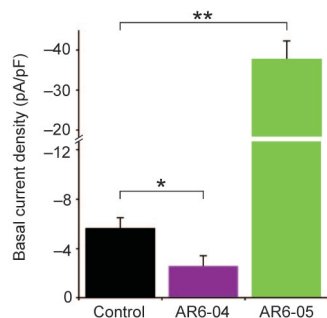


Figure 4. Intracellular application of 40 μM AR6-05 or AR6-04 increases ($n = 4$, $p = 0.0046$) or decreases ($n = 5$, $p = 0.027$) basal GIRK currents, respectively. Current densities are determined by normalization with the individual cell capacitance. The control contains <0.5% (v/v) DMSO.

after ~ 2 min from the dopamine washout, indicating that the L19GPR peptide effect is transient, since a persistent effect should have resulted in higher basal activity.

R6A exhibited effects similar to those of the L19GPR peptide. R6A increased τ_b moderately, whereas the negative control peptide C-GPR had no effect (Figure 3, panel d). R6A had minimal effect on the basal GIRK channel activity ($n = 7$, $p = 0.18$), which suggests that, like the GPR peptide, R6A is unable to dissociate $\text{G}\alpha\beta\gamma$ heterotrimers *in vivo*. In contrast to R6A, intracellular application of AR6-05 increased basal activity dramatically, suggesting that AR6-05 actively dissociates $\text{G}\beta\gamma$ from $\text{G}\alpha$ *in vivo* (Figure 4).

AR6-04 had no effect on the deactivation kinetics (Figure 3, panel d) but instead directly reduced basal GIRK activity (Figure 4). This coincides with the *in vitro* binding data, where AR6-04 stabilizes a heterotrimer complex and presumably reduces the active $\text{G}\beta\gamma$ available for GIRK channel activation. It is not clear how AR6-04 stabilizes the heterotrimer despite being selected against the $\text{G}\alpha$ subunit alone. The peptide sequence differs greatly from the original R6A-1 core motif. The flanking regions of AR6-04, however, share modest sequence similarity to

the short $\text{G}\beta\gamma$ -binding motifs previously identified (12), suggesting that other molecules that shut down G protein signaling may be constructed by fusing known $\text{G}\alpha$ - and $\text{G}\beta\gamma$ -specific ligands.

The R6A-1 based peptide library should be useful for the selection of peptides that are specific for various G protein subclasses or nucleotide-bound states. Distinct functions, such as specificity for $\text{G}\alpha_{i1}$ over other $\text{G}\alpha$ subunits, may not yet be identified from the large number of unique $\text{G}\alpha_{i1}$ -binding peptides recovered in our selection. While it is clear that AR6-04 and AR6-05 affect GIRK channel activity, an effector of $\text{G}\beta\gamma$, the peptide effects on $\text{G}\alpha$ -regulated pathways and $\text{G}\alpha$ nucleotide-bound states have yet to be determined. For example, because GAPs have been shown to catalyze GTP hydrolysis by stabilizing a transition state (21, 22), selected peptides that bind $\text{G}\alpha_{i1}$ -GDP-AIF may act as small-molecule GAPs. Although further technological advances are necessary for the facile conversion of peptides to therapeutics, determining the mechanism of action of the AR6-04 and AR6-05 peptides will facilitate the rational design of more potent modulators of G protein signaling for use as biological tools and potential drug leads.

METHODS

Materials. Human complementary DNA (cDNA) clones for G proteins were obtained from the University of Missouri-Rolla cDNA Resource Center (www.cdna.org) in the pcDNA3.1+ vector (Invitrogen). The $\text{G}\gamma_2$ cDNA vector encoded an N-terminal HA tag. Anti-HA monoclonal antibody (clone HA-7) was obtained from Sigma. Expression of ^{35}S -Met labeled G proteins by *in vitro* translation was performed as described previously (19).

mRNA Display Selection. The doped R6A-1 library was constructed by polymerase chain reaction amplification of oligo 115.1 [5'-AGC AGA CAG ACT AGT GTA ACC GCC (SNN)₆ (S13) (641) (542) (521) (521) (641) (S13) (543) (642) (SNN)₆ CAT TGT AAT TGT AAA TAG TAA TTG TCC C; 1 = 7:1:1:1, 2 = 1:7:1:1, 3 = 1:1:7:1, 4 = 1:1:1:7, A:C:G:T; 5 = 9:1, 6 = 1:9, C:G; N = A, C, G, OR T; S = C or G (ratios have been adjusted for synthesis incorporation rates)] with primers 4777FP (5'-GGA TTC TAA TAC GAC TCA TAG GGA CAA TTA CTA TTT ACA ATT AC) and 22.9 (5'-AGC AGA CAG ACT AGT GTA ACC G) to produce double-stranded DNA encoding

M-X₆-DQLYWWEYL-X₆-GGYTSLSA, with the core 9-residues conserved ~ 40 –50%. Sequencing of randomly chosen clones from the initial pool revealed a distribution of wild-type residues in the core motif that agreed with theoretical calculations (data not shown). *In vitro* transcription, ligation of the mRNA to the puromycin linker, and purification of the RNA-F30P template were performed as described previously, except that the splint oligo 23.8 (5'-TTT TTT TTT TTN AGC AGA CAG AC) was used for the ligation reaction (18). RNA-peptide fusions were prepared, purified on oligo-dT cellulose, reverse-transcribed with oligo 22.9, and selected against immobilized Nb-G α_{i1} as described previously using a modified selection buffer [25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-KOH at pH 7.5, 150 mM NaCl, 0.05% (v/v) Tween 20, 1 mM β -mercaptoethanol, 10 μM GDP, 20 μM EDTA, 5 mM MgCl₂, 10 mM NaF, 25 μM AlCl₃, 0.05% (w/v) bovine serum albumin, and 1 mg mL⁻¹ yeast transfer RNA] (18).

RNA-Peptide Fusion Binding Assay. Purified RNase-treated mRNA display peptide fusions of individual clones were assayed for binding as described previously (27). Briefly, aliquots of ^{35}S -labeled fusions were added to ~ 15 μL of Nb-G α_{i1} (~ 15 μg protein) on streptavidin agarose (immobilized NeutrAvidin on agarose, Pierce) in 1 mL of selection buffer. After binding for 1 h, the matrices were washed with 3×0.6 mL of selection buffer in a 0.45 μm cellulose acetate spin filter (CoStar Spin-X, Corning). Input ^{35}S counts for binding assays were determined by scintillation counting of the washes and the matrix. Bound ^{35}S counts were divided by the input counts to calculate the fraction bound. Binding of RNase-treated peptide fusions to the immobilization matrix alone was <0.001 . Assays for binding to $\text{G}\alpha_{i1}$ -GDP were performed in selection buffer without AIF.

In Vitro Peptide Studies. Peptides were synthesized with the first three residues of the C-terminal constant tag (GGY) and purified by Bio-Synthesis, Inc. Three additional C-terminal lysines were added to the AR6-04 peptide to enhance solubility. SPR affinity measurements were made on immobilized Nb-G α_{i1} as described previously (18). Peptide effects on $\text{G}\beta\gamma$ association with immobilized Nb-G α_{i1} were assayed by mixing an aliquot of ^{35}S -labeled $\text{G}\beta_1\gamma_2$ with ~ 15 μL of immobilized Nb-G α_{i1} in 1 mL of selection buffer without AIF. After rotating at 4 $^\circ\text{C}$ for 1 h, samples were washed with 3×0.6 mL of the binding buffer in a spin filter, as described above. Binding was determined by scintillation counting and scaled to the amount of $\text{G}\beta\gamma$ pulled down in the absence of peptide. Data are background subtracted from binding to matrix without immobilized $\text{G}\alpha_{i1}$ ($\sim 10\%$ of overall binding). AIF reduced $\text{G}\beta_1\gamma_2$ pull down on Nb-G α_{i1} to $\sim 50\%$.

G $\alpha\beta\gamma$ Heterotrimer Immunoprecipitation. AR6-04 and AR6-05 were expressed as MBP fusion proteins and immobilized by random amine coupling on CNBr-Sepharose 4B (GE Healthcare) as described previously (19). ^{35}S -Labeled $\text{G}\alpha\beta\gamma$ heterotrimer was immunoprecipitated with anti-HA monoclonal antibody and protein G-sepharose or pulled down on immobilized MBP as described

previously (19). Recovered proteins were separated by SDS-PAGE. Gels were imaged by autoradiography (Storm PhosphorImager, GE Healthcare).

Electrophysiology. We used HEK293 cell lines stably expressing GIRK1, GIRK2a, and the $G_{i/o}$ -coupled dopamine (D_{2s}) receptor (23). The pipet solution contained 107 mM KCl, 1.2 mM $MgCl_2$, 1 mM $CaCl_2$, 10 mM EGTA, 5 mM HEPES at pH 7.4, 2 mM MgATP, and 0.3 mM Na_2GTP . Peptides were added to the pipet solution immediately prior to recording. The final DMSO concentration was 0.5% (v/v) or less. The bath solution contained 2.6 mM $CaCl_2$, 1.2 mM $MgCl_2$, 5 mM HEPES at pH 7.4, and either 140 mM KCl (High- K^+) or 140 mM NaCl (Zero- K^+). Membrane currents were recorded in a whole-cell patch-clamp mode with an Axopatch 200B amplifier (Axon Instruments) and a patch pipet resistance of 2.5–4.5 M Ω . Data were filtered at 1 kHz and digitized at 5 kHz. Cell capacitance was 12–18 pF, and series resistance (5–20 M Ω) was at least 75% compensated on-line. Current recording was acquired after equilibration for ~5 min in gap-free mode at -80 mV. Dopamine (2 μ M, Sigma) was applied in bath solution via an N_2 -pressurized perfusion system (ALA Scientific Instruments).

Data Analysis. Data acquisition and analysis was done by a Digidata 1200A interface (Axon Instruments) and pClamp 8.2 and Microcal Origin 6.0 software. The deactivation time constants (τ) were determined in pClamp (standard exponential). Currents were averaged over 17 ms to reduce 60 Hz background noise. All data are presented as mean \pm SEM. Statistical significance was determined by non-paired, two-tailed Student's *t* tests. Significance in figures: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.0001$.

Acknowledgment: We thank D. S. Waugh (National Cancer Institute at Frederick) for providing the original pDW363 vector and P. J. Bjorkman for the use of the Biacore 2000 instrument. This work was supported by funding from the National Institutes of Health (GM60416) and the Beckman Foundation (R.W.R.), the Howard Hughes Medical Institute and National Institute of Mental Health (MH63981) (L.Y.J.), in the Silvio Conte Center of Neuroscience at the University of California-San Francisco, the European Molecular Biology Organization (O.W.), and a Department of Defense National Defense Science and Engineering Graduate Fellowship (W.W.J.).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals, *Annu. Rev. Biochem.* **56**, 615–649.
- Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G protein pathways, *Science* **296**, 1636–1639.
- Clapham, D. E., and Neer, E. J. (1997) G protein β subunits, *Annu. Rev. Pharmacol. Toxicol.* **37**, 167–203.
- Nürnberg, B., Tögel, W., Krause, G., Storm, R., Breitweg-Lehmann, E., and Schunack, W. (1999) Non-peptide G-protein activators as promising tools in cell biology and potential drug leads, *Eur. J. Med. Chem.* **34**, 5–30.
- Höller, C., Freissmuth, M., and Nanoff, C. (1999) G proteins as drug targets, *Cell. Mol. Life Sci.* **55**, 257–270.
- Freissmuth, M., Waldhoer, M., Bofill-Cardona, E., and Nanoff, C. (1999) G protein antagonists, *Trends Pharmacol. Sci.* **20**, 237–245.
- Spiegel, A. M., and Weinstein, L. S. (2004) Inherited diseases involving G proteins and G protein-coupled receptors, *Annu. Rev. Med.* **55**, 27–39.
- Ja, W. W., and Roberts, R. W. (2005) G-protein-directed ligand discovery with peptide combinatorial libraries, *Trends Biochem. Sci.* **30**, 318–324.
- Chahdi, A., Daeffler, L., Gies, J. P., and Landry, Y. (1998) Drugs interacting with G protein α subunits: selectivity and perspectives, *Fundam. Clin. Pharmacol.* **12**, 121–132.
- Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2000) Stabilization of the GDP-bound conformation of $G_{i\alpha}$ by a peptide derived from the G-protein regulatory motif of AGS3, *J. Biol. Chem.* **275**, 33193–33196.
- Willard, F. S., Kimple, R. J., and Siderovski, D. P. (2004) Return of the GDI: the GoLoco motif in cell division, *Annu. Rev. Biochem.* **73**, 925–951.
- Scott, J. K., Huang, S. F., Gangadhar, B. P., Samoriski, G. M., Clapp, P., Gross, R. A., Taussig, R., and Smrcka, A. V. (2001) Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein beta gamma subunits is used for recognition of a subclass of effectors, *EMBO J.* **20**, 767–776.
- Bonacci, T. M., Mathews, J. L., Yuan, C., Lehmann, D. M., Malik, S., Wu, D., Font, J. L., Bidlack, J. M., and Smrcka, A. V. (2006) Differential targeting of $G\beta\gamma$ -subunit signaling with small molecules, *Science* **312**, 443–446.
- Dower, W. J., and Mattheakis, L. C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides, *Curr. Opin. Chem. Biol.* **6**, 390–398.
- Lin, H., and Comish, V. W. (2002) Screening and selection methods for large-scale analysis of protein function, *Angew. Chem., Int. Ed.* **41**, 4402–4425.
- Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12297–12302.
- Takahashi, T. T., Austin, R. J., and Roberts, R. W. (2003) mRNA display: ligand discovery, interaction analysis and beyond, *Trends Biochem. Sci.* **28**, 159–165.
- Ja, W. W., and Roberts, R. W. (2004) *In vitro* selection of state-specific peptide modulators of G protein signaling using mRNA display, *Biochemistry* **43**, 9265–9275.
- Ja, W. W., Adhikari, A., Austin, R. J., Sprang, S. R., and Roberts, R. W. (2005) A peptide core motif for binding to heterotrimeric G protein α subunits, *J. Biol. Chem.* **280**, 32057–32060.
- LaBean, T. H., and Kauffman, S. A. (1993) Design of synthetic gene libraries encoding random sequence proteins with desired ensemble characteristics, *Protein Sci.* **2**, 1249–1254.
- Berman, D. M., Kozasa, T., and Gilman, A. G. (1996) The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis, *J. Biol. Chem.* **271**, 27209–27212.
- Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) Structure of RGS4 bound to AlF_4^+ -activated $G_{i\alpha 1}$: stabilization of the transition state for GTP hydrolysis, *Cell* **89**, 251–261.
- Leaney, J. L., Milligan, G., and Tinker, A. (2000) The G protein α subunit has a key role in determining the specificity of coupling to, but not the activation of, G protein-gated inwardly rectifying K^+ channels, *J. Biol. Chem.* **275**, 921–929.
- Benians, A., Leaney, J. L., and Tinker, A. (2003) Agonist unbinding from receptor dictates the nature of deactivation kinetics of G protein-gated K^+ channels, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6239–6244.
- Webb, C. K., McCudden, C. R., Willard, F. S., Kimple, R. J., Siderovski, D. P., and Oxford, G. S. (2005) D2 dopamine receptor activation of potassium channels is selectively decoupled by G_{α} -specific GoLoco motif peptides, *J. Neurochem.* **92**, 1408–1418.
- Peterson, Y. K., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2002) Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins, *J. Biol. Chem.* **277**, 6767–6770.
- Ja, W. W., Olsen, B. N., and Roberts, R. W. (2005) Epitope mapping using mRNA display and a unidirectional nested deletion library, *Protein Eng. Des. Sel.* **18**, 309–319.

Evolving the Substrate Specificity of O^6 -Alkylguanine-DNA Alkyltransferase through Loop Insertion for Applications in Molecular Imaging

Christian Heinis[†], Simone Schmitt, Maik Kindermann[‡], Guillaume Godin, and Kai Johnsson^{*}

Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Chemical Sciences and Engineering, CH-1015 Lausanne, Switzerland, [†]Current address: Division of Protein and Nucleic Acid Chemistry, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K. [‡]Current address: Sanofi-Aventis Deutschland; Chemical Sciences; D-65926 Frankfurt am Main, Germany.

Redesigning the substrate specificity of enzymes is one of the major challenges in protein engineering. As well as advancing our understanding of the relationship between structure and function, the generation of proteins with new substrate specificities has important medical and industrial applications. Classical examples are the conversion of lactate dehydrogenase into a malate dehydrogenase, which required the mutation of only one single amino acid to switch substrate specificity (1), or the redesign of trypsin into a protease with chymotrypsin-like specificity, which required much more extensive mutagenesis (2). Unfortunately, it is the latter case that is representative of the amount of effort typically needed to engineer enzyme substrate specificity. The protocols that are commonly employed to alter the substrate specificity and selectivity of catalysts have been iteratively improved over the past decades, to the point where a new activity can be introduced into an existing unrelated protein scaffold, a recent example being the introduction of β -lactamase activity into the $\alpha\beta/\beta\alpha$ metallohydrolase scaffold of glyoxalase II (3). However, a recurring problem in protein engineering is that the desired change in substrate specificity can rarely be achieved by a small number of point mutations and more often requires extensive probing of the protein-sequence space (4). Frequently, directed evolution of enzymes based on libraries generated by conventional approaches such as error-prone polymerase chain reaction (PCR) or saturation mutagenesis does not yield an enzyme with the desired specificity at all. The difficulty of altering the substrate specific-

ABSTRACT We introduce a strategy for evolving protein substrate specificity by the insertion of random amino acid loops into the protein backbone. Application of this strategy to human O^6 -alkylguanine-DNA alkyltransferase (AGT) led to the isolation of mutants that react with the non-natural substrate O^6 -propargylguanine. Libraries generated by conventional random or targeted saturation mutagenesis, by contrast, did not yield any mutants with activity towards this new substrate. The strategy of loop insertion to alter enzyme specificity should be general and applicable to other classes of proteins. An important application of the isolated AGT mutant is in molecular imaging, where the mutant and parental AGTs are used to label two different AGT fusion proteins with different fluorophores in the same living cell or *in vitro*. This allowed the establishment of fluorescence-based assays to detect protein–protein interactions and measure enzymatic activities.

*Corresponding author,
kai.johnsson@epfl.ch.

Received for review July 27, 2006
and accepted September 15, 2006.

Published online October 6, 2006
10.1021/cb6003146 CCC: \$33.50

© 2006 by American Chemical Society

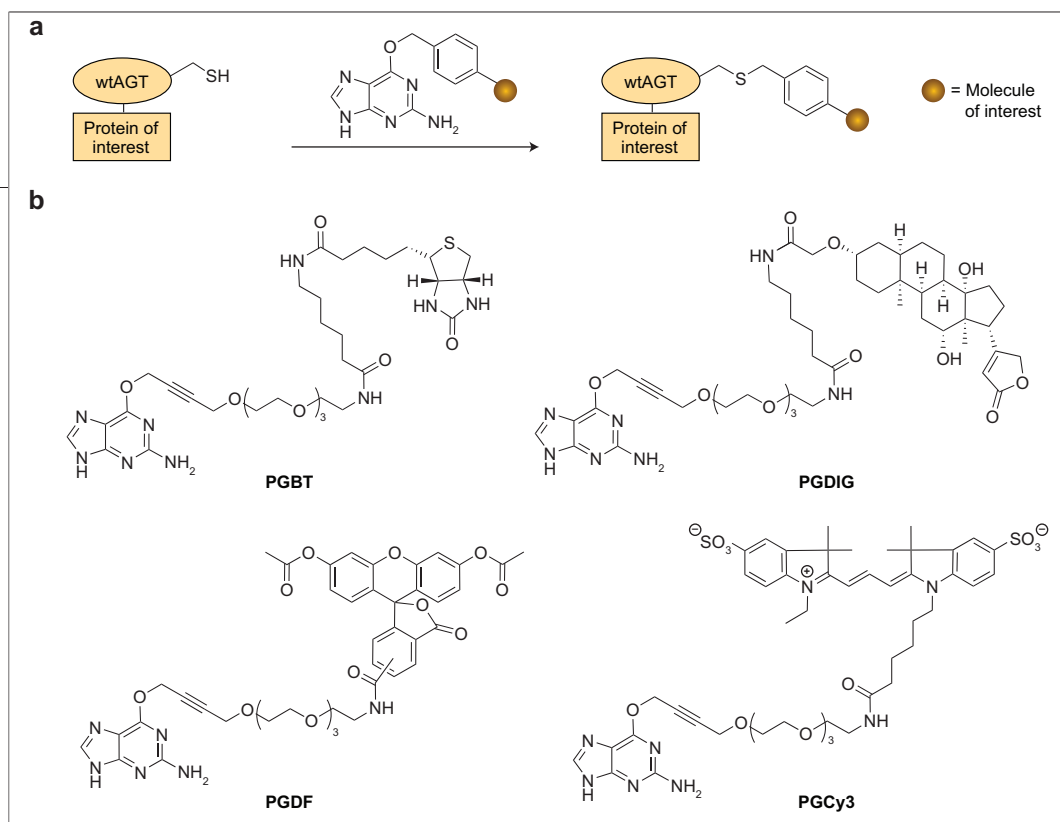


Figure 1. Mechanism and substrates used for labeling of AGT fusion proteins. a) General mechanism for the labeling of AGT fusion proteins with O^6 -benzylguanine (BG) derivatives. b) Structure of labeled PG derivatives: PGBT, PGDIG, PGDF, and PGCy3.

ity of enzymes also became apparent in this work as we tried to reprogram the substrate specificity of the DNA repair protein O^6 -alkylguanine-DNA alkyltransferase (AGT). AGT repairs O^6 -alkylated guanine in DNA by irreversibly transferring the alkyl group to its active cysteine residue. AGT has previously been exploited by our group as a tag that can be covalently labeled with chemical probes in living cells. The labeling is based on the reaction of AGT fusion proteins with O^6 -benzylguanine (BG) derivatives carrying a probe attached to the 4-position of the benzyl ring, leading to specific transfer of the probe to the fusion protein (Figure 1, panel a) (5, 6). AGT mutants with orthogonal substrate specificities could be used for simultaneous and specific labeling of different AGT fusion proteins with different synthetic probes, which would be particularly attractive for applications in molecular imaging. We show in this work that attempts to alter the substrate specificity of AGT by directed evolution using protein repertoires generated by traditional mutagenesis techniques (saturation and random mutagenesis) failed. As an alternative to these mutagenesis strategies, we tested here a new approach in which additional amino acid sequences were inserted into the polypeptide chain of the protein to provide new substrate binding elements. AGT mutants with completely random amino acid loops on the surface were generated, and those that catalyzed the desired reaction

were identified through selection and screening techniques. The best AGT mutant isolated can be used in conjunction with previously generated AGT mutants to achieve selective labeling of two different fusion proteins with two different fluorophores *in vitro* and in living cells. Such protein pairs can be used to explore protein–protein interactions or enzymatic activities using FRET measurements.

RESULTS

Substrate Design. We sought a substrate that does not react with the AGT mutants in current use, but that might be able to react with an appropriately engineered AGT. Numerous O^6 -alkylguanine derivatives have been previously synthesized for the inhibition of AGT in tumor therapy (7–10). We focused on O^6 -propargylguanine derivatives (PG; Figure 1, panel b) as potential substrates for an appropriately mutated AGT because PG has been reported to be a very poor inhibitor of human wild-type AGT (7). Furthermore, its reactivity in S_N2 reactions should be comparable to that of benzylguanine. The low reactivity of PG with AGT most likely arises from the poor binding of the substrate to the active site of AGT, which we reasoned could be improved by redesigning the substrate binding site. The difference in size between PG and BG also has the advantage that a mutant that reacts with PG might not react with the more

bulky substrate BG, thus leading to two orthogonal AGT–substrate pairs based on steric complementation (11).

PG derivatives were synthesized in which PG was coupled to biotin (PGBT), digoxigenin (PGDIG), diacetylfluorescein (PGDF), and Cy3 (PGCy3) (Figure 1, panel b; see Supporting Information for syntheses of these molecules). Immunoassay of wild-type AGT and previously evolved improved AGT mutants confirmed that PGBT and PGDIG are indeed very poor substrates. PGBT did not compete with BG substrates for reaction with AGT even when used at a 300-fold higher concentration (results not shown). A precise rate for the reaction of AGT with PGBT could not be determined under these conditions because very little product was detected, even after long incubations at high substrate concentrations. We estimated an upper limit of $0.5 \text{ s}^{-1} \text{ M}^{-1}$ for the apparent second-order rate constant of the reaction of the mutant ^MAGT (4) with PGBT. This value is ~ 1000 -fold lower than the rate constant of ^MAGT for BG. ^MAGT is an engineered mutant with improved expression properties, reduced DNA binding, and increased activity towards BG substrates.

Selection of Activities from Random and Saturation Mutagenesis Libraries. Numerous directed evolution strategies have been applied in search of enzyme variants that exhibit activities towards new substrates. In the majority of attempts, one or several amino acid residues were exchanged, either through random or saturation mutagenesis, to obtain mutants that display activities towards the desired substrates. Such classical enzyme engineering strategies have been particularly successful in the improvement of the robustness of catalysts (e.g., thermal stability, solubility, activity in organic solvents) (12) to increase the catalytic turnover of enzymes or for the creation of enantioselective catalysts (13). Following these examples, two repertoires of mutated AGT were generated by random and saturation mutagenesis of the AGT mutant ^MAGT (4). Members of the randomly mutated library (4×10^7 transformants) contained an average of 2–3 amino acid mutations per gene, as determined by sequencing 10 clones chosen at random. For the second library, amino acids that are in close proximity to the benzyl group (residues 138–143, 157, 159, and 160) were conservatively mutated using degenerate primers to increase the chance of encoding the wild-type amino acid in each position (5.5×10^7 transformants). The AGT mutants were displayed on fila-

mentous phage as a fusion of the minor coat protein pIII. AGT-phage reacting with PGDIG were isolated on a solid support coated with anti-digoxigenin antibody. In earlier work we have successfully used this strategy to evolve AGT mutants with increased activity towards BG (14, 15) and mutants resistant to an inhibitor of wild-type AGT (4). Four consecutive phage panning rounds were performed with the two libraries in parallel using either 2 or 20 μM PGDIG substrate. The number of phage isolated after each round of selection was similar to the level of phage isolated in reactions without the substrate (10^5 – 10^6 transducing units (t.u.)), indicating that the activities of the selected proteins were not improved. Activity measurements of 100 mutants of each library isolated in the fourth selection round confirmed that none of the mutants reacted with the PGDIG substrate with a rate constant significantly $>0.5 \text{ s}^{-1} \text{ M}^{-1}$.

Insertion of Amino Acid Loops into the Polypeptide Chain of AGT. Because mutation of the existing amino acids in AGT failed to generate activity towards the PG substrate, we envisioned adding new sequence space to AGT by inserting additional amino acid loops into the protein. We reasoned that libraries of AGT with random loops inserted into selected regions of the polypeptide chain could offer a rich source of potential catalysts, as the catalytic center of the enzyme is expected to remain functional. In fact, it has been demonstrated in simulations that DNA insertion is a key step in searching protein space in natural evolution (16). Furthermore, a number of studies have shown that amino acid stretches inserted into selected surface exposed regions of a protein (e.g., loop regions) often leave the structure and the function of a protein unperturbed (17). Insertion of random sequences of 120 amino acids into a surface loop of RNase H, for example, left the activity of 10% of the mutants intact (18), whereas TEM-1 β -lactamase can accommodate insertions of random sequences in two loops surrounding its active site without compromising its activity (19, 20). We chose three positions in AGT that are in close proximity to the substrate binding site but are surface exposed. Three libraries were generated by inserting a stretch of five random amino acids into the polypeptide chain of ^MAGT at each of the three positions (Figure 2). The length of the inserted loop was chosen to be long enough to form a turn and hence to link the opened ends of the enzymes polypeptide chain without impairing protein folding. For library 1, peptide loops were inserted into a turn between the short β strand 7

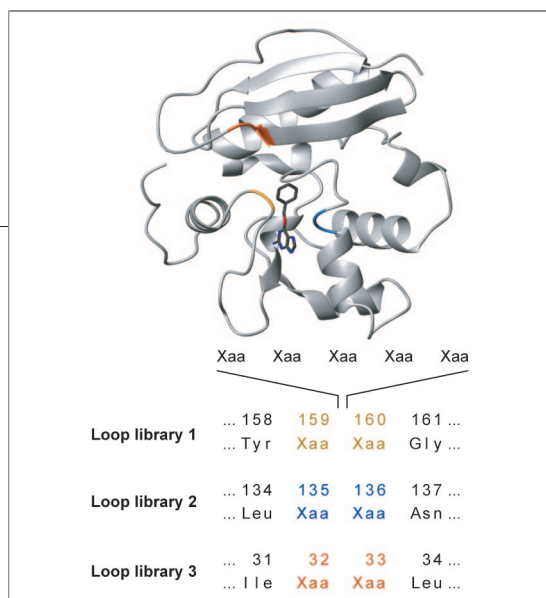


Figure 2. Structure of AGT with BG modeled into the active site (14). For the creation of the loop libraries, stretches of five random amino acids were inserted between amino acids in the highlighted regions. One amino acid on each side of the insert was also randomized in each library.

and α helix 8 (amino acid residues 159/160); for library 2, amino acids were inserted at the terminus of helix 6, close to the asparagine hinge (amino acid residues 135/136); and for library 3, the additional amino acids were inserted at the terminus of the third β strand of the N-terminal domain of AGT (amino acid residues 32/33). One amino acid on either side of the inserted pentapeptide was also randomized, on the basis of the assumption that these junctions might require a different conformation to accommodate the inserted loop. Transformation of the three libraries into bacterial cells resulted in library sizes of 5.5×10^7 (library 1), 9.2×10^7 (library 2), and 2.5×10^7 (library 3) with $>95\%$ of clones containing the gene of an AGT mutant. The number of AGT mutants that retained their activity against the BG substrate (and hence their structural integrity) was estimated by phage capture experiments with the substrate BG digoxigenin (BGDIG) and phage mixtures of each library. Phages displaying the mutagenized AGT were enriched 150-fold (library 1) and 10-fold (libraries 2 and 3) above background levels at a BGDIG concentration of $2 \mu\text{M}$. This indicates that, in contrast to mutant libraries created by saturation mutagenesis of active site residues, a very large portion of the library remained catalytically active.

Four consecutive phage panning rounds were performed with all three libraries in parallel using either 2 or $20 \mu\text{M}$ PGDIG substrate. Phage titers were measured after each round of selection, and an increase in the titer from library 1 was noticed after round 3 (20- and 200-fold enrichment at 2 and $20 \mu\text{M}$ PGDIG) and round 4 (30- and 600-fold enrichment at 2 and $20 \mu\text{M}$ PGDIG),

indicating that phage enzymes with improved activities against the PGDIG substrate had become enriched. No enrichment was observed in the selections with libraries 2 and 3.

After the fourth round of selection, phage DNA was isolated from each of the libraries, and the genes coding for the enzyme mutants were cloned into the plasmid pGEX-2T for cytoplasmic expression of glutathione S-transferase (GST)-AGT fusion proteins. About 100 clones from each selected library were expressed on a small scale, and the activity was measured by immunoassay as follows: cell lysate was incubated with PGDIG for time periods ranging from 0 to 30 min and blotted onto a hydrophobic membrane. Mutants with activity towards PGDIG became labeled with DIG and were quantified with an anti-digoxigenin antibody-horseradish peroxidase (HRP) conjugate (Figure 3, panel a). Of 100 mutants selected from library 1, eight reacted efficiently with the PGDIG substrate, whereas no active mutants were found among the clones selected from libraries 2 and 3. Library 1 also contained a significantly higher percentage of clones with activity against BG compared with libraries 2 and 3, underscoring the importance of the choice of the position for the loop insertion.

Characterization of AGT Mutants. The sequences of the eight mutants with improved PGDIG reactivity revealed high homology in the added loop (Figure 3, panel b). Amino acid 159, which flanks the inserted pentapeptide at the N-terminus, and amino acid 159a, the first amino acid of the inserted loop, were found to be glycine and proline in all sequences. Positions 159c, 159d, and 160 were frequently occupied by the amino acids glycine, tryptophan, and glycine, respectively. At positions 159b and 159e, the amino acid side chains varied. One of the active mutants (clone 24) contained an insertion of only three amino acids, probably as a result of an impurity in the wobble DNA primer used in library construction. This clone also contained the conserved motif Gly-Pro-Xaa-Gly seen in the other seven active mutants, which indicates that AGT with different loop sizes can also react with PGDIG.

The mutant that showed the highest activity in the immunoassay with the PGDIG substrate (clone 9; $45 \text{ s}^{-1} \text{ M}^{-1}$) was tested for its activity towards PGBT (Figure 1, panel b), a substrate in which digoxigenin is replaced with biotin. The mutant reacted with PGBT with the same kinetics as with PGDIG ($45 \text{ s}^{-1} \text{ M}^{-1}$), indicat-

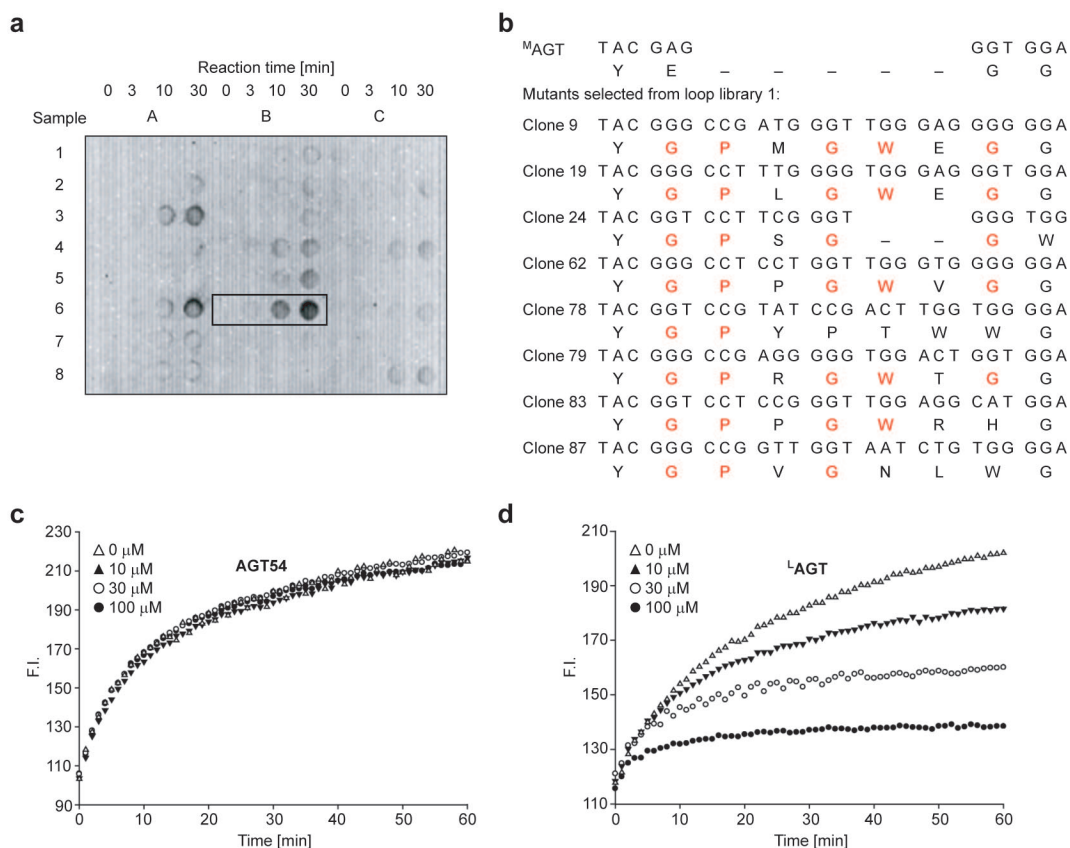


Figure 3. Characterization of selected AGT mutants. **a)** Example of a typical dot-blot activity screen. Cell lysates were incubated with PGDIG for 0, 3, 10, or 30 min and blotted onto a hydrophobic membrane. The lysates of 24 mutants (including ^MAGT; line 1, column A) were arrayed on eight lines (1–8) and three columns (A–C). Product formation was detected using an anti-digoxigenin-antibody-HRP conjugate and a chemiluminescence-based assay. As an example of an activity assay of one mutant, the dots of the most active clone in this screen (6B) are highlighted in a rectangular box. **b)** Alignment of DNA and amino acid sequences of ^LAGT and AGT mutants that reacted with PGDIG in the activity screen. Numbers were newly assigned to the mutants and do not relate to the ones in panel a. **c, d)** Time course of the reactions of AGT54 or ^LAGT with 1 μM BGCy3 in the presence of varying concentrations of PGBT. The assay is based on an increase of the fluorescence intensity (F.I.) of Cy3 upon reaction of AGT with BGCy3 (4).

ing that the evolved activity is independent of the label attached to the poly(ethylene glycol) linker. In contrast to our expectations, clone 9 retained its activity towards BG derivatives; the activity of clone 9 with BG coupled to Cy3 was measured as $480 \text{ s}^{-1} \text{ M}^{-1}$ compared to $520 \text{ s}^{-1} \text{ M}^{-1}$ for the parental clone ^MAGT (Table 1). The thermal stability of the evolved mutants was measured by incubation of purified enzyme for 15 min at temperatures ranging from 4 to 60 °C and measurement of the residual activity. Despite the insertion of a 5-amino acid loop, the selected mutants were as stable under these

conditions as their progenitor ^MAGT ($T_M \sim 55$ °C; see Supplementary Figure 1). The high sequence similarity in the loop of the active mutants isolated from library 1 indicates a unique mode of interaction between the mutants and the PG substrate. The higher activity towards the novel substrate may derive from a better positioning of PG in the active site or a higher affinity of the binding site for PG. The fact that the selected mutants retain their activity with BG argues against drastic changes to the structure of the binding site, but a detailed analysis of the mutant will require structural information.

TABLE 1. Properties of AGT mutants used in this work

Mutant	Description	Activity versus BG k_{obs} ($\text{s}^{-1} \text{M}^{-1}$)	Activity versus PG k_{obs} ($\text{s}^{-1} \text{M}^{-1}$)
^{W160} AGT	Human wild-type AGT with the single mutation G160W	400	≤ 0.5
^M AGT	Previously evolved mutant with increased activity towards BG and optimized properties for applications in protein labeling (4)	520	≤ 0.5
clone 9	Mutant selected from library 1 with loop inserted at position Glu159/Gly160 of ^M AGT	480	45
AGT54	Previously evolved mutant with increased activity towards BG (4)	1300	≤ 0.5
^L AGT	Mutant with loop from clone 9 inserted at position Glu159/Gly160 of AGT54	1250	115

To validate that the new activity originates from the inserted loop and to generate a clone with even higher activity towards PG substrates, the five additional amino acids 159a–159e, plus the two neighboring amino acids 159 and 160, were grafted into another AGT mutant, termed AGT54, which has an activity 2.5-fold higher than that of ^MAGT with BG substrates. The superior activity of AGT54 derives from mutations in the residues that interact with the purine heterocycle of BG (115, 116, 150, 151, and 152), which is identical in PG. The mutant resulting from the loop grafting, termed ^LAGT, has a second-order rate constant of $115 \text{ s}^{-1} \text{ M}^{-1}$ with PGDIG, which is elevated 2.6-fold compared with that of clone 9 (Table 1). This confirms that the inserted loop is responsible for the activity towards PG derivatives and shows that, in this case, the changes generated through loop insertion and site-directed mutagenesis are additive. The latter point should be important for future protein engineering experiments with AGT.

To assess the substrate specificity of ^LAGT and AGT54, we measured the reaction of each protein with BGCy3, a BG derivative labeled with the fluorophore Cy3, and then measured the inhibition of the reaction by competition with the alternative substrate PGBT. The reaction between AGT54 and BGCy3 was unaffected even by a 100-fold excess of PGBT over BGCy3, reflecting the lack of activity of AGT54 against PG derivatives (Figure 3, panel c). In contrast, the reaction of ^LAGT with BGCy3 was quenched by 50% by a 30-fold excess of PGBT over BGCy3 (Figure 3, panel d). This result is in agreement with the higher activity of ^LAGT towards BG derivatives compared to that towards PG derivatives.

This strategy of inserting additional amino acid sequences into surface regions of a protein to enhance activity towards a new substrate should, in principle, be applicable to other proteins. Surface loops play an important role in the substrate recognition of many enzymes. For example, in many kinases, the difference in specificity originates from variations in the loops connecting the secondary structure elements (21). The insertion of amino acid sequences into surface regions of enzymes may thus also be applied to multi-turnover catalysts. Loop insertion at random positions might be an important alternative to insertion at rationally chosen

sites. An innovative methodology to insert or delete amino acids in random positions has recently been presented (22).

Specific Double-Labeling of AGT Fusion Proteins *In Vitro* and *In Living Cells*. To achieve specific labeling of different proteins with different chemical probes, we first needed to determine whether the mutant ^LAGT could be used as a tag to label fusion proteins in living cells. ^LAGT was fused either to three copies of the simian virus 40 large T-antigen nuclear localization signal (NLS) or to β -galactosidase and transiently expressed in a Chinese hamster ovary (CHO) cell line. Cells were incubated with PGDF (Figure 1, panel b) for 30 min. Confocal microscopy showed specific labeling of ^LAGT-NLS in the nucleus and ^LAGT- β -galactosidase in the cytoplasm (Figure 4, left panels). As a control, we chose ^{W160}AGT, human wild-type AGT with the single mutation G160W, whose activity towards BG derivatives is comparable to that of ^LAGT. Cells overexpressing ^{W160}AGT did not become labeled, even when incubated with PGDF for extended periods of time (Figure 4, right panels). Next we transiently co-expressed ^LAGT- β -galactosidase and ^{W160}AGT-NLS to assess whether the evolved ^LAGT

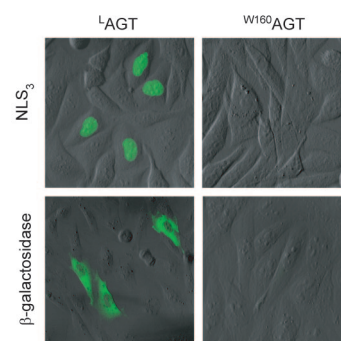


Figure 4. Selective fluorescence labeling of ^LAGT fusion proteins in living CHO cells with PGDF. Left panels: CHO cells transiently expressing ^LAGT-NLS₃ or ^LAGT- β -galactosidase were incubated with PGDF (20 μM) and analyzed by fluorescence microscopy. Right panels: CHO cells transiently expressing ^{W160}AGT-NLS₃ or ^{W160}AGT- β -galactosidase were incubated with PGDF (20 μM) and analyzed by fluorescence microscopy. Fluorescence and differential interference contrast (DIC) images of confocal micrographs are overlaid.

mutants could be selectively labeled in the presence of an AGT with wild-type specificity. Incubation of the cells with PGDF led to selective labeling of ^1AGT in the cytoplasm, while labeling of $^{W160}\text{AGT-NLS}$ in the nucleus could not be detected (Supplementary Figure 3).

Because ^1AGT can react with both PG and BG derivatives, selective labeling of ^1AGT and AGT fusion proteins relies on sequential addition of the substrates. First, the ^1AGT fusion protein must be labeled to completion with a PG derivative, and then a BG derivative can be added to label the wild-type specific AGT fusion protein. To test the feasibility of this sequential labeling strategy, an equimolar mix of pure $^1\text{AGT-GST}$ and ^{W160}AGT was incubated first with PGDF and then with BG linked to tetramethylrhodamine (BGTM). Aliquots of the reaction mixture were taken at different time points, subjected to SDS-PAGE, and analyzed for fluorescein and TMR labeling using a laser-based fluorescence scanner with appropriate filters (Figure 5, panel a). Under these conditions $^1\text{AGT-GST}$ reacts with PGDF to completion within 90 min, and BGTM added after 90 min reacts exclusively with ^{W160}AGT (Figure 5, panel a). Next we attempted a similar sequential labeling in CHO cells transiently co-expressing $^1\text{AGT-}\beta\text{-galactosidase}$ and $^{W160}\text{AGT-NLS}$. Sequential labeling with PGDF followed by a red-emitting SNARF-1 derivative of BG (BGSF) (6) led to specific labeling of the corresponding proteins in their different cellular compartments (Figure 5, panel b).

The successful reprogramming of the substrate specificity of AGT through loop insertion broadens the scope of application of the AGT labeling approach, as it allows selective labeling of two fusion proteins within the same cell. The use of spectroscopic distinguishable fluorophores is especially useful for studying protein localization and function in living cells.

Specific Double-Labeling of AGT Fusion Proteins for FRET Measurements. The ability to perform selective labeling of different AGT proteins with different spectroscopic probes may be of particular use for FRET. This powerful technique is used to investigate protein–protein interactions or to construct sensors for enzyme activities conformational changes of proteins and concentration changes of secondary messengers. So far, FRET measurements in living cells have predominantly used autofluorescent proteins (23, 24). Cyan and yellow fluorescent proteins (CFP and YFP, respectively) are currently the FRET system of choice, and the two proteins have been optimized by various groups with respect to

spectral overlap, dynamic range, and various other properties (25–27). Despite all these efforts, the optimized CFP–YFP pairs still have a number of drawbacks. For example, CFP and YFP in their current form are still reported to be weak dimers, which might affect the interaction of the corresponding fusion proteins (27). Also, it has been reported that the YFP mutant optimized for FRET applications, CyPet, folds poorly at 37 °C (27). Furthermore, efficient FRET pairs of autofluorescent proteins that are orthogonal to the CFP–YFP pairs are not yet available (27). The latter point would be important for the simultaneous monitoring of multiple biochemical processes in living cells. We wanted, therefore, to evaluate whether the new AGT mutant can be used together with parental AGT for the creation of efficient FRET pairs. The proteins FK506 binding protein (FKBP) and

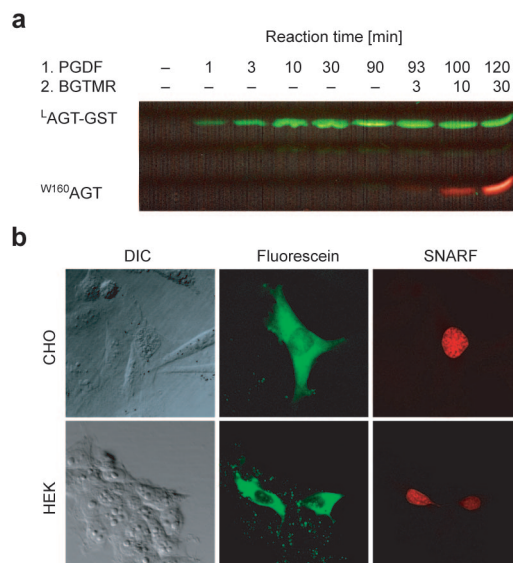


Figure 5. Specific labeling of ^1AGT and ^{W160}AGT fusion proteins with different fluorophores *in vitro* and in living cells. a) Sequential incubation of an equimolar mixture of purified $^1\text{AGT-GST}$ and ^{W160}AGT ($0.5\ \mu\text{M}$ each) with PGDF ($10\ \mu\text{M}$) and BGTM ($1\ \mu\text{M}$) and subsequent analysis using SDS-PAGE and a laser-based fluorescence scanner. Green color represents fluorescence resulting from fluorescein and red color represents fluorescence resulting from TMR. b) Sequential incubation of human embryonic kidney (HEK) and CHO cells transiently co-expressing $^1\text{AGT-}\beta\text{-galactosidase}$ and $^{W160}\text{AGT-NLS}$ with PGDF ($20\ \mu\text{M}$) and BGSF ($5\ \mu\text{M}$) and subsequent analysis using laser scanning confocal microscopy. Green color represents fluorescence resulting from fluorescein, and red color represents fluorescence resulting from SNARF-1.

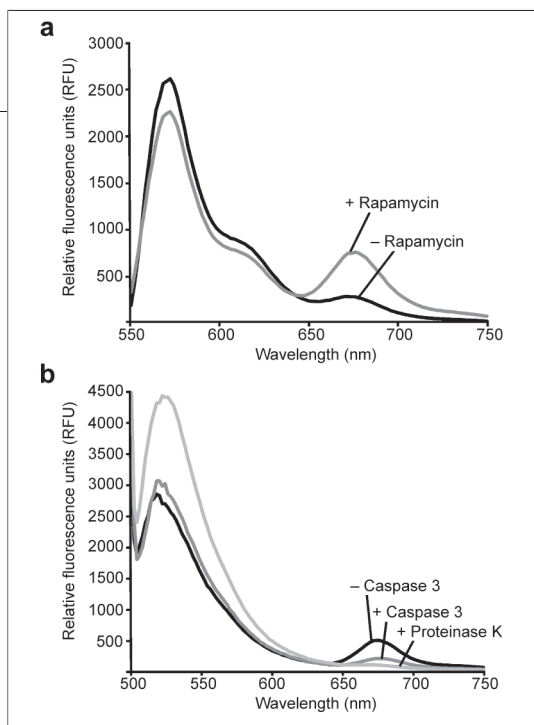


Figure 6. FRET between fluorescence labeled AGT fusion proteins. **a)** Emission spectra of PGCy3/BGCy5 labeled FKBP-^LAGT and FRB-^MAGT in the presence and absence of rapamycin. **b)** Emission spectra of PGFL/BGCy5 labeled ^LAGT-DEVD-^MAGT before and after treatment with caspase 3 or proteinase K.

incubated with equimolar concentrations of PGCy3 (Figure 1, panel b) followed by BG linked to Cy5 (BGCy5) and subsequently used without further purification. The selective labeling of FKBP-^LAGT with Cy3 and FRB-^MAGT with Cy5 was confirmed by SDS-PAGE, visualized using a fluorescence scanner (Supplementary Figure 4, panel a). Excitation of the mixture of the two labeled proteins at 520 nm and measurement of the Cy5 emission at 674 nm showed that addition of rapamycin induced a 250% increase of the Cy5 emission and a change of the emission ratio of Cy3 and Cy5 of 2.9-fold (Figure 6, panel a). When an alternative substrate pair was used, PGFL (Figure 1, panel b) and BGCy5, with excitation at 485 nm, the increase in FRET through addition of rapa-

mycin was even more pronounced (350% increase of Cy5 emission and a 3.8-fold change in the emission ratio of Cy3 and Cy5), emphasizing the need for well-optimized pairs of fluorophores; see Supplementary Figure 4, panel b.

We then tested whether the two AGT variants concatenated in a single polypeptide chain could be specifically labeled with two different fluorophores. Such constructs would be valuable for the construction of FRET-based sensors to measure concentration changes of important metabolites or enzymatic activities (23, 24). Towards this end, ^MAGT and ^LAGT were linked by the DEVD substrate motif of the endopeptidase caspase 3. The resulting protein ^LAGT-DEVD-^MAGT could be efficiently labeled with the substrates PGFL and BGCy5, as confirmed by SDS-PAGE and subsequent analysis using a fluorescence scanner; see Supplementary Figure 5.

Excitation of the doubly labeled fusion protein at 485 nm led to intramolecular FRET and emission at 640 nm (Figure 6, panel b). When the construct was subjected to proteolysis by caspase 3, under conditions where ~75% of the protein was cleaved, the FRET-dependent Cy5 emission decreased by ~60% and the emission ratio at 540 and 640 nm changed by a factor of 2.7. Complete digestion by proteinase K reduced the Cy5 emission by 77% (Figure 6, panel b and Supplementary Figure 5).

mycin was even more pronounced (350% increase of Cy5 emission and a 3.8-fold change in the emission ratio of Cy3 and Cy5), emphasizing the need for well-optimized pairs of fluorophores; see Supplementary Figure 4, panel b.

Together these experiments demonstrate that doubly labeled AGT fusion proteins are attractive candidates for the construction of FRET-based sensors. Furthermore, different AGT fusion proteins can now be tagged with a large variety of different synthetic fluorophores, which is important for the construction of efficient FRET pairs orthogonal to those of autofluorescent proteins. The emission wavelengths of fluorophores available for labeling of AGT fusion proteins in living cell ranges from blue- to far-red (28). Future experiments will focus on establishing such FRET pairs in living cells.

Together these experiments demonstrate that doubly labeled AGT fusion proteins are attractive candidates for the construction of FRET-based sensors. Furthermore, different AGT fusion proteins can now be tagged with a large variety of different synthetic fluorophores, which is important for the construction of efficient FRET pairs orthogonal to those of autofluorescent proteins. The emission wavelengths of fluorophores available for labeling of AGT fusion proteins in living cell ranges from blue- to far-red (28). Future experiments will focus on establishing such FRET pairs in living cells.

METHODS

Standard chemicals were purchased from Sigma-Aldrich. Enzymes for recombinant DNA work were purchased from MBI Fermentas or New England Biolabs. Sequences of PCR primers used in this work are listed in Supporting Information.

Phage Library Creation. The phagemid vector pAK100 (29) was used for the creation of filamentous phage displaying variants of AGT. Mutated ^MAGT genes were ligated into the two *Sfi*I restriction sites that are located between the gene of the minor coat protein pIII and its leader sequence. The creation of the random mutagenesis library is described in Gronemeyer *et al.* (15). In

this library, the gene of ^MAGT is randomly mutated with an average of 2–3 base mutations per gene. For the creation of the saturation mutagenesis library, the amino acid positions 138–143, 157, 159, and 160 of ^MAGT were mutated using partially degenerate primers in a PCR-based method using pAK100-^MAGT as a template (4). Two overlapping PCR products were formed with the primer pairs p1/p2 and p3/p4. The gel-purified products were assembled in a PCR reaction with the primer pair p1/p4.

The three libraries containing an additional stretch of five random residues (loop libraries 1–3) were created by an overlap-

extension PCR of two PCR products. The overlapping PCR products were created with the primer pairs p5/p6 and p7/p8 for the library with the insertion at the amino acid position 159/160 (loop library 1), with the primer pairs p1/p9 and p10/p4 for the library with the insertion at the amino acid position 135/136 (loop library 2), and with the primer pairs p1/p11 and p12/p4 for the library with the insertion at the amino acid position 32/33 (loop library 3). The degenerated primers were designed in a such way that the two amino acid positions flanking the five residue insert were also randomized. The gel-purified PCR products were assembled in a reaction using the primer pair p13/p14. For all libraries, ligated plasmids were purified and electroporated into *Escherichia coli* JM101 cells.

Phage Selection. Phage were produced in 50-mL cultures essentially as described in Juillerat *et al.* (14). Phage were separated from cells by centrifugation at 4000 rpm for 20 min at 4 °C. Then, 0.5 mL of the supernatant containing the phage particles (typically 10^{11} t.u./mL) was blocked with 0.5 mL of phosphate buffered saline (PBS)/milk (4% (w/v) skimmed milk in PBS pH 7.4) and incubated with PGDIG (2 or 20 μ M) for 10 min. As a positive and negative control of the phage panning procedure, phage of the same preparation were incubated in parallel either with or without the substrate BGDIG. Nonreacted substrate was removed by two rounds of PEG precipitation as follows: Phage were incubated with 250 μ L of 20% (w/v) PEG 8000, 2.5 M NaCl on ice for 10 min and spun at 13,000 rpm for 10 min at 4 °C. Phage were resuspended in 0.5 mL of PBS/milk, added to pre-blocked anti-digoxigenin magnetic beads (50 μ L of beads in 0.5 mL of PBS/milk; Roche), and incubated on a rotating wheel at RT for 30 min. The magnetic beads were then washed once with PBS/milk, six times with PBS/Tween 20 (0.01% (v/v)), and twice with PBS. The reaction tube was replaced at least once during the washing procedure. The phage were eluted with 100 μ L of 50 mM glycine buffer, pH 2.5. The pH of the eluate was neutralized by addition of 50 μ L of 1 M Tris-HCl, pH 8. Eluted phage (100 μ L) were incubated with 10 mL of exponentially growing JM101 for 30 min at 37 °C and plated on a large (20 cm diameter) agar plate containing kanamycin (70 μ g/mL). For phage titer determination, 20 μ L of the eluted phage (and eight parallel 10-fold dilutions) were incubated with 180 μ L of exponentially growing JM101 cells for 30 min at 37 °C. Then, 20 μ L of the infected cells was spotted on kanamycin-containing agar plates. The infected cells grown on large agar plates were collected and subjected to 2–3 further rounds of phage selection.

Activity Screening. Cells harboring the DNA of selected phage were pooled by resuspension of bacterial colonies resulting from phage infection with PBS. Plasmid DNA was extracted from cells, and the genes of the AGT mutants were PCR amplified with the primer pair p15/p16 to attach the sequences of the *Bam*HI and *Eco*RI restriction sites and a C-terminal polyhistidine tag. The PCR product was ligated into the pGEX-2T vector for expression of AGT mutants as fusion proteins with GST. The plasmids were electroporated into *E. coli* XL1-Blue. Individual colonies were picked and grown overnight in 3 mL of 2YT/ampicillin media. Cells of these cultures were diluted in 20 mL of Sabouraud dextrose broth (SB) media containing ampicillin to an OD₆₀₀ of 0.1 and grown at 37 °C. Protein expression was induced when the OD₆₀₀ reached 0.6 by addition of 1 mM IPTG, and cells were incubated at 24 °C for 4 h. The cells were then pelleted, resuspended in 1 mL of PBS pH 7.4 containing 1 mg/mL lysozyme and 10 μ g/mL DNase I, and sonicated for 10 s. The cell debris was removed by centrifugation for 30 min at 13,000 rpm at 4 °C, and 50 μ L of supernatant was incubated with either 20 μ M PGDIG or 2 μ M BGDIG for 0, 1, 3, 10, 30, or 90 min. The reactions were quenched by addition of 100 μ M BG. The reaction mixtures were spotted on a hydrophobic mem-

brane with a 96-well dot-blot system (Minifold I, Schleicher & Schuell Biosciences). The membrane was blocked in PBS/milk, washed with PBS, blocked in PBS/bovine serum albumin (BSA) (1 mg/mL), and incubated with preblocked anti-digoxigenin-HRP antibody conjugate (dilution, 1:1000). The membrane was washed, and the HRP was detected with a chemiluminescence-based assay.

Characterization of Active AGT Mutants. Active mutants were expressed in cultures of 100 mL of 2YT containing ampicillin (100 μ g/mL) at 24 °C for 4 h and purified by either nickel or glutathione affinity chromatography. The eluted protein was dialyzed overnight at 4 °C in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7, containing 1 mM DTT.

The activity of the purified protein was measured with the dot-blot-based assay described above using either the substrates PGDIG (20 μ M) or BGDIG (2 μ M) in combination with the anti-digoxigenin-HRP antibody conjugate or the substrates PGBT or BGBT in combination with streptavidin-HRP. The same assay was used to measure residual activity of the AGT mutants after heat incubation for 15 min at temperatures ranging from 4 to 60 °C.

The selectivity of AGT mutants towards different substrates was measured by incubating the protein simultaneously with the two substrates BGCy3 (1 μ M) and PGBT (1–200 μ M) at various molar ratios in 50 mM HEPES buffer and 1 mM DTT. The kinetics of the reaction of AGT mutants with BGCy3 was measured as described previously (4).

In Vitro FRET Assays. Construction and purification of the fusion proteins FKBP-¹AGT, FRB-^MAGT, and ¹AGT-DEVD-^MAGT is described in Supporting Information. A solution containing FKBP-¹AGT and FRB-^MAGT (5 μ M each) in 50 mM HEPES, pH 7.2, containing 1 mM DTT was incubated sequentially with either PGCy3 or PGFL (4 μ M, 3 h at RT), then PG (40 μ M, 1 h, RT), to quench non-reacted ¹AGT, and finally BGCy5 (4 μ M, 1 h, RT). FRET was measured by exciting at 485 nm (or 520 nm) and monitoring the emission at wavelengths between 500 and 750 nm (or between 530 and 750 nm) in the presence and absence of rapamycin (5 μ M) using a SPECTRAMax GEMINI fluorescence spectrometer (Molecular Devices Corp.). Purification and sequential labeling of ¹AGT-DEVD-^MAGT was performed as described for FKBP-¹AGT and FRB-^MAGT. The fusion protein was incubated with caspase 3 (5.8 μ g/mL; Sigma-Aldrich) at RT overnight or proteinase K (250 μ g/mL; AppliChem) for 20 min on ice. FRET was measured as described for the FKBP and FRB fusion proteins.

Fluorescence Labeling of AGT Fusion Proteins in Living Cells. The construction of expression plasmids for transient expression in mammalian cells and the protocols for transient transfection are given in Supporting Information. Twenty-four hours after transient transfection, AGT-deficient CHO-9-neo-C5 cells and HEK cells in F-12(Ham) medium were incubated with either the substrate PGDF (20 μ M), BGDF (5 μ M), or BGSF (5 μ M) in PBS containing 0.1% (v/v) DMSO for time periods ranging from 5 min to 2 h. In the case of transient co-expression of ^{W160}AGT-NLS₃ (4) with an ¹AGT fusion protein, PGDF (20 μ M) was added to the cells for 1 h, followed by incubation with BGSF (5 μ M) for 30 min. The cells were then washed three times with PBS to remove excess substrate. Cells were imaged in PBS 30 min after the last washing step with a laser-scanning confocal microscope (Leica TCS SP2 AOBs) using 488-nm argon, 515-nm argon, or 561-nm HeNe laser lines and a 20 \times water objective. Emission was recorded at wavelengths from 520 to 550 nm and 600 to 700 nm, respectively. DIC optics were used to image non-labeled cellular structures. Scanning speed and laser intensity were adjusted to avoid photobleaching of the fluorescent probes and damage or morphological changes of the cells.

Acknowledgments: This work was supported by the Swiss National Science Foundation, the Commission for Technology and Innovation of the Bundesamt für Berufsbildung und Technologie, and the Human Frontier Science Program. The authors would like to thank Beatrice Kunz for technical assistance and Helen O'Hare for critical reading of the manuscript.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R., and Holbrook, J. J. (1988) A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework, *Science* **242**, 1541–1544.
- Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992) Converting trypsin to chymotrypsin: the role of surface loops, *Science* **255**, 1249–1253.
- Park, H. S., Nam, S. H., Lee, J. K., Yoon, C. N., Mannervik, B., Benkovic, S. J., and Kim, H. S. (2006) Design and evolution of new catalytic activity with an existing protein scaffold, *Science* **311**, 535–538.
- Juillerat, A., Heinis, C., Sielaff, I., Barnikow, J., Jaccard, H., Kunz, B., Terskikh, A., and Johnsson, K. (2005) Engineering substrate specificity of O6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells, *ChemBioChem* **6**, 1263–1269.
- Kepler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo, *Nat. Biotechnol.* **21**, 86–89.
- Kepler, A., Pick, H., Arrivoli, C., Vogel, H., and Johnsson, K. (2004) Labeling of fusion proteins with synthetic fluorophores in live cells, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9955–9959.
- Griffin, R. J., Arris, C. E., Bleasdale, C., Boyle, F. T., Calvert, A. H., Curtin, N. J., Dalby, C., Kanugula, S., Lembic, N. K., Newell, D. R., Pegg, A. E., and Golding, B. T. (2000) Resistance-modifying agents.
- Inhibition of O(6)-alkylguanine-DNA alkyltransferase by O(6)-alkenyl-, O(6)-cycloalkenyl-, and O(6)-(2-oxoalkyl)guanines and potentiation of temozolomide cytotoxicity in vitro by O(6)-(1-cyclopentenylmethyl)guanine, *J. Med. Chem.* **43**, 4071–4083.
- Arris, C. E., Bleasdale, C., Calvert, A. H., Curtin, N. J., Dalby, C., Golding, B. T., Griffin, R. J., Lunn, J. M., Major, G. N., and Newell, D. R. (1994) Probing the active site and mechanism of action of O6-methylguanine-DNA methyltransferase with substrate analogues (O6-substituted guanines), *Anticancer Drug Des.* **9**, 401–408.
- McElhinney, R. S., Donnelly, D. J., McCormick, J. E., Kelly, J., Watson, A. J., Rafferty, J. A., Elder, R. H., Middleton, M. R., Willington, M. A., McMurry, T. B., and Margison, G. P. (1998) Inactivation of O6-alkylguanine-DNA alkyltransferase. 1. Novel O6-(hetaryl methyl) guanines having basic rings in the side chain, *J. Med. Chem.* **41**, 5265–5271.
- Moschel, R. C., McDougall, M. G., Dolan, M. E., Stine, L., and Pegg, A. E. (1992) Structural features of substituted purine derivatives compatible with depletion of human O6-alkylguanine-DNA alkyltransferase, *J. Med. Chem.* **35**, 4486–4491.
- Ulrich, S. M., Buzko, O., Shah, K., and Shokat, K. M. (2000) Towards the engineering of an orthogonal protein kinase/nucleotide triphosphate pair, *Tetrahedron* **56**, 9495–9502.
- Farinas, E. T., Bulter, T., and Arnold, F. H. (2001) Directed enzyme evolution, *Curr. Opin. Biotechnol.* **12**, 545–551.
- Reetz, M. T. (2004) Changing the enantioselectivity of enzymes by directed evolution, *Methods Enzymol.* **388**, 238–256.
- Juillerat, A., Gronemeyer, T., Kepler, A., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2003) Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo, *Chem. Biol.* **10**, 313–317.
- Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C., and Johnsson, K. (2006) Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling, *Eng., Des. Sel.* **19**, 309–316.
- Bogard, L. D., and Deem, M. W. (1999) A hierarchical approach to protein molecular evolution, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2591–2595.
- Mathonet, P., and Fastrez, J. (2004) Engineering of non-natural receptors, *Curr. Opin. Struct. Biol.* **14**, 505–511.
- Doi, N., Itaya, M., Yomo, T., Tokura, S., and Yanagawa, H. (1997) Insertion of foreign random sequences of 120 amino acid residues into an active enzyme, *FEBS Lett.* **402**, 177–180.
- Legendre, D., Soumillion, P., and Fastrez, J. (1999) Engineering a regulatable enzyme for homogeneous immunoassays, *Nat. Biotechnol.* **17**, 67–72.
- Legendre, D., Vucic, B., Hougardy, V., Girboux, A. L., Henriou, C., Van Haute, J., Soumillion, P., and Fastrez, J. (2002) TEM-1 beta-lactamase as a scaffold for protein recognition and assay, *Protein Sci.* **11**, 1506–1518.
- Hubbard, S. R., and Till, J. H. (2000) Protein tyrosine kinase structure and function, *Annu. Rev. Biochem.* **69**, 373–398.
- Murakami, H., Hohsaka, T., and Sisido, M. (2002) Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs, *Nat. Biotechnol.* **20**, 76–81.
- Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002) Creating new fluorescent probes for cell biology, *Nat. Rev. Mol. Cell Biol.* **3**, 906–918.
- Giepmans, B. N., Adams, S. R., Ellisman, M. H., and Tsien, R. Y. (2006) The fluorescent toolbox for assessing protein location and function, *Science* **312**, 217–224.
- Kremers, G. J., Goedhart, J., van Munster, E. B., and Gadella, T. W., Jr. (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius, *Biochemistry* **45**, 6570–6580.
- Nguyen, A. W., and Daugherty, P. S. (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET, *Nat. Biotechnol.* **23**, 355–360.
- Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins, *Nat. Methods* **2**, 905–909.
- Kepler, A., Arrivoli, C., Sironi, L., and Ellenberg, J. (2006) Fluorophores for live cell imaging of AGT fusion proteins across the visible spectrum, *BioTechniques* **41**, 167–175.
- Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R., and Pluckthun, A. (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system, *J. Immunol. Methods* **201**, 35–55.

Design and Characterization of a Thyroid Hormone Receptor α (TR α)-Specific Agonist

Cory A. Ocasio and Thomas S. Scanlan*

Chemistry and Chemical Biology Graduate Program and the Departments of Pharmaceutical Chemistry and Cellular & Molecular Pharmacology, University of California–San Francisco, 600 16th Street, San Francisco, California 94143-2280

Because of the extensive role of 3,5,3'-triiodo-L-thyronine (T_3) (Table 1) in vertebrate physiology, thyroid hormone analogues with tissue-specific actions, or selective thyromimetics, are highly desirable (1–3). T_3 exerts its actions by translocating into the nucleus of target cells and binding to the ligand binding domain of its cognate receptor, TR. Upon binding, TRs undergo a conformational change leading to recruitment of coregulator proteins and hence induction or repression of target gene transcription (4). TR is a member of the nuclear receptor superfamily, a group of ligand-activated transcription factors that are involved in a variety of cellular processes (4). There are two genes for TR, TR α and TR β , that are differentially processed by alternative splicing or differential promoter usage, producing an ensemble of four different major isoforms (4, 5). Although most of these isoforms are ubiquitously expressed, the TR isoform ratio in some tissues is different, giving rise to tissue-specific isoform actions (5, 6).

Selective thyromimetics bind to and preferentially activate or inactivate one of three physiologically relevant TR subtypes: TR α_1 , TR β_1 , and TR β_2 . A complete panel of thyromimetics that either activate or inactivate these receptor subtypes would provide a set of chemical tools that may assist in understanding the molecular-level basis behind TR physiology.

RESULTS AND DISCUSSION

The genesis of CO23 (Table 1) came about by first looking at a class of structurally related compounds, the TR β -selective agonists (3, 7–11). All of these compounds were derivatized from the thyronine backbone of T_3 by modifying either the moiety in the C_1 region of the scaffold, consisting of a linker capped by an acid group (*i.e.*, carboxylic, oxamic, and malonic acids), or both the C_1 region and the 3' position of the outer ring; it has been shown that placing large hydrophobic groups

ABSTRACT Thyroid hormone is a classical endocrine signaling molecule that regulates a diverse array of physiological processes ranging from energy metabolism to cardiac performance. The active form of thyroid hormone, 3,5,3'-triiodo-L-thyronine or T_3 , exerts many of its actions through its receptor, the thyroid hormone receptor (TR), of which there are two subtypes for two isoforms: TR α_1 , TR α_2 , TR β_1 , and TR β_2 . Although TR isoforms, with the exception of TR β_2 , are expressed in all tissues, they display different patterns of expression in different tissues, giving rise to tissue-specific isoform actions. Currently, several TR β -selective agonists have been developed; however, TR α -selective agonists have remained elusive. Herein, we report the synthesis and biological evaluation of CO23, the first potent thyromimetic with TR α -specific effects *in vitro* and *in vivo*.

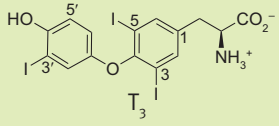
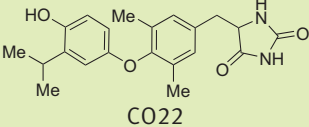
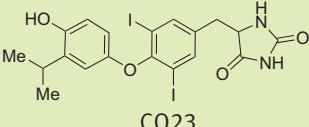
*Corresponding author,
scanlan@cgl.ucsf.edu.

Received for review July 24, 2006
and accepted September 1, 2006.

Published online October 13, 2006
10.1021/cb600311v CCC: \$33.50

© 2006 by American Chemical Society

TABLE 1. Binding affinity and potency of CO22 and CO23

Compound	K_d and EC_{50} values (nM)					
	Binding affinity (K_d) ^a		Transactivation in U2OS cells (EC_{50}) ^b		Transactivation in HeLa cells (EC_{50}) ^b	
	TR α	TR β	TR α	TR β	TR α	TR β
	0.058	0.081	2.4 ± 0.4	11 ± 2	2.4 ± 0.5	2.4 ± 0.5
	286 ± 27	280 ± 102	3870 ± 1	c	c	c
	1.2 ± 0.2	1.7 ± 0.3	34 ± 4	390 ± 3	11 ± 1	58 ± 1

^aDetermined by means of an ¹²⁵I-T₃ competitive binding assay, and data are reported as the mean K_d ± standard error of the mean, $n = 3$. ^bDetermined through use of a TRE-driven dual-luciferase reporter assay in U2OS or HeLa cells, and the data are reported as the mean EC_{50} value ± standard error of the mean, $n = 3$ (T₃, CO22, and CO23 in HeLa cells) and $n = 6$ (CO23 in U2OS cells). ^cNot applicable.

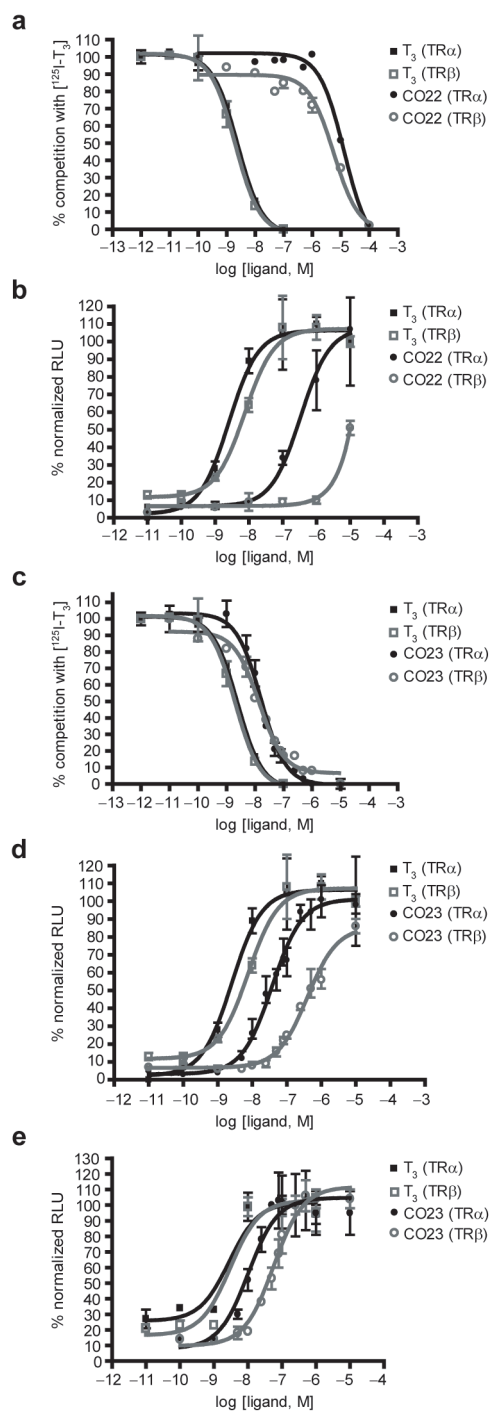
in the 3' position improves the β -selectivity for this class of thyromimetic (β). It has also been reported that compounds with a phenyl-naphthylene core bind to TR β in the sub-nanomolar range and display low to modest selectivity for TR β . Interestingly, these thyromimetics are the first to display a structure activity relationship that diverges from other thyromimetics that are based on the biaryl ether core found in the thyronine backbone of T₃ (12).

After surveying a list of β -selective agonists, we observed that heterocycles were underutilized in generating chemical diversity at the C₁ region, with few exceptions (10, 11, 13). Recent examples of thyromimetics with heterocycles in the C₁ region and outer ring include thyromimetics that incorporate inner-ring fused quinoline-2-carboxylates and indole-2-carboxylates in the C₁ region and indoles and indazoles forming heterocycle-fused outer rings (14, 15). Despite these examples, thyromimetics containing heterocycles are rare, and hence the infrequent use of heterocycles in thyromimetic design coupled with their relatively low cost and availability, chemical diversity (particularly with

respect to pK_d), and ease of synthetic incorporation into the C₁ region prompted the synthesis of a small panel of thyroid hormone analogues bearing heterocycles at the 1 position with or without a linker. The halogen-free 3,5-dimethyl,3'-isopropyl inner- and outer-ring substitution pattern was selected for this collection of thyroid hormone analogues due to ease of synthesis.

An *in vitro* evaluation of a small panel of thyroid hormone analogues identified one lead compound, CO22 (Table 1), that displays modest binding affinity and potency but at the highest concentration is as efficacious as T₃ when tested for TR α -induced transactivation in U2OS cells (Table 1 and Figure 1, panels a and b). CO22 showed significantly lower potency and efficacy when assayed for TR β -induced transactivation. In fact, at the highest concentration tested, transcriptional activity did not plateau, and it displayed about half the efficacy of T₃, indicating that CO22 is a poor agonist against TR β (Table 1 and Figure 1, panels a and b).

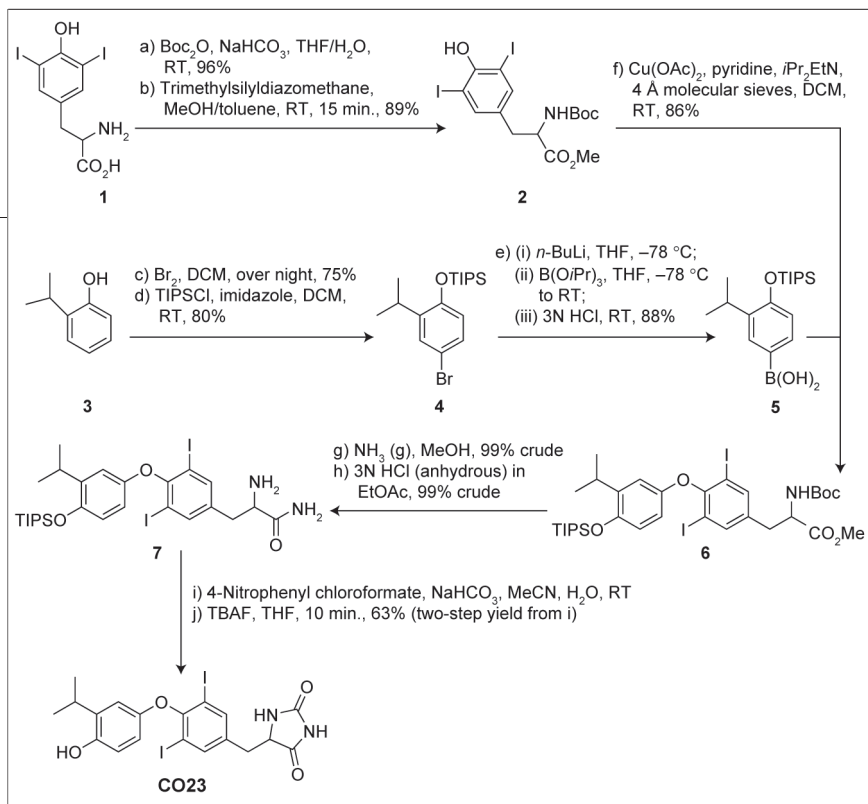
The structure–activity data revealed a way to circumvent the problem in potency of CO22. A similar thyroid



hormone analogue displaying an EC₅₀ value of ~6 nM when tested for TR_α-induced transactivation in COS-1 cells is different from CO22 in two ways: 1) there is a thiazolidinedione in the C₁ region instead of an imidazolidinedione and 2) the inner ring consists of iodides rather than methyl groups (11). Although this compound is structurally similar to CO23, it is unknown whether this compound displays TR_α specificity *in vitro* because it was not tested for TR_β-induced transactivation in COS-1 cells. In terms of CO23, the imidazolidinedione was deemed necessary for conferring TR_α-specificity, and thus we modified the inner ring of CO22 by replacing the methyl groups with iodides (Scheme 1).

CO23 proved superior to CO22 with respect to binding and transactivation because it exhibits a >200-fold improvement in binding affinity as well as a >100-fold improvement in potency (as determined by transactivation in U2OS cells) compared with CO22 (Table 1 and Figure 1, panels a–d). Although CO23, like CO22, shows no preference in binding to TR_α in an ¹²⁵I-T₃ competitive binding assay, it shows selective activation of TR_α in a DR4-driven dual-luciferase reporter assay using U2OS cells (Table 1 and Figure 1, panels c and d). Note that the replacement of the inner-ring methyl groups with iodides results in diminished TR_α selectivity (Figure 1, panel d). It is likely that the polarizability and increased electronegativity of iodides favor binding to the TR binding pocket in a steric and electronic sense, making it more potent to both receptors. However, the methyl groups, which are smaller and less polarizable and electronegative than iodides, may cause negative interactions that decrease both binding affinity and transactivation through TR but affect TR_α activation to a lesser extent than TR_β activation. With this in mind, modification of the inner-ring substituents may serve to optimize the selectivity of an already selective thymi-

Figure 1. *In vitro* evaluation of CO22 and CO23. a, c) ¹²⁵I-T₃ competitive binding curves for T₃, CO22, and CO23 against hTR_α and hTR_β; b, d) a TRE-driven dual-luciferase reporter assay showing transactivation curves for T₃, CO22, and CO23 against hTR_α and hTR_β in U2OS cells; e) a TRE-driven dual-luciferase reporter assay showing transactivation curves for T₃ and CO23 against hTR_α and hTR_β in HeLa cells. Plots show mean of triplicates with standard deviation.



Scheme 1. Synthesis of **CO23**.

metic, whereas the C_1 moiety may be responsible for setting the selectivity of a thyromimetic for $\text{TR}\alpha$ or $\text{TR}\beta$.

When determining selective activation of TRs using this cell line, it is necessary to compare the potency of the test ligand to the control ligand T_3 because T_3 shows a difference in activation of $\text{TR}\alpha_1$ and $\text{TR}\beta_1$ using a synthetic TRE-driven luciferase reporter construct; **CO23** selectively activated $\text{TR}\alpha_1$ by ~ 3 -fold in U2OS cells relative to T_3 (Table 1 and Figure 1, panel d). Due to this difference in T_3 activation of $\text{TR}\alpha_1$ and $\text{TR}\beta_1$ in U2OS cells, it was necessary to directly determine the $\text{TR}\alpha_1$ selectivity of **CO23**. In HeLa cells, a cell line where T_3 is equipotent with respect to $\text{TR}\alpha_1$ and $\text{TR}\beta_1$ activation, **CO23** showed a ~ 5 -fold preference in $\text{TR}\alpha_1$ activation (Table 1 and Figure 1, panel e).

After the $\text{TR}\alpha$ selectivity of **CO23** was established *in vitro*, **CO23** was further investigated in precociously induced amphibian metamorphosis. Amphibian metamorphosis occurs in three distinct stages: premetamorphosis, prometamorphosis, and climax (16). The premetamorphic tadpole primarily undergoes larval growth; however, the onset of prometamorphosis is marked by secretion of thyroid hormone from the developing thyroid gland, giving rise to a number of morphological and biochemical changes such as hind limb (HL) proliferation, differentiation, and induction of genes, including $\text{TR}\beta$ (16, 17). Fore leg emergence and the rapid and complete resorption of gills and tail mark metamorphic climax and, thus, complete the developmental program (16, 17).

proliferation and differentiation such as the limb buds, brain, and skin (17, 18). The gene for $\text{TR}\beta$ is itself an early-response gene of thyroid hormone, and its messenger RNA levels increase as thyroid hormone levels increase during the progression of metamorphosis, reaching a peak at metamorphic climax where death and resorption of larval tissue predominate (16, 19). This correlation between metamorphic stage, TR-isoform expression, and morphogenic response is ideal for the study of a $\text{TR}\alpha$ -selective thyromimetic.

Stage-53/54 tadpoles treated with **CO23** (50, 100, and 200 nM) experienced HL growth as extensive as or greater than that of tadpoles treated with 30 nM T_3 after 1 week (Figure 2, panel a). In addition, **CO23**-treated tadpoles exhibited less tail, gill, and head resorption at concentrations (50 and 100 nM) that yielded greater or equal HL growth as compared with 30 nM T_3 -treated tadpoles (Figure 2, panel a). In a 4-d time course, tadpoles treated with **CO23** (50, 100, and 200 nM) showed more extensive HL development after each successive day compared with tadpoles treated with 30 nM T_3 (Supplementary Figure 1, panels a and b). Treating tadpoles in a dose-dependent manner with T_3 (10, 20, and 30 nM) and **CO23** (10, 50, 100, 200, and 300 nM) showed that after 4 d, concentrations of up to 200 nM **CO23** were more effective than or as effective as 30 nM T_3 in promoting HL development. The same **CO23**-treated tadpoles displayed less head and tail resorption as did 30 nM T_3 -treated tadpoles (Supplementary Figure 2, panel a). Furthermore, to show that **CO23** works through

Using *Xenopus laevis* metamorphosis as a model system for studying biologically active and selective thyroid hormone agonists offers several advantages. First, *Xenopus* and mammalian TRs, their heterodimer partners, and the receptor-associated coregulators are structurally and functionally well conserved (17). Furthermore, *Xenopus* metamorphosis has been extensively studied, and detailed molecular events throughout the process are well known (17). *Xenopus* $\text{TR}\alpha$ ($\text{xTR}\alpha$) is expressed early on in the developing embryo and before the larval tadpole has a functional thyroid gland (17, 18). Just prior to metamorphosis, $\text{xTR}\alpha$ expression becomes widespread with locally high levels occurring in tissues that are destined to undergo

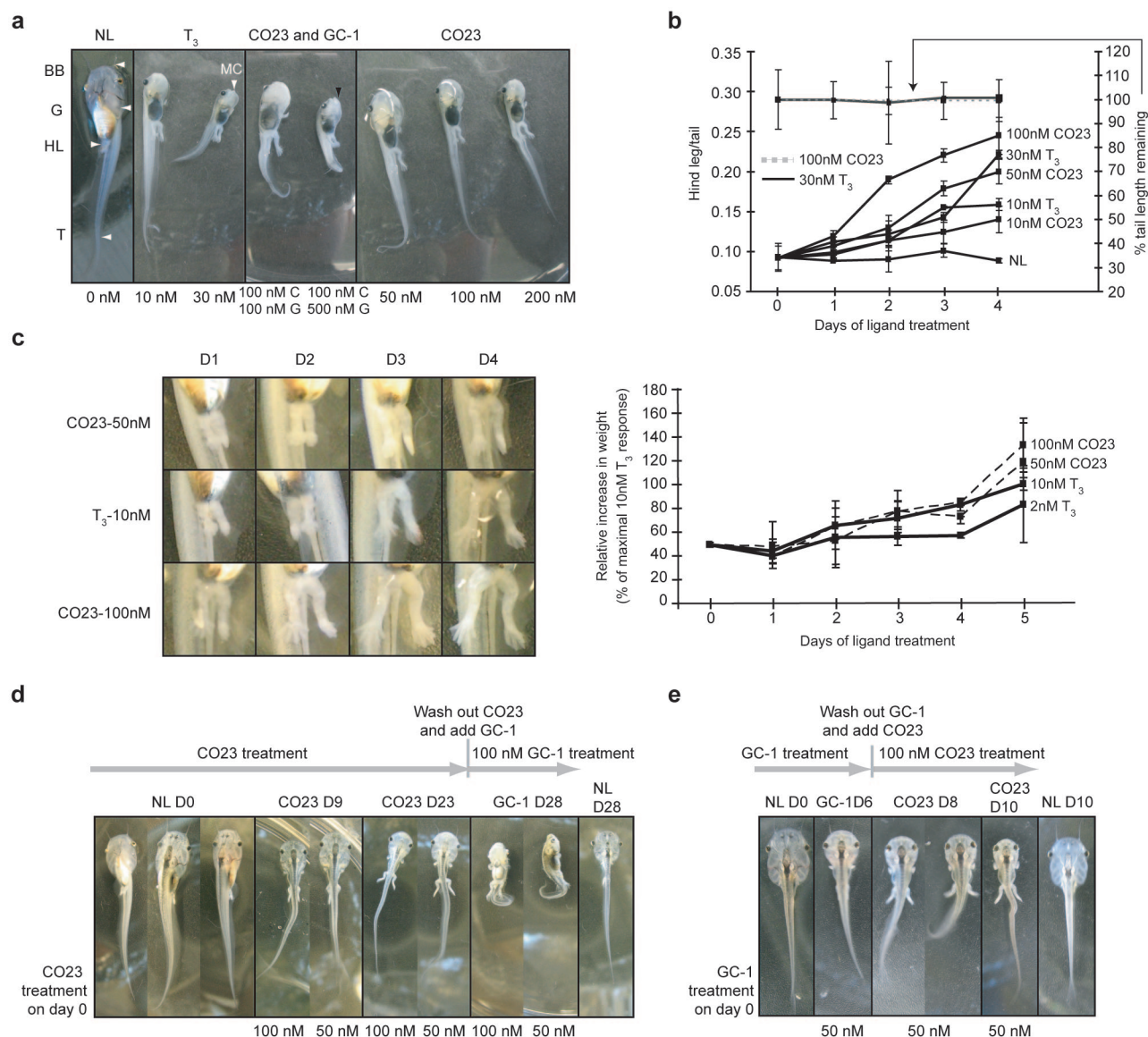


Figure 2. *In vivo* analysis of CO23. **a**) Induced metamorphosis of stage-53/54 tadpoles, $n = 3$, treated for 7 d with vehicle, T₃ (10 and 30 nM), CO23 (50, 100, and 200 nM), and a combination of CO23 (100 nM) and GC-1 (100 and 500 nM); **b**, **c**) HL development of tadpoles, $n = 3$, treated without **b**) and with **c**) 1 mM methimazole and vehicle, T₃ (2, 10, and 30 nM), and CO23 (10, 50, and 100 nM) up to 4 d as measured by H/T ratio (**b**) or relative increase in weight (**c**, represented as a percentage of the maximal T₃ response); **d**) tadpoles were treated with vehicle for 23 d, then placed in 100 nM GC-1 for 5 d more. NL, vehicle control; **e**) tadpoles were treated with vehicle for 10 d and 50 nM GC-1 for 6 d, then placed in 100 nM CO23 for 4 d more. All images are to scale for ease of comparison.

TR, tadpoles were treated with CO23 and NH-3, a TR antagonist (17). Tadpoles treated with 100 nM CO23 and 500 nM NH-3 displayed less HL and fore leg development and resorption compared with the CO23-treated control and only slightly more HL development com-

pared with the no ligand control after 5 d (Supplementary Figure 2, panel b).

A combination experiment whereby stage-53/54 tadpoles were subjected to 100 nM CO23 and the TR β -selective agonist GC-1 (100 or 500 nM) showed that tad-

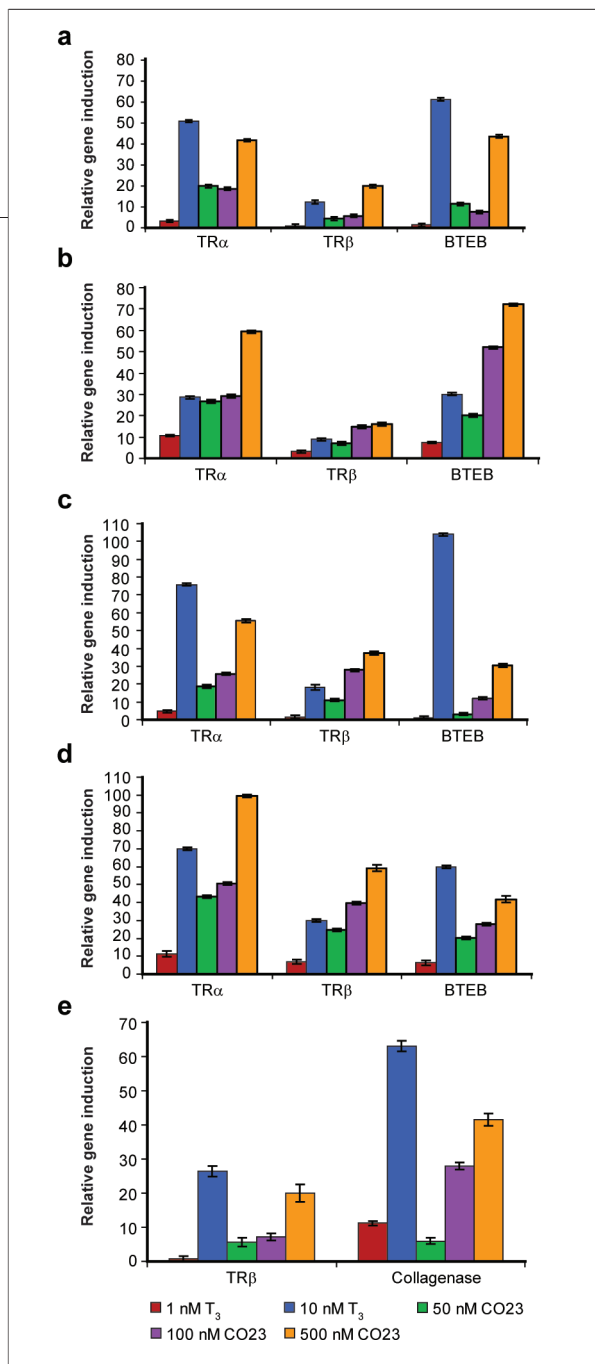


Figure 3. Relative gene induction as determined by quantitative PCR. a, b) Relative gene induction by T₃ and CO₂₃ in HL tissue after 2 (a) and 6 (b) d of treatment; c, d) relative gene induction by T₃ and CO₂₃ in head tissue after 2 (c) and 6 (d) d of treatment; e) relative gene induction by T₃ and CO₂₃ in tail tissue after 6 d of treatment.

poles treated with CO₂₃ and a low dose of GC-1 resembled 10 nM T₃-treated tadpoles and tadpoles treated with CO₂₃ and a high dose of GC-1 resembled 30 nM T₃-treated tadpoles (7, 17) (Figure 2, panel a). Therefore, this combination treatment rescues the β -isoform effects (resorption of gills, head, and bar bells, reorganization of the brain, and development of Meckel's cartilage in the jaw area) at the CO₂₃ concentration tested.

The effects of CO₂₃ on HL development were further explored by measuring the length and mass of growing HLs in response to CO₂₃ and T₃ treatment. In the first experiment, tadpoles treated with 100 nM CO₂₃ yielded the longest limbs after each successive day up to 4 d, whereas the progression in length of 50 nM CO₂₃-treated tadpole limbs was similar to that of 30 nM T₃-treated tadpole limbs. At low concentrations of T₃ and CO₂₃ (10 nM), HL lengthening was less rapid and extensive compared with the no ligand control (Figure 2, panel b). In the presence of a thyroid hormone biosynthesis inhibitor (methimazole), the HLs of CO₂₃ (50 and 100 nM)-treated tadpoles grew more rapidly than (100 nM CO₂₃) or in similar fashion (50 nM CO₂₃) to those of 10 nM T₃-treated tadpoles (Figure 2, panel c). It should be noted that 1 mM methimazole renders tadpoles much more sensitive to T₃.

Further studies with CO₂₃ and GC-1 indicated that sequential treatment with isoform-selective agonists can temporally control tadpole morphogenesis. Treating tadpoles with CO₂₃ (50 and 100 nM) for up to 23 d led to massive HL growth with slight resorption of head and gills, and replacing it with 100 nM GC-1 on the 23rd day led to the rapid resorption of head and tail, shrinking of body size, and a hunch-back appearance after 5 d of treatment (Figure 2, panel d). Conversely, treating with 50 nM GC-1 and replacing it with 100 nM CO₂₃ yielded opposite results, although some limb growth did occur with GC-1 (Figure 2, panel e). In fact, limb growth resulting from GC-1 may be attributed to GC-1 activation of TR α . This is likely because in the early stages of tadpole metamorphosis TR α is the most abundant isoform present and its activation leads to induction of TR β , which then carries out the β -isoform effects upon activation by GC-1.

Next, we sought evidence for the TR α specificity of CO₂₃ by monitoring induction of early and late thyroid hormone responsive genes. This was done by performing quantitative real-time polymerase chain reaction (rt-PCR) on complementary DNAs derived from various *Xenopus* tissues. TR β and xBTEB, a zinc-finger transcription factor, are early response genes that are both rapidly induced by T₃, and in the case of xBTEB, by GC-1 as well (17, 18). As such, xBTEB can be viewed as a TR β -induced early gene, and hence, its induction by CO₂₃ should be less intense compared with that by T₃. The next gene of interest, collagenase-3, is a late responsive gene that is up-regulated in tail tissue and should not

respond to CO23 as effectively as T_3 , particularly on day 6 (19). In this experiment, CO23 acted as hypothesized; in HL and head tissue after 2 and 6 d of treatment, CO23 (100 and 500 nM)-induced TR β levels were higher than those induced by T_3 (10 nM), except in HL tissue treated with 100 nM CO23 for 2 d (Figure 3, panels a–d). In terms of xBTEB induction, CO23 (50, 100, and 500 nM) is less effective than T_3 (10 nM) (Figure 3, panels a, c, and d). It is only in HL tissue after 6 d of treatment (late prometamorphosis) that CO23 (100 and 500 nM) induced xBTEB more effectively than T_3 (Figure 3, panel b). Up-regulation of the late response gene collagenase-3 by CO23 is also less effective at all

concentrations compared with that by T_3 (10 nM) (Figure 3, panel e). As for xTR α , only the highest dose of CO23 (500 nM) was as effective as T_3 (10 nM) at its up-regulation.

CO23 is the first thyroid hormone analogue to demonstrate TR α specificity *in vitro* and *in vivo*. The implications of this may extend to mammalian species because TR α_1 in mammals is known to regulate several cardiac parameters that maintain healthy cardiac performance. Finally, CO23, along with current thyromimetics, may serve as chemical probes of TR signaling pathways, leading to new insights regarding TR biology.

EXPERIMENTAL METHODS

General Synthesis. All chemicals used for organic synthesis were purchased from Aldrich, Sigma-Aldrich, Fluka, or Acros and were used without further purification. Anhydrous conditions were maintained under argon using standard Schlenk line techniques and oven-dried glassware. Anhydrous tetrahydrofuran (THF), dichloromethane (DCM), pyridine, and diisopropyl ethylamine were available in house and dispensable from a solvent purification system. Compounds were purified by either flash chromatography using silica gel (VWR Scientific) or preparatory thin layer chromatography (prep TLC) using Analtech prep TLC plates (20 cm \times 20 cm, 1000 μ m). ^1H NMR spectra were taken on the Varian Utility 400 MHz spectrometer in CDCl_3 or $\text{DMSO}-d_6$ solvents, and chemical shifts were reported as δ (parts per million) downfield of the internal control, trimethylsilane (TMS), for all solvents. High-resolution mass spectrometry (HRMS) using electrospray ionization was performed by the National Bio-Organic, Biomedical Mass Spectrometry Resource at UCSF, and specific rotation was determined using a Perkin Elmer 241 polarimeter. Purity of final compounds was assessed using Rainin HPLC pumps, an Altech Nucleosil 100 (C_{18}) 10 μ m 4.6 mm \times 250 mm column with a 7.5 mm guard column, and a Varian ProStar 330 photo diode array detector controlled by a Varian Star chromatography workstation. HPLC grade acetonitrile and H_2O were purchased from Fisher.

Synthesis of CO23. Replacement of the inner-ring methyl groups with iodides was accomplished by coupling of a Boc-protected diiodotyrosine methyl ester, **9**, with a triisopropylsilyl (TIPS)-protected 4-hydroxy-3-isopropylphenyl boronic acid, **12**, resulting in the biaryl ether intermediate **13** (20–23). Subsequent amidation of the methyl ester and deprotection of the Boc group set up the amino acid amide side chain for cyclization into the imidazolidinedione by 4-nitrophenyl chloroformate and water (20, 24, 25). The synthesis of CO23 was finalized by deprotection of the TIPS group with tetrabutylammonium fluoride (20) (Scheme 1). Full experimental details for both CO22 and CO23 are described in Supplementary Methods. Analytical data for CO22 and CO23 are as follows. CO22 ^1H NMR (400MHz, CDCl_3): δ 1.21 (d, 6H, J = 8.0 Hz), 2.10 (s, 6H), 2.75 (dd, 1H, J = 4.0 Hz, J = 12.0 Hz), 3.14 (heptet, 1H, J = 8.0 Hz), 3.22 (dd, 1H, J = 8.0 Hz, J = 12.0 Hz), 4.28 (dd, 1H, J = 4.0 Hz, J = 8.0 Hz), 4.97 (s, 1H), 5.57 (s, 1H), 6.22 (dd, 1H, J = 4.0 Hz, J = 8.0 Hz), 6.58 (d, 1H, J = 8.0 Hz), 6.73 (d, 1H, J = 4.0 Hz), 6.91 (s, 2H), 7.27 (s, 1H). CO22 HRMS (m/z): $[\text{M}]^+$ calcd for $\text{C}_{21}\text{H}_{24}\text{N}_2\text{O}_4$, 368.1736; found, 368.1749. CO22 HPLC (MeCN/water,

65–100%, 12 min): retention time 3.6 min; 99% pure. CO23 $[\alpha]_D^{20}$ = -18.7 (c 0.05, MeOH). CO23 ^1H NMR (400MHz, $\text{DMSO}-d_6$): δ 1.11 (d, 6H, J = 8.0 Hz), 2.81 (dd, 1H, J = 8.0 Hz, J = 14.0 Hz), 2.93 (dd, 1H, J = 4.0 Hz, J = 14.0 Hz), 3.15 (heptet, 1H, J = 8.0 Hz), 4.35 (dd, 1H, J = 4.0 Hz, J = 8.0 Hz), 6.16 (dd, 1H, J = 4.0 Hz, J = 8.0 Hz), 6.61 (d, 1H, J = 8.0 Hz), 6.63 (d, 1H, J = 4.0 Hz), 7.75 (s, 2H), 7.98 (s, 1H), 8.96 (s, 1H), 10.60 (s, 1H). CO23 HRMS (m/z): $[\text{M}]^+$ calcd for $\text{C}_{19}\text{H}_{18}\text{I}_2\text{N}_2\text{O}_4$, 591.9356; found, 591.9356. CO23 HPLC (MeCN/water, 65–100%, 12 min): retention time 3.9 min; 99% pure.

Thyroid Hormone Competition Binding Assay. Full-length hTR α_1 and hTR β_1 were expressed using a TNT T7 quick-coupled transcription translation system (Promega). Competition assays for binding of unlabeled T_3 and CO23 were performed using 1 nM ^{125}I - T_3 in a gel filtration binding assay as described (26).

Transient Transfection Assays. Human bone osteosarcoma epithelial (U2OS) cells or human uterine cervical cancer (HeLa) cells (Cell Culture Facility, UCSF) were grown to $\sim 80\%$ confluency in Dulbecco's modified Eagles (DME)/H-21, 4.5 g L^{-1} glucose medium containing 10% newborn calf serum (NCS) or fetal bovine serum (FBS), respectively (both heat-inactivated), 2 mM glutamine, 50 units mL^{-1} penicillin, and 50 $\mu\text{g mL}^{-1}$ streptomycin. Cells (~ 1.5 – 2×10^6) were collected and resuspended in 0.5 mL of electroporation buffer (Dulbecco's phosphate-buffered saline (PBS) containing 0.1% glucose and 10 mg mL^{-1} bioprene) with 1.5 μg of a TR expression vector (full-length hTR α_1 -CMV or hTR β_1 -CMV), 0.5 μg of pRL-TK constitutive *Renilla* luciferase reporter plasmid (Promega), 5 μg of a reporter plasmid containing a synthetic TR response element (DR-4) containing two copies of a direct repeat spaced by four nucleotides (AGGTCACaggAGGTC) cloned immediately upstream of a minimal thymidine kinase promoter linked to a luciferase coding sequence (7). Cells were electroporated using a Bio-Rad gene pulser at 350 V and 960 μF in 0.4 cm cuvettes, pooled in DME/F-12 Ham's 1:1 without phenol red (U2OS) or DME/H-21 (HeLa), supplemented as above except that NCS and FBS were hormone-stripped using dextrose-coated charcoal, and plated in 96-well (U2OS) or 12-well (HeLa) plates to a final density of 20,000 cells per well and 100,000 cells per well, respectively. After a 2-h incubation period, compounds in 1% DMSO were added to the cell culture medium in triplicate. After an additional 16-h incubation period, cells were harvested and assayed for luciferase activity using the Promega dual-luciferase kit (Promega) and an Analyst AD (Molecular Devices). Data normalized to the *Renilla* internal control were analyzed

with GraphPad Prism, v4, using the sigmoid dose response model to generate EC_{50} values; EC_{50} values were obtained by fitting data to the following equation: $Y = \text{bottom} + (\text{top} - \text{bottom}) / (1 + 10^{(\log EC_{50} - X) \times \text{HillSlope}})$.

Preparation of Chemicals. Stocks of T_3 , CO23, GC-1, and NH-3 were prepared with DMSO at a concentration of 10 mM and stored at -20°C until use; GC-1 and NH-3 were prepared as described previously (27, 28). All other chemicals were purchased from Sigma unless otherwise indicated. Methimazole (Aldrich) was dissolved in sterile water to a final concentration of 1 M and stored at -20°C . Aminobenzoic acid ethyl ester (0.1%, Tricaine or MS222) was made fresh in sterile ddH_2O and kept at 4°C for no longer than 1 week.

General *X. laevis* Tadpole Procedures. *X. laevis* stage-53/54 tadpoles were purchased from NASCO, Inc., and staged according to Nieuwkoop and Faber (29). Upon receipt, tadpoles were allowed to set overnight at RT ($18-25^\circ\text{C}$) in order to recover from shipping shock, after which half of the initial rearing water was replaced with $0.1\times$ Marc's Modified Ringer's (MMR) buffer ($10\times$ solution consists of 100 mM NaCl (Fisher), 2 mM KCl (Fisher), 1 mM MgCl_2 , 2 mM CaCl_2 , 0.1 mM EDTA, and 5 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid, pH 7.8) and in some cases, the tadpoles were maintained at a concentration of 1 mM methimazole. Tadpoles were ultimately maintained in fresh $0.1\times$ MMR buffer, changed every 2 d, with or without methimazole. After completion of experiments, live tadpoles were euthanized by treatment with 0.01% Tricaine, exposure to an ice bath, and either fixation in PBS containing 3.5% formalin or decapitation in order to ensure death. Animals were photographed with a Canon PowerShot A510, and images were processed with Adobe Photoshop CS, v8, and Adobe Illustrator CS, v11. All tadpole experiments were conducted in accordance with Institutional Animal Care and Use Committee approval (animal protocol no. A7228-23070-01).

Induced Metamorphosis Experiments. Stage-53/54 tadpoles were added to extra-deep Petri dishes (Fisher) in triplicate containing 50 mL of $0.1\times$ MMR buffer and vehicle or the appropriate concentration(s) and combination of ligand(s) (T_3 , CO23, GC-1, and NH-3) with or without methimazole. The final DMSO concentration was 0.1%. Induced metamorphosis experiments were repeated at least three times.

Quantification of HL Development. Groups of tadpoles were treated with the appropriate concentration of ligand and photographed live daily, every 24 h, for 4 d or sacrificed before excision of HLs. One method of quantifying HL development was through measuring the HL length (pixels) to tail length (pixels) (H/T) ratio, because this ratio correlates reasonably well with metamorphic stage and serves to normalize for differences in initial tadpole size (30). The percentage of tail length remaining was also determined for the highest ligand concentrations tested (30 nM T_3 and 100 nM CO23) in order to show that tail length remains constant and, thus, leaves the H/T ratio unaffected. These experiments were repeated twice, and each experimental point consisted of the mean H/T ratio and percent tail length remaining for three tadpoles. HL development was also monitored by recording the increase in mass as compared with day zero of HLs over a 4 d time period. Tadpole HLs from different groups of tadpoles in 1 mM methimazole were excised every 24 h, placed on pretared weighing paper, dried in ambient air for 3 h, and then weighed on a Sartorius balance accurate to 0.01 mg. These experiments were repeated twice, and each experimental point consisted of the mean increase in weight normalized to maximal T_3 response for three to five tadpoles.

Temporal Control Over Morphogenesis. These experiments were carried out in the presence of 1 mM methimazole as in induced metamorphosis experiments except that at the indicated time point, $0.1\times$ MMR buffer containing either CO23 (50

and 100 nM) or GC-1 (50 nM) was replaced with $0.1\times$ MMR containing GC-1 (50 nM) or CO23 (100 nM), respectively. Each experimental point consisted of three tadpoles, and each experiment was repeated twice.

Quantitative rt-PCR Assay. Total RNA was extracted from head, HL, and tail tissue from groups of 6–10 tadpoles using TRIzol reagent (Invitrogen) according to the manufacturer's specifications. The total RNA was processed as described previously (13), and the C_t method (Applied Biosystems User Bulletin no. 2) was employed to quantify gene induction normalized to the *Xenopus* 18S ribosomal RNA subunit (RL8) and relative to a physiological calibrator. Relative gene induction was quantified with the equation $2^{-\Delta\Delta C_T}$ in sextuplicate, and the standard deviation was calculated using the comparative method described in User Bulletin no. 2. rt-PCR reactions were carried out on a DNA Engine Opticon2, and the data were analyzed using Opticon software. Primers used to detect RL8 and collagenase-3 were the same as reported previously (13). Primers used to detect all other target genes were designed using the Primer3 website (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), and the sequences are as follows: xTR α f, 5'-CTA CGA TCC AGA CAG CGA GAC-3'; xTR α r, 5'-GTT CAA AGG CGA GAA GGT AGG-3'; xTR β f, 5'-ATG GCA ACA GAC TTG GTT TTG-3'; xTR β r, 5'-CGC ATT AAC TAT GGG AGC TTG-3'; xBTB f, 5'-CCA TCT CAA AGC CCA CTA CAG-3'; xBTB r, 5'-GAA TTG GAC CTT TTG GAC CTT-3'.

Acknowledgments: We would like to thank J. David Furlow, Eric Neff, and Cindy Chen for their advice and guidance and for taking time to critically analyze the *X. laevis* induced metamorphosis experiments. We are also grateful to Suzana T. Cunha Lima, for her technical expertise with the ^{125}I - T_3 competitive binding assay. Finally, we are grateful to the National Institutes of Health (Grant DK-52798, T.S.S.) and the Ford Foundation for financial support.

Supporting information available: This material is free of charge via the Internet.

REFERENCES

- Scanlan, T., Yoshihara, H., Nguyen, N.-H., and Chiellini, G. (2001) Selective thyromimetics: tissue-selective thyroid hormone analogs, *Curr. Opin. Drug Discovery Dev.* 4, 614–622.
- Ocasio, C., and Scanlan, T. (2005) Clinical prospects for new thyroid hormone analogues, *Curr. Opin. Endocrinol. Diabetes* 12, 363–370.
- Morkin, E., Ladenson, P., Goldman, S., and Adamson, C. (2004) Thyroid hormone analogs for treatment of hypercholesterolemia and heart failure: past, present, and future prospects, *J. Mol. Cell. Cardiol.* 37, 1137–1146.
- Greenspan, F. S., and Gardner, D. G. (2001) The thyroid gland, in *Basic and Clinical Endocrinology* (Greenspan, F., Ed.) 6th ed., pp 201–272, Lange Medical Books/McGraw-Hill, New York.
- Lazar, M. (1993) Thyroid hormone receptors: multiple forms, multiple possibilities, *Endocr. Rev.* 14, 184–193.
- Brent, G. (2000) Tissue-specific actions of thyroid hormone: insights from animal models, *Rev. Endocr. Metab. Disord.* 1, 27–33.
- Chiellini, G., Apriletti, J., Yoshihara, H., Baxter, J., Ribeiro, R., and Scanlan, T. (1998) A high-affinity subtype-selective agonist ligand for thyroid hormone receptor, *Chem. Biol.* 5, 299–306.
- Borggraeber, S., Budny, M.-J., Chiellini, G., Cunha-Lima, S., Togashi, M., Webb, P., Baxter, J., Scanlan, T., and Fletterick, R. (2003) Ligand selectivity by seeking hydrophobicity in the thyroid hormone receptor, *Proc. Natl. Acad. Sci. U.S.A.* 100, 15358–15363.
- Hayashi, M., Ohnata, H., Tamura, T., Kuroda, J., Shibata, N., Akahane, M., Moriwaki, H., Machida, N., and Mitsumori, K. (2004) Inhibitory effects of KAT-681, a liver-selective thyromimetic, on development of hepatocellular proliferative lesions in rats induced by 2-acetylaminofluorene and partial hepatectomy after diethylnitrosamine initiation, *Arch. Toxicol.* 78, 460–466.

10. Dow, R., Schneider, S., Paight, E., Hank, R., Chiang, P., Cornelius, P., Lee, E., Newsome, W., Swick, A., Spitzer, J., Hargrove, D., Patterson, T., Pandit, J., Chrunk, B., LeMotte, P., Danley, D., Rosner, M., Ammirati, M., Simons, S., Schulte, G., Tate, B., and DaSilva-Jardine, P. (2003) Discovery of a novel series of 6-azauracil-based thyroid hormone receptor ligands: potent, TR β subtype-selective thyromimetics, *Bioorg. Med. Chem. Lett.* **13**, 379–382.
11. Ebisawa, M., Inoue, N., Fukasawa, H., Sotome, T., and Kagechika, H. (1999) Thiazolidinediones with thyroid hormone receptor agonistic activity, *Chem. Pharm. Bull.* **47**, 1348–1350.
12. Hangeland, J., Friends, T., Doweiko, A., Mellström, K., Sandberg, J., Grynfarb, M., and Ryono, D. (2005) A new class of high affinity thyromimetics containing a phenyl-naphthylene core, *Bioorg. Med. Chem. Lett.* **15**, 4579–4584.
13. Chiang, Y.-C. (Pfizer Products Inc., USA) Preparation of [(hydroxyphenoxy)benzyl]thiazolidinediones and analogs as thyroid receptor ligands. *Eur. Pat. Appl.* EP-1148054, 17 Apr 2004.
14. Collazo, A.-M., Koehler, K., Garg, N., Färegårdh, M., Husman, B., Ye, L., Ljunggren, J., Mellström, K., Sandberg, J., Grynfarb, M., Ahola, H., and Malm, J. (2006) Thyroid receptor ligands. Part 5: Novel bicyclic agonist ligands selective for the thyroid hormone receptor β , *Bioorg. Med. Chem. Lett.* **16**, 1240–1244.
15. Hanig, H., Woltering, M., Mueller, U., Schmidt, G., Schmeck, C., Voehringer, V., Kretschmer, A., and Pernerstorfer, J. (2005) Novel heterocyclic thyromimetics, *Bioorg. Med. Chem. Lett.* **15**, 1835–1840.
16. Kaltenbach, J. C. (1996) Endocrinology of amphibian metamorphosis, in *Metamorphosis postembryonic reprogramming of gene expression in amphibian and insect cells* (Gilbert, L., Tata, J., and Atkinson, B., Eds.), pp 403–431, Academic Press, San Diego, CA.
17. Furlow, J., Yung-Yang, H., Hsu, M., Lim, W., Ermio, D., Chiellini, G., and Scanlan, T. (2004) Induction of larval tissue resorption in *Xenopus laevis* tadpoles by the thyroid hormone receptor agonist GC-1, *J. Biol. Chem.* **279**, 26555–26562.
18. Lim, W., Nguyen, N.-H., Yang, H., Scanlan, T., and Furlow, J. (2002) A thyroid hormone antagonist that inhibits thyroid hormone action *in vivo*, *J. Biol. Chem.* **277**, 35664–35670.
19. Wang, Z., and Brown, D. (1993) Thyroid hormone-induced gene expression program for amphibian tail resorption, *J. Biol. Chem.* **268**, 16270–16278.
20. Hart, M., Suchland, K., Miyakawa, M., Bunzow, J., Grandy, D., and Scanlan, T. (2006) Trace amine receptor-associated agonists: synthesis and evaluation of thyronamines and related analogues, *J. Med. Chem.* **49**, 1101–1112.
21. Hodnett, N. (2003) Trimethylsilyldiazomethane, *Synlett* **13**, 2095–2096.
22. Yoshihara, H., Apriletti, J., Baxter, J., and Scanlan, T. (2001) A designed antagonist of the thyroid hormone receptor, *Bioorg. Med. Chem. Lett.* **11**, 2821–2825.
23. Evans, D., Katz, J., and West, T. (1998) Synthesis of diaryl ethers through the copper-promoted arylation of phenols with arylboronic acids - an expedient synthesis of thyroxine, *Tetrahedron Lett.* **39**, 2937–2940.
24. Mishra, A., Panwar, P., Chopra, M., Sharma, R., and Chatal, J. (2003) Synthesis of novel bifunctional Schiff-base ligands derived from condensation of 1-(*p*-nitrobenzyl)ethylenediamine and 2-(*p*-nitrobenzyl)-3-monooxo-1,4,7-triazapeptane with salicylaldehyde, *New J. Chem.* **27**, 1054–1058.
25. Yamaguchi, J., Harada, M., Kondo, T., Noda, T., and Suyama, T. (2003) A facile method for preparation of optically active hydantoin, *Chem. Lett.* **32**, 372–373.
26. Apriletti, J., Baxter, J., Lau, K., and West, B. (1995) Expression of the Rat α 1 Thyroid Hormone Receptor Ligand Binding Domain in *Escherichia coli* and the Use of a Ligand-Induced Conformation Change as a Method for Its Purification to Homogeneity, *Protein Expression Purif.* **6**, 363–370.
27. Chiellini, G., Nguyen, N.-H., Yoshihara, H., and Scanlan, T. (2000) Improved synthesis of the iodine-free thyromimetic GC-1, *Bioorg. Med. Chem. Lett.* **10**, 2607–2611.
28. Nguyen, N.-H., Apriletti, J., Cunha-Lima, S., Webb, P., Baxter, J., and Scanlan, T. (2002) Rational design and synthesis of a novel thyroid hormone antagonist that blocks coactivator recruitment, *J. Med. Chem.* **45**, 3310–3320.
29. Nieuwkoop, P. D., and Faber, J. (1994) *Normal Table of Xenopus laevis (Daudin): A Systematical and Chronological Survey of the Development From the Fertilized Egg Till the End of Metamorphosis*, 2nd ed., Garland Publishing, New York and London.
30. Li, H. C. (1978) *Hormonal Proteins and Peptides*, 6th ed., Academic Press, New York.

Conjugating Berberine to a Multidrug Resistance Pump Inhibitor Creates an Effective Antimicrobial

Anthony R. Ball[†], Gabriele Casadei[†], Siritron Samosorn[‡], John B. Bremner^{*,*}, Frederick M. Ausubel[§], Terence I. Moy[§], and Kim Lewis^{†,*}

[†]Department of Biology and Antimicrobial Discovery Center, Northeastern University, Boston, Massachusetts 02115, [‡]Institute for Biomolecular Science and Department of Chemistry, University of Wollongong, Wollongong, New South Wales 2522, Australia, and

[§]Department of Genetics, Harvard Medical School, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

ABSTRACT In bacteria, multidrug-resistance pumps (MDRs) confer resistance to chemically unrelated amphipathic toxins. A major challenge in developing efficacious antibiotics is identifying antimicrobial compounds that are not rapidly pumped out of bacterial cells. The plant antimicrobial berberine, the active component of the medicinal plants echinacea and golden seal, is a cation that is readily extruded by bacterial MDRs, thereby rendering it relatively ineffective as a therapeutic agent. However, inhibition of MDR efflux causes a substantial increase in berberine antimicrobial activity, suggesting that berberine and potentially many other compounds could be more efficacious if an effective MDR pump inhibitor could be identified. Here we show that covalently linking berberine to INF₅₅, an inhibitor of Major Facilitator MDRs, results in a highly effective antimicrobial that readily accumulates in bacteria. The hybrid molecule showed good efficacy in a *Caenorhabditis elegans* model of enterococcal infection, curing worms of the pathogen.

*Corresponding authors,
k.lewis@neu.edu,
jbremner@uow.edu.au.

Received for review June 5, 2006
and accepted September 14, 2006.

Published online October 13, 2006

10.1021/cb600238x CCC: \$33.50

© 2006 by American Chemical Society

Efflux by multidrug-resistance pumps (MDRs) is a universal mechanism by which microorganisms resist a broad variety of antimicrobials (1–4). Bacterial MDRs are found in all microorganisms and make up five distinct independently evolved protein families (5). The first MDR pump described was the human ABC (ATP binding cassette)-family P-glycoprotein transporter (6), which protects a number of tissues from xenobiotics and is an essential component of the blood–brain penetration barrier (7–9). Overexpression of P-glycoprotein plays an important role in tumor resistance to chemotherapeutic agents (10). Although a few bacterial P-glycoprotein homologues have been described, such as the LmrA MDR of *Lactococcus lactis* (11), bacterial ATP-dependent MDRs are uncommon, and efflux of clinically relevant compounds is due primarily to the drug/proton antiporters of the resistance nodulation cell division (RND) type MDRs of Gram-negative species (12) and major facilitator (MF) MDRs present in all groups of microorganisms (13). Some of the MF MDRs, such as the *Staphylococcus aureus* QacA pump, are carried on transmissible genetic elements (14).

Unlike specialized transporters, MDRs recognize their substrates largely on the basis of polarity. In order to cross the lipid bilayer of the membrane, drugs must be amphipathic, containing both hydrophilic and hydrophobic components, whereas cytoplasmic compounds are hydrophilic, which prevents their escape from the cell (15, 16). Any amphipathic compound could potentially be a toxin, providing a simple basis for an MDR pump to discriminate self from harmful foreign mol-

ecules. The crystal structure of the *Escherichia coli* RND pump AcrAB revealed an unusually large “binding site” capable of accommodating a vast variety of substrates (17). At the same time, MDRs have preferences, and the best substrates for all studied MDR groups are hydrophobic cations (16, 18). A positive charge allows a molecule to accumulate in the cell, driven by the transmembrane potential. An ability to accumulate up to 1000-fold makes cations potentially highly toxic to the cells, and this threat may have been responsible for the origin of MDRs. Benzalkonium chloride is an example of a hydrophobic cation that is widely used as an antiseptic and whose efficiency is limited by MDR efflux (19). It was recently shown that creating a polymer of the hydrophobic cation hexyl pyridinium makes the compound insensitive to MDR efflux (20). Apparently, the pumps can extrude small molecules but not large polymers. This enables creation of effective “sterile surface” materials to prevent the spread of pathogens (21). At the same time, it would be very useful to have small molecules in our arsenal of potential pharmaceuticals that avoid MDR efflux. As described below, plants provide a natural example of a chemical strategy to block MDR efflux, thereby allowing antimicrobial compounds synthesized by the plant to inhibit the growth of microbial pathogens.

The alkaloid berberine is a natural product and a hydrophobic cation that is the principal component of the medicinal plants golden seal (*Hydrastis canadensis*) and echinacea (*Echinacea* species). Berberine is a potentially excellent antimicrobial, because it accumulates in cells driven by the membrane potential (19) and hits two immutable targets, the membrane and DNA (22). Accumulation of hydrophobic cations in the membrane causes leaks, and berberine is also an excellent DNA intercalator (23). Resistance to berberine is thus unlikely to develop due to target modification. It was previously shown that resistance to berberine is based on MDRs (18); for example, it is readily pumped out of *S. aureus* cells by NorA, an MDR pump responsible for efflux of cationic antiseptics and fluoroquinolones. However, *Berberis* species of plants produce, in addition to berberine, 5'-methoxyhydrnocarpin (5-MHC), an inhibitor of MF MDRs. 5-MHC strongly potentiates the action of berberine (24). The synergistic combination of an anti-

microbial with an MDR inhibitor results in an effective antimicrobial that avoids bacterial resistance.

Not surprisingly, the presence of MDRs has been an important impediment in the development of new synthetic antibiotics. One approach to solving the penetration problem has been to develop MDR inhibitors (25–27), similar to the natural strategy that plants use to combat microbial pathogens (24). A potential challenge of this approach, however, is to match pharmacokinetics and other properties of two unrelated molecules.

We reasoned that the challenge of developing an efficacious MDR inhibitor could potentially be met by covalently linking an antimicrobial compound with a MDR inhibitor to create a well-penetrating molecule. Here, we report that combining berberine with the MDR inhibitor INF₅₅ produces a novel hybrid antibacterial that is insensitive to MDR efflux.

RESULTS AND DISCUSSION

To test the concept of an antimicrobial/MDR inhibitor hybrid, we synthesized a conjugate between berberine, a hydrophobic cation that is an excellent MDR substrate, and INF₅₅, an inhibitor of MF family MDRs (25).

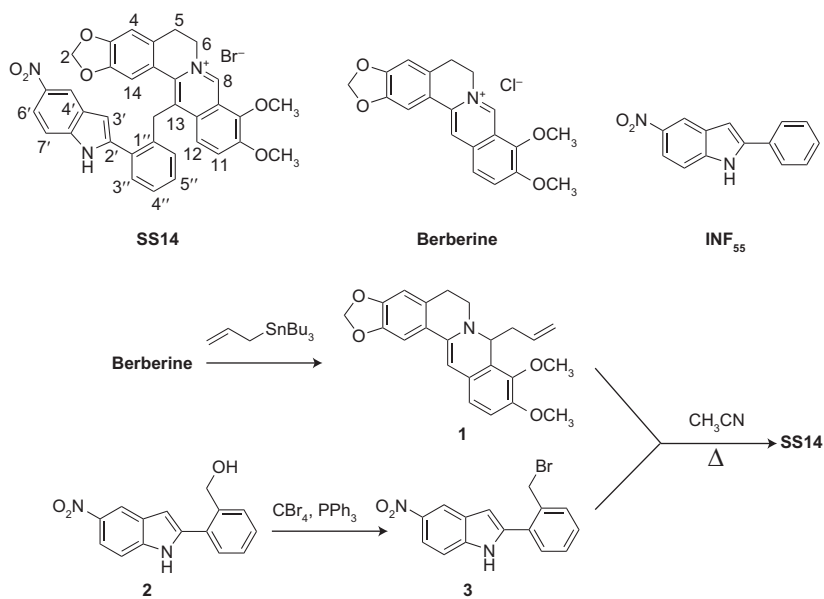


Figure 1. Synthesis of the SS14 hybrid from berberine and INF₅₅ derivatives. The hybrid compound SS14 was synthesized in fair yield by reaction of the 8-allyldihydroberberine (1) with the indole derivative (3). Compound 1 was prepared in turn from berberine as described previously, while 3 could be accessed from the indole alcohol (2).

TABLE 1. Minimum inhibitory concentrations of hybrid SS14, berberine, and MDR inhibitor

	MIC (μM)			
	Hybrid	Berberine + INF ₅₅ ^a	Berberine	INF ₅₅
<i>S. aureus</i>				
K1758 ΔnorA	3.125	6.25	40	>525
8325-4 Wild-type	3.125	12.5	325	>525
K2361 NorA++	6.25	50	>650	>525
K2378 NorA++	3.125	50	>650	>525
<i>En. faecalis</i>				
OG1RF	6.25	50	650	>525
MMH594	6.25	50	>650	>525
V583	6.25	50	>650	>525
<i>En. faecium</i>				
DO	3.125	25	80	>525
<i>B. anthracis</i>				
Sterne	6.25	12.5	325	>525
<i>B. cereus</i>				
569	3.125	12.5	325	>525
T	3.125	12.5	650	>525

^aBerberine and INF₅₅ were added at equimolar concentrations.

Design of the Hybrid Antimicrobial. Berberine and INF₅₅ were linked *via* a methylene group (Figure 1). A short linker was chosen in order to minimize any potential problems from increased steric volumes in the hybrid molecule. The hybrid compound, SS14, was synthesized by reaction of 8-allyldihydroberberine (compound **1**) with the indole derivative (compound **3**). Compound **1** was prepared in turn from berberine as described previously (28), while compound **3** could be accessed from the indole alcohol (compound **2**) (29).

Antibacterial Activity. The resulting conjugate had excellent antimicrobial activity (Table 1). When used against *S. aureus*, SS14 was 100 \times more active than berberine. The difference was even greater (200–400 fold) in the case of *S. aureus* mutants overexpressing the NorA MDR. The activity of berberine, as measured by the minimal inhibitory concentration (MIC), against an assortment of tested bacteria varied widely, from 50 to 1300 μM . This is to be expected and likely depends on the level of expression of a variety of MDRs in these organisms. In contrast to berberine, the MIC of the hybrid was essentially the same in all strains tested, 3–6 μM (note that ≤ 2 -fold differences are considered

insignificant in the MIC test). *Enterococcus faecalis*, which is known for its high levels of “intrinsic antibiotic resistance” (30, 31), was especially resistant to berberine (MIC 650–1300 μM) but susceptible to the hybrid. The hybrid was also more active than an equimolar combination of berberine and INF₅₅ (Table 1). As expected, berberine in the presence of INF₅₅ was more effective against a strain lacking the NorA pump as compared with wild-type and overexpressing mutants. This is apparently due to the presence of additional MDRs in *S. aureus* and is consistent with previous findings from our group (18) and from other authors (32). By contrast, the activity of the hybrid was not affected by the presence or absence of NorA, apparently indicating that penetration of this molecule was largely unaffected by MDRs.

Similarly to the combination of INF₅₅ and berberine, the hybrid had limited activity against Gram-negative species. The MICs were 192 μM against *Pseudomonas aeruginosa*, 96 μM against *E. coli*, and 24 μM against *Salmonella typhimurium*. This is to be expected, since Gram-negative species possess RND MDRs, which are insensitive to INF₅₅. Apparently, the hybrid is unable to bypass RND MDRs as well.

A more detailed examination of growth inhibition of *S. aureus* by the hybrid was made (Figure 2). Measuring inhibition of *S. aureus* growth as a function of concentration shows rapidly increasing inhibition of growth with the hybrid at ~ 100 -fold lower concentration than berberine. A combination of berberine and INF₅₅ was less effective than the hybrid. At higher concentrations, INF₅₅ showed the paradoxical effect of decreased growth inhibi-

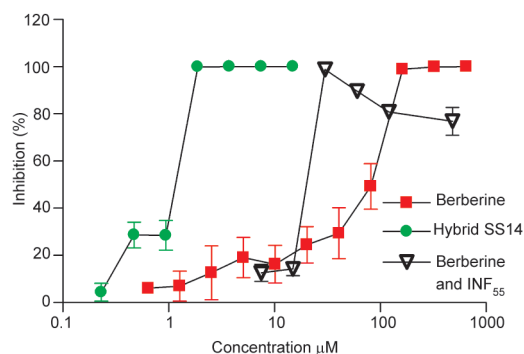


Figure 2. Potentiation of berberine action against *S. aureus* by disabling MDR efflux. Actively growing wild-type *S. aureus* cells were treated with berberine or hybrid SS14. Berberine at a fixed concentration (1.87 μM) was potentiated by varying amounts of INF₅₅.

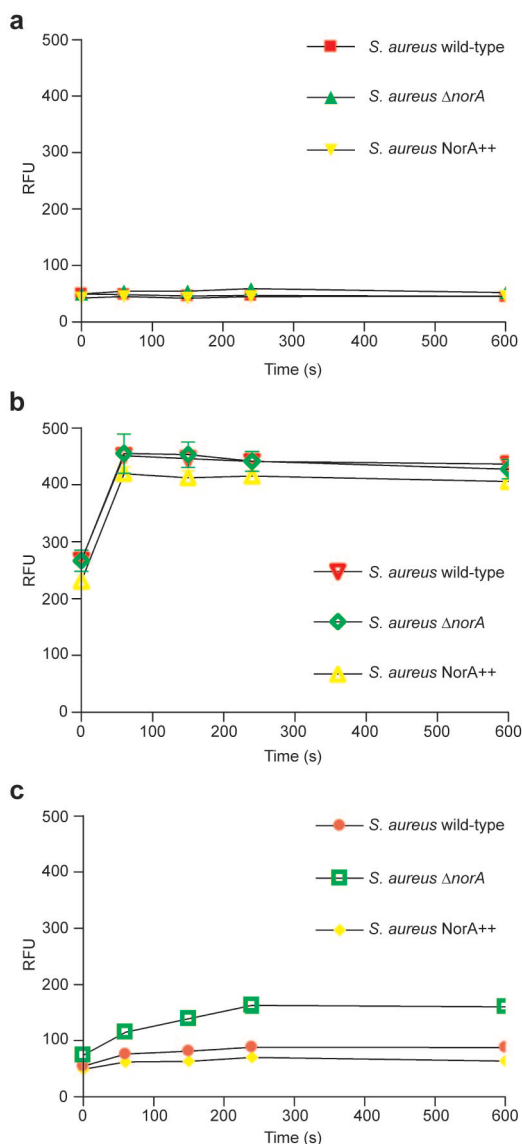


Figure 3. Accumulation of berberine and the hybrid SS14 by *S. aureus*. a) Berberine at 3 μM . b) Hybrid at 3 μM . c) Berberine at 3 μM plus INF_{55} at 3 μM . Accumulation was measured by increase in fluorescence and expressed as relative fluorescence units (RFU).

tion, possibly due to complex formation at high concentrations of INF_{55} and berberine (J. B. Bremner, University of Wollongong, unpublished), which effectively decreases the concentration of free berberine and may decrease its ability to bind DNA, for example. We did not observe a comparable paradoxical effect with SS14. This is perhaps

not surprising, given the steric hindrance between the berberine and INF_{55} moieties and the consequent different conformational preferences of the bulky *o*-substituted 2'-aryl group on the indole moiety, which would not apply to the complex of two free molecules. Specifically, we see good DNA interaction for SS14, comparable to that of berberine, as observed by increased fluorescence.

Hybrid Transport. We had previously monitored increase in fluorescence upon accumulation of berberine in cells (29). The same approach was used to monitor uptake of SS14 (Figure 3). We chose the MIC concentration of SS14, 3 μM , for this experiment. Rapid accumulation of SS14 was observed with all strains tested (ΔnorA , wild-type, and $\text{norA}++$), whereas there was essentially no accumulation of berberine at the same concentration. This is consistent with the dramatically lower antimicrobial activity of berberine compared with the hybrid. Accumulation of berberine was observed in the presence of INF_{55} , although the rate and level of uptake were lower compared with the hybrid, which also agrees with the antimicrobial activity data. To study efflux in *S. aureus*, we loaded cells with the hybrid or berberine and then transferred them into fresh medium. It was not possible to load the cells, however, using the same concentration of SS14 and berberine. When loaded with berberine at 3 μM , the MIC of the hybrid, there was no noticeable efflux of berberine due to the nominal amount of accumulation. The level of berberine had to be increased substantially in the loading phase to 80 μM in order to observe efflux (Figure 4). At this concentration, SS14 caused rather rapid lysis of cells, which is to be expected, since the membrane is an important site of action of hydrophobic cations, and this concentration is 26-fold above its MIC (a comparable level of berberine would have been 8450 μM , above the limit of solubility). We therefore compared the relative changes in efflux using berberine at 80 μM and the hybrid at 3 μM (Figure 4). Hybrid efflux eventually leveled off, as did berberine efflux in the presence of INF_{55} , whereas berberine efflux continued unabated.

Antibacterial Activity *in Vivo*. To test the potential of the berberine MDR inhibitor conjugate in an *in vivo* model of infection, we took advantage of a pathogenesis model that utilizes the well-studied nematode worm *Caenorhabditis elegans* persistently infected with *En. faecalis* (33). We have recently shown that *C. elegans* infected with a variety of human bacterial pathogens that normally kill *C. elegans* can be cured by treatment

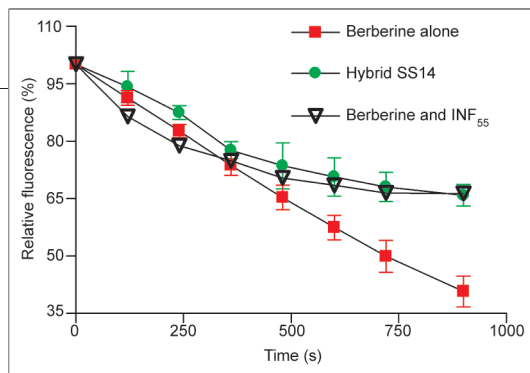


Figure 4. Efflux of berberine and hybrid by *S. aureus* *norA*+. Cells were loaded with berberine at 80 μM and then resuspended in buffer in the absence of inhibitor or in the presence of inhibitor INF₅₅ or loaded with hybrid at 3 μM and resuspended in buffer.

with conventional antibiotics (34). *C. elegans* were allowed to ingest cells of *En. faecalis*, which causes a persistent intestinal infection leading to the death of the nematode (33). Vancomycin is an effective antibiotic used to treat enterococcal infections, and it produced a substantial curative effect (Figure 5). The hybrid also showed good anti-infective activity. By contrast, berberine alone had no effect, consistent with its low activity against *En. faecalis*. Interestingly, co-administering berberine with the INF₅₅ was even worse than the mock-treated control, apparently due to toxicity of INF₅₅. It appears that the hybrid not only improves on the activity of berberine but possibly also cancels the toxicity of the MDR inhibitory moiety. Indeed, INF₅₅ was toxic to human cells at 50 $\mu\text{g mL}^{-1}$ (P. Markham, Grants abc, Clifton, VA, personal communication), while SS14 did not show cytotoxicity at >100 $\mu\text{g mL}^{-1}$ (J. B. Bremner, University of Wollongong, unpublished). In summary, these results indicate that conjugating an antimicrobial

prone to efflux to an MDR inhibitory moiety can produce an antimicrobial with excellent activity.

Conjugating molecules with different functionalities to produce a hybrid antimicrobial has been reported previously (35, 36), but this work is the first example of an anti-infective/MDR inhibitor conjugate. Future work will examine the detailed mechanism by which this compound penetrates into the cell. In this study, we intentionally chose a preferred MDR substrate, a hydrophobic cation. This suggests that other compounds extruded by MDRs can be similarly potentiated by implementation of the design principle of conjugation to a suitable MDR inhibitor.

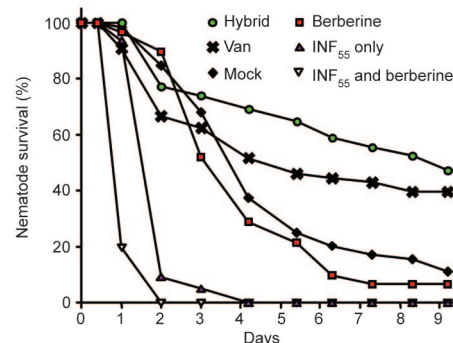


Figure 5. The effect of the hybrid on the survival of nematodes infected with *En. faecalis*. Infected nematodes were mock treated or treated with berberine, INF₅₅ only, INF₅₅ in combination with berberine, vancomycin, or hybrid. In pairwise comparison log rank tests, the difference in survival curves between mock and hybrid treatments was $p < 0.0001$.

METHODS

Bacterial Strains and Culture Conditions. The following bacterial strains were used in this study: *S. aureus* 8325-4 (wild-type); K1758 (8325-4 Δ *norA*) (32); K2361 (K1758/pK364:*norA*, with *norA* from *S. aureus* SA1199B (37)); K2378 (K1758/pK374:*norA* with *norA* from *S. aureus* SA1199 (37)) (G. Kaatz, VA Medical Center, Detroit, MI, personal communication).

S. aureus was grown in Mueller–Hinton broth, *Bacillus cereus* T and 594 were grown in Luria–Bertani broth, and *En. faecalis* V583, MMH594, and OG1RF were grown in brain–heart infusion (BHI) broth. Growth of K1758 was supplemented with erythromycin (20 $\mu\text{g mL}^{-1}$) and that of K2361 and K2378 with chloramphenicol (20 $\mu\text{g mL}^{-1}$).

Chemicals and Antibiotics. Berberine chloride, chloramphenicol, and erythromycin were purchased from Sigma Chemical Co. INF₅₅ and INF₂₇₁ were purchased from ChemBridge. Compound

SS14 was determined to be >95% pure by HPLC analysis (Phenomenex Luna C18 (5 μm) column, 150 mm \times 4.60 mm; UV detection, 254 nm; gradient elution, solvent A = 100% H₂O (0.1% conc HCl) and solvent B = 10% H₂O, 90% CH₃CN (0.1% conc HCl); gradient timing: 0 min 70% solvent A, 30% solvent B; 2 min 70% solvent A, 30% solvent B; 32 min 0% solvent A, 100% solvent B; R_t = 24.7 min) and by ¹H NMR spectroscopy; commercial compounds were also noted as >95% pure.

Antimicrobial Susceptibility. Cells (10^5 mL^{-1}) were inoculated into broth and dispensed at 0.2 mL well⁻¹ in 96-well microtiter plates. MICs were determined in triplicate by serial 2-fold dilution of the test compound. The MIC was defined as the concentration of an antimicrobial that completely inhibited cell growth during an 18 h incubation at 37 °C or 20 h incubation at 30 °C. Growth was assayed with a microtiter plate reader (Spectramax PLUS384; Molecular Devices) by monitoring absorption at 600 nm.

Uptake Assay. *S. aureus* was cultured with aeration at 37 °C to an optical density of 1.5 at 600 nm (OD₆₀₀), pelleted by centrifugation.

gation for 2 min at 12,000 RPM, and washed twice with 25 mM phosphate-buffered saline (PBS). Cells were then resuspended to an OD₆₀₀ of 0.8 in PBS buffer containing 10 mM dextrose and incubated for 1 h at 37 °C with aeration. Cells were washed twice by centrifugation for 2 min at 12,000 RPM with PBS containing dextrose. Cells were then resuspended to an OD₆₀₀ of 0.3 in PBS. Assays were performed in 96-well flat-bottom white microtiter plates (NUNC) in a final volume of 200 μL. Hybrid was added at 3 μM, berberine at 3 μM, and INF₅₅, when present, at 3 μM. Fluorescence was measured with a Spectramax GeminiXS spectrofluorometer (Molecular Devices) at a 355-nm excitation wavelength and a 517-nm emission wavelength. Both SS14 and berberine showed low and equivalent background fluorescence, which was blanked, in the absence of cells.

Efflux Assay. *S. aureus* norA++ was cultured with aeration at 37 °C to an OD₆₀₀ between 0.9 and 1, pelleted by centrifugation for 2 min at 12,000 RPM, and then washed and resuspended in 25 mM PBS (pH 7.3) containing 0.05 g L⁻¹ MgSO₄, 7 g L⁻¹ K₂HPO₄, 0.5 g L⁻¹ sodium citrate·3H₂O, 1 g L⁻¹ (NH₄)₂SO₄, 0.01 mg L⁻¹ folic acid, 0.05 mg L⁻¹ pyridoxine hydrochloride, 0.025 mg L⁻¹ riboflavin, 0.01 mg L⁻¹ biotin, 0.025 mg L⁻¹ thiamine, 0.025 mg L⁻¹ nicotinic acid, 0.025 mg L⁻¹ calcium pantothenate, 0.5 μg L⁻¹ vitamin B12, 0.025 mg L⁻¹ *p*-aminobenzoic acid, 0.025 mg L⁻¹ thiotic acid, and 4.5 mg L⁻¹ monopotassium phosphate. Cells were then resuspended to an OD₆₀₀ of 0.8 in buffer with 10 mM dextrose. Cells were then loaded with either 80 μM berberine and 30 μg mL⁻¹ reserpine or 3 μM SS14 and incubated at 37 °C with aeration for 20 min. Cells were then centrifuged for 2 min at 12,000 RPM in a 4 °C cold room, washed in ice-cold PBS, and added at an OD₆₀₀ of 0.3 to a chilled 96-well flat-bottom white microtiter plate (NUNC) containing ice-cold 25 mM PBS and 10 mM dextrose in a final volume of 200 μL. Some wells contained INF₅₅ at 5 μg mL⁻¹. Fluorescence was measured with a Spectramax GeminiXS spectrofluorometer (Molecular Devices) at a 355-nm excitation wavelength and a 517-nm emission wavelength.

Nematode Curing Assay. *C. elegans* strain *glp-4(bn2ts);sek-1(km4)*, which is a temperature-sensitive sterile mutant that has enhanced susceptibility to pathogens (38), was synchronized by isolating eggs and hatching them overnight in M9 buffer. The L1 stage nematodes were grown on *E. coli* strain OP50 on NGM agar media at 25 °C to the young adult stage. Nematodes were suspended in M9 buffer and transferred onto a lawn of *En. faecalis* strain OG1RF on BHI agar containing 80 μg mL⁻¹ kanamycin to inhibit *E. coli* growth. Nematodes were infected for 9 h at 25 °C and resuspended with M9 buffer, and then 30–40 worms were pipetted onto 35 mm Petri plates containing 3 mL of BHI agar plus the appropriate compounds: the test compounds were incorporated into BHI media at 75 μM and vancomycin at 25 μg mL⁻¹. The BHI media for all of the samples included 1% DMSO, 3% ethanol, and 80 μg mL⁻¹ kanamycin. Worm survival was monitored, and worms were considered to be dead when they were unresponsive to touch. Statistical analysis was performed according to the method of Kaplan–Meier.

Synthesis of Compound 3 and SS14. For general experimental details, see ref 29.

2-(2-Bromomethyl-phenyl)-5-nitro-1H-indole (3). A suspension of the alcohol 2 (29) (200 mg, 0.75 mmol), triphenylphosphine (390 mg, 1.5 mmol), and carbon tetrabromide (490 mg, 1.5 mmol) in dry diethyl ether (60 mL) was stirred with warming at 40 °C under a nitrogen atmosphere for 2 d. The reaction mixture was then filtered, and the filtrate was concentrated. The residual yellow oil was chromatographed on silica gel by vacuum liquid chromatography (silica gel, 20% EtOAc in petroleum spirit, bp 40–60 °C) to give 3 (102.3 mg, 41%) as a yellow solid, mp 164–166 °C. ¹H NMR (300 MHz, CDCl₃): δ 4.64 (s, 2H, CH₂Br), 6.93 (d, *J* = 1.2 Hz, 1H, H-3), 7.43–7.58 (m, 5H, H-7, aro-

matic), 8.16 (dd, *J* = 9.0, 2.1 Hz, H-6), 8.64 (d, *J* = 2.1 Hz, 1H, H-4), 9.14 (s, 1H, NH). ¹³C NMR (75 MHz, CDCl₃): δ 33.0 (CH₂Br), 105.1 (C3), 111.1 (C7), 117.8 (C4), 118.6 (C6), 127.9 (C3a), 129.4 (aromatic), 129.6 (aromatic), 130.5 (aromatic), 131.7 (aromatic), 131.8 (C2), 135.6 (C1'), 139.3 (C7a)^a, 139.6 (C2')^a, 142.3 (C5). Superscript letters in NMR data indicate interchangeable assignments. High-resolution mass spectrometry (HRMS) (electron ionization (EI)): *m/z* calcd for C₁₅H₁₁N₂O₂⁺ [M]⁺, 330.0003; found, 329.9982.

9,10-Dimethoxy-13-[2-(5-nitro-1H-indol-2-yl)benzyl]-5,6-dihydrobenzo[*g*]-1,3-benzodioxolo[5,6-*a*]quinolinizinium Bromide, SS14. A solution of the dihydroberberine 1 (28) (91.0 mg, 0.24 mmol) and the benzyl bromide 3 (102.3 mg, 0.30 mmol) in dry CH₃CN (7 mL) was heated at reflux for 24 h under a nitrogen atmosphere. The mixture was then concentrated and triturated with diethyl ether. The precipitate was filtered and washed with diethyl ether. The solid was chromatographed on silica gel (6% MeOH in dichloromethane) to give SS14 (55.0 mg, 35%) as a yellow solid; mp 206 °C (dec.). ¹H NMR (300 MHz, CD₃OD): δ 3.03 (t, *J* = 5.5 Hz, 2H, H-5), 4.01 (s, 3H, OCH₃), 4.17 (s, 3H, OCH₃), 4.80 (br.s, 2H, H-6), 4.84 (s, 2H, CH₂Ph), 5.96 (s, 2H, OCH₂O), 6.72 (s, 1H, H3'), 6.86 (s, 1H, H-4)^a, 6.90 (s, 1H, H-14)^a, 6.96 (d, *J* = 7.8 Hz, 1H, H-6'), 7.27 (td, *J* = 7.7, 1.5 Hz, 1H, H-5'), 7.37 (br.t, *J* = 7.5 Hz, 1H, H-4'), 7.42 (d, *J* = 9.0 Hz, 1H, H-7'), 7.58 (dd, *J* = 7.7, 1.1 Hz, 1H, H-3''), 7.78 (d, *J* = 9.3 Hz, 1H, H-11)^b, 7.88 (dd, *J* = 9.0, 2.4 Hz, 1H, H-6'), 7.94 (d, *J* = 9.3 Hz, 1H, H-12)^b, 8.34 (d, *J* = 2.1 Hz, 1H, H-4'), 9.8 (s, 1H, H-8). ¹³C NMR (75 MHz, CD₃OD): δ 29.1 (C5), 36.4 (CH₂Ar), 57.5 (OCH₃), 58.8 (C6), 62.7 (OCH₃), 103.6 (OCH₂O), 105.6 (C3'), 109.3 (C4)^c, 109.8 (C14)^c, 112.3 (C7'), 118.2 (C6'), 118.3 (C4'), 121.4 (C4a), 122.5 (C11)^d, 122.9 (C8a)^e, 127.3 (C12)^d, 128.6 (C4''), 129.2 (C3a'), 130.4 (C6', C5''), 131.8 (C3''), 132.8 (C12a)^e, 133.4 (C2'), 135.0 (C13b), 135.1 (C13), 138.6 (C1'')^f, 139.0 (C13a), 141.2 (C7a'), 141.5 (C2'')^f, 142.9 (C5'), 146.1 (C8), 146.2 (C9)^g, 148.6 (C3a)^h, 151.4 (C14a)^h, 151.7 (C10)^g. HRMS (electrospray): *m/z* calcd for C₃₅H₂₈N₃O₆ [M]⁺, 586.1978; found, 586.1984.

Acknowledgments: This work was supported by Grant AIO59483-01 from the National Institutes of Health to K. Lewis, by the University of Wollongong, Australia, to J. B. Bremner, and by Srinakharinwirot University, Thailand, to S. Samosorn. We thank G. Kaatz for kindly providing *S. aureus* strains.

REFERENCES

- Lewis, K., and Lomovskaya, O. (2002) in *Bacterial Resistance to Antimicrobials: Mechanisms, Genetics, Medical Practice and Public Health* (Lewis, K., Salyers A., Taber H., and Wax, R., Eds.) pp 6190, Marcel Dekker, New York.
- Levy, S. B. (2002) The 2000 Garrod lecture. Factors impacting on the problem of antibiotic resistance, *J. Antimicrob. Chemother.* 49, 25–30.
- Li, X. Z., and Nikaido, H. (2004) Efflux-mediated drug resistance in bacteria, *Drugs* 64, 159–204.
- Poole, K. (2005) Efflux-mediated antimicrobial resistance, *J. Antimicrob. Chemother.* 56, 20–51.
- Saier, M. H., Jr., and Paulsen, I. T. (2001) Phylogeny of multidrug transporters, *Semin. Cell Dev. Biol.* 12, 205–213.
- Chen, C. J., Chin, J. E., Ueda, K., Clark, D. P., Pastan, I., Gottesman, M. M., and Roninson, I. B. (1986) Internal duplication and homology with bacterial transport proteins in the *mdr1* (P-glycoprotein) gene from multidrug-resistant human cells, *Cell* 47, 381–389.
- Thiebaut, F., Tsuruo, T., Hamada, H., Gottesman, M. M., Pastan, I., and Willingham, M. C. (1989) Immunohistochemical localization in normal tissues of different epitopes in the multidrug transport protein P170: Evidence for localization in brain capillaries and cross-reactivity of one antibody with a muscle protein, *J. Histochem. Cytochem.* 37, 159–164.

8. Schinkel, A. H., Smit, J. J. M., van Tellingen, O., Beijnen, J. H., Wagenaar, E., van Deemter, L., Mol, C. A. A. M., van der Valk, M. A., Robanus-Maandag, E. C., te Riele, H. P. J., Bems, A. J. M., and Borst, P. (1994) Disruption of the mouse *mdr1a* P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs, *Cell* **77**, 491–502.
9. Loscher, W., and Potschka, H. (2005) Drug resistance in brain diseases and the role of drug efflux transporters, *Nat. Rev. Neurosci.* **6**, 591–602.
10. Gottesman, M. M., Fojo, T., and Bates, S. E. (2002) Multidrug resistance in cancer: role of ATP-dependent transporters, *Nat. Rev. Cancer* **2**, 48–58.
11. Konings, W. N., and Poelarends, G. J. (2002) Bacterial multidrug resistance mediated by a homologue of the human multidrug transporter P-glycoprotein, *IUBMB Life* **53**, 213–218.
12. Zgurskaya, H. I., and Nikaido, H. (2000) Multidrug resistance mechanisms: drug efflux across two membranes, *Mol. Microbiol.* **37**, 219–225.
13. Markham, P. N., and Neyfakh, A. A. (2001) Efflux-mediated drug resistance in Gram-positive bacteria, *Curr. Opin. Microbiol.* **4**, 509–514.
14. Paulsen, I. T., Brown, M. H., and Skurray, R. A. (1998) Characterization of the earliest known *Staphylococcus aureus* plasmid encoding a multidrug efflux system, *J. Bacteriol.* **180**, 3477–3479.
15. Gros, P., Talbot, F., Tang-Wai, D., Bibi, E., and Kaback, H. R. (1992) Lipophilic cations: a group of model substrates for the multidrug-resistance transporter, *Biochemistry* **31**, 1992–1998.
16. Lewis, K. (2001) In search of natural substrates and inhibitors of MDR pumps, *J. Mol. Microbiol. Biotechnol.* **3**, 247–254.
17. Yu, E. W., McDermott, G., Zgurskaya, H. I., Nikaido, H., and Koshland, D. E., Jr. (2003) Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump, *Science* **300**, 976–980.
18. Hsieh, P. C., Siegel, S. A., Rogers, B., Davis, D., and Lewis, K. (1998) Bacteria lacking a multidrug pump: A sensitive tool for drug discovery, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6602–6606.
19. Severina, I. I., Muntyan, M. S., Lewis, K., and Skulachev, V. P. (2001) Transfer of cationic antibacterial agents berberine, palmatine and benzalkonium through bimolecular planar phospholipid film and *Staphylococcus aureus* membrane, *IUBMB Life Sci.* **52**, 321–324.
20. Lin, J., Tiller, J. C., Lee, S. B., Lewis, K., and Klibanov, A. M. (2002) Insights into bactericidal action of surface-attached poly(vinyl-N-hexylpyridinium) chains, *Biotechnol. Lett.* **24**, 801–805.
21. Lewis, K., and Klibanov, A. M. (2005) Surpassing nature: rational design of sterile-surface materials, *Trends Biotechnol.* **23**, 343–348.
22. Amin, A. H., Subbaiah, T. V., and Abbasi, K. M. (1969) Berberine sulfate: antimicrobial activity, bioassay, and mode of action, *Can. J. Microbiol.* **15**, 1067–1076.
23. Davidson, M. W., Lopp, I., Alexander, S., and Wilson, W. D. (1977) The interaction of plant alkaloids with DNA. II. Berberinium chloride, *Nucleic Acids Res.* **4**, 2697–2712.
24. Stermitz, F. R., Lorenz, P., Tawara, J. N., Zenewicz, L., and Lewis, K. (2000) Synergy in a medicinal plant: antimicrobial action of berberine potentiated by 5'-methoxyhydronecarpin, a multidrug pump inhibitor, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1433–1437.
25. Markham, P. N., Westhaus, E., Klyachko, K., Johnson, M. E., and Neyfakh, A. A. (1999) Multiple novel inhibitors of the NorA multidrug transporter of *Staphylococcus aureus*, *Antimicrob. Agents Chemother.* **43**, 2404–2408.
26. Lomovskaya, O., Warren, M. S., Lee, A., Galazzo, J., Fronko, R., Lee, M., Blais, J., Cho, D., Chamberland, S., Renau, T., Leger, R., Hecker, S., Watkins, W., Hoshino, K., Ishida, H., Lee, V. J. (2001) Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: Novel agents for combination therapy, *Antimicrob. Agents Chemother.* **45**, 105–116.
27. Lomovskaya, O., and Bostian, K. A. (2006) Practical applications and feasibility of efflux pump inhibitors in the clinic—a vision for applied use, *Biochem. Pharmacol.* **71**, 910–918.
28. Bremner, J.B., and Samosom, S. (2003) 8-Allyldihydroberberine as an alternative precursor for the synthesis of 13-substituted berberine derivatives, *Austr. J. Chem.* **56**, 871–873.
29. Samosom, S., Bremner, J. B., Ball, A., and Lewis, K. (2006) Synthesis of functionalized 2-aryl-5-nitro-1 *H*-indoles and their activity as bacterial NorA efflux pump inhibitors, *Bioorg. Med. Chem.* **14**, 857–865.
30. Lynch, C., Courvalin, P., and Nikaido, H. (1997) Active efflux of antimicrobial agents in wild-type strains of enterococci, *Antimicrob. Agents Chemother.* **41**, 869–871.
31. Gilmore, M. S., Sahm, D. F., and Huycke, M. M. (1998) Multi-drug resistant enterococci: The nature of the problem and an agenda for the future, *Emerging Infect. Dis.* **4**, 239–249.
32. Price, C. T., Kaatz, G. W., and Gustafson, J. E. (2002) The multidrug efflux pump NorA is not required for salicylate-induced reduction in drug accumulation by *Staphylococcus aureus*, *Int. J. Antimicrob. Agents* **20**, 206–213.
33. Garsin, D. A., Sifri, C. D., Mylonakis, E., Qin, X., Singh, K. V., Murray, B. E., Calderwood, S. B., and Ausubel, F. M. (2001) A simple model host for identifying Gram-positive virulence factors, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10892–10897.
34. Moy, T. I., Ball, A. R., Anklesaria, Z., Casadei, G., Lewis, K., and Ausubel, F. M. (2006) Identification of novel antimicrobials using a live-animal infection model, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10414–10419.
35. Tegos, G. P., Anbe, M., Yang, C., Demidova, T. N., Satti, M., Mroz, P., Janjua, S., Gad, F., and Hamblin, M. R. (2006) Protease-stable polycationic photosensitizer conjugates between polyethyleneimine and chlorin(e6) for broad-spectrum antimicrobial photoinactivation, *Antimicrob. Agents Chemother.* **50**, 1402–1410.
36. Merrifield, R. B., Juvvadi, P., Andreu, D., Ubach, J., Boman, A., Boman, H. G. (1995) Retro and retroantio analogs of cecropin-melittin hybrids, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3449–3453.
37. Kaatz, G. W., and Seo, S. M. (1997) Mechanisms of fluoroquinolone resistance in genetically related strains of *Staphylococcus aureus*, *Antimicrob. Agents Chemother.* **41**, 2733–2737.
38. Kim, D. H., Feinbaum, R., Alloing, G., Emerson, F. E., Garsin, D. A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M. W., and Ausubel, F. M. (2002) A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity, *Science* **297**, 623–626.

ACS Chem. Biol. 2006, 1, 481–484

Correction: Symbiosis: Chemical Biology at Wisconsin

Laura L. Kiessling*, Sally Garbo Wedde, and Ronald T. Raines

Because of a production error, the following references were misformatted: 5, 7, 10–17. These errors do not affect the scientific integrity of the article. The electronic version was corrected and reposted to the web on October 20, 2006.

10.1021/cb600417c
Published online October 20, 2006

ACS Chem. Biol. 2006, 1, 487–488

Correction: Graduate Education in Chemical Biology at the University of Michigan

Tonia J. Buchholz, Bruce Palfey, Anna K. Mapp, and Gary D. Glick*

The name of author Tonia J. Buchholz was inadvertently misspelled. This error does not affect the scientific integrity of the article. The electronic version was corrected and reposted to the web on October 20, 2006.

10.1021/cb600416x
Published online October 20, 2006

ACS Chem. Biol. 2006, 1, 515–524

Correction: Utilizing RNA Aptamers To Probe a Physiologically Important Heme-Regulated Cellular Network

Jacquin C. Niles and Michael A. Marletta

Because of a production error, the following references were misformatted: 1–42. These errors do not affect the scientific integrity of the article. The electronic version was corrected and reposted to the web on October 20, 2006.

10.1021/cb6004189
Published online October 20, 2006

ACS Chem. Biol. 2006, 1, 525–533

Correction: Design, Synthesis, and Biological Activity of a Potent Smac Mimetic That Sensitizes Cancer Cells to Apoptosis by Antagonizing IAPs

Kerry Zobel, Lan Wang, Eugene Varfolomeev, Matthew C. Franklin, Linda O. Elliott, Heidi J. A. Wallweber, David C. Okawa, John A. Flygare, Domagoj Vucic, Wayne J. Fairbrother, and Kurt Deshayes

Because of a production error, the following references were misformatted: 1–39, 41–44. These errors do not affect the scientific integrity of the article. The electronic version was corrected and reposted to the web on October 20, 2006.

10.1021/cb6004195
Published online October 20, 2006

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stiltz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

Congratulations, You're a Winner Too!

As the Nobel Prize winners head to Stockholm, we have to face the reality that most of us will never receive a call from the Nobel Foundation in the wee hours of the morning. Don't despair: there's always hope that each of our small contributions to science, whether literary or experimental, will be recognized by another, equally visible but less lucrative award, the Ig Nobel Prize (<http://improbable.com/ig.html>).

The Ig Nobel Prize is organized by the magazine *Annals of Improbable Research*. "The winners have all done things that first make people LAUGH, then make them THINK." These Ig Nobel Prizes will not make anyone rich, but the winners are proud of their imaginative and sometimes unusual ways of increasing people's interest in science and technology. These awards are presented by real Nobel laureates, and in 2006 it all happened at the 16th First Annual Ig Nobel Prize Ceremony at Harvard's Sanders Theater. Some of the 2006 winners are presented below, with comments from the editor.

We'll begin with the Ornithology prize, because it is the only 2006 Ig Nobel Prize that is remotely related to chemical biology. Ivan R. Schwab and the late Philip R. A. May received the prize for exploring and explaining why woodpeckers don't get headaches. Schwab claims that the "woodpeckers enjoy a cushioned choroid with an as yet unknown mucopolysaccharide filling the interstices. The pecten probably also has a role in maintaining an effective cushion as the pecten can fill with blood to briefly elevate intraocular pressure thus maintaining firm pressure on the lens and retina to prevent damage" (1). We will leave it to the chemical biologists to determine the nature of the mucopolysaccharide in woodpecker heads and to the physicians to develop a mechanism to deliver it into our craniums for those moments when we feel like we are banging our heads against a wall.

For the cheeseheads out there, know that the Ig Nobels have honored (disonored?) your favorite food with the Ig Nobel Prizes in Chemistry and in Biology. Antonio Mulet, José Javier Benedito, José Bon, and Carmen Rosselló received the Chemistry award for their study "Ultrasonic velocity in cheddar cheese as affected by temperature" (2). Bart Knols and Ruurd de Jong received the Biology prize for showing that the female malaria mosquito *Anopheles gambiae* is attracted equally to the smell of Limburger cheese and to the smell of human feet (3, 4). Chemical biologists interested in receptor function should find this field ripe for further investigations at the molecular level.

For those wishing to disperse the packs of undergraduates outside their office door and get some peace to write their grants, consider using the Mosquito (available through Merthyr Tydfil, Compound Security Systems, www.compoundsecurity.co.uk). Howard Stapleton received the Ig Nobel Peace Prize for inventing an electromechanical teenager-repellent, the Mosquito, a device that makes a high-pitched noise audible to teenagers but not to adults. He invented the device, which emits a 17-kHz high-frequency pulsing sound that can be heard by most people younger than 20 and almost no one over 30, to disperse teenagers loitering outside of storefronts. The device works very well, but how? As we age, we lose the ability to hear higher-frequency sounds. Not to be outsmarted by the adults, teenagers are using that same technology for telephone ringtones that are audible to them but

not to their teachers and parents. For their contributions, teenagers everywhere shared the prize with Stapleton.

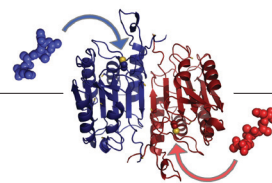
When we find a moment of peace to write that grant or paper, let's choose our words carefully. Daniel Oppenheimer, the winner of the Ig Nobel Prize in Literature for his report "Consequences of erudite vernacular utilized irrespective of necessity: problems with using long words needlessly", suggests we use simpler words (5). "It turns out that somewhere between two-thirds and three-quarters of people admit to replacing short words with longer words in their writing in an attempt to sound smarter," he said. "The problem is that this strategy backfires—such writing is reliably judged to come from less intelligent authors." Let's remember that the next time we're preparing a manuscript, a grant, or an Ig Nobel acceptance speech.



Evelyn Jabri
Executive Editor

REFERENCES

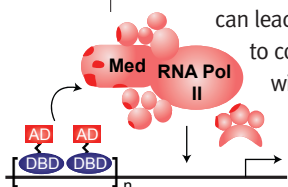
1. Schwab, I. R. (2002) Cure for headache, *Br. J. Ophthalmol.* 86, 843.
2. Mulet, A., Benedito, J. J., Bon, J., and Rosselló, C. (1999) Ultrasonic velocity in cheddar cheese as affected by temperature, *J. Food Sci.* 64, 1038–1041.
3. Knols, B. G. J. (1996) On human odour, malaria mosquitoes, and Limburger cheese, *Lancet* 348, 1322.
4. Knols, B. G. J., and De Jong, R. (1996) Limburger cheese as an attractant for the malaria mosquito *Anopheles gambiae* s.s., *Parasitol. Today* 12 (4), 159–161.
5. Oppenheimer, D. M. (2006) Consequences of erudite vernacular utilized irrespective of necessity: problems with using long words needlessly, *Appl. Cognit. Psychol.* 20, 139–156.



Transcriptional Activation, Artificially

There are many pathways through which transcriptional activation can go awry, which can lead to disease. The ability to control gene transcription with artificial transcriptional activators is a compelling strategy to help scientists understand the transcription process and to manipulate it. Lum *et al.* (p 639) present the design and evaluation of a new class of artificial activators that incorporate a masking interaction, which profoundly enhances their activity.

Starting with several eight-residue peptides that bind to the transcriptional machinery but have no transcriptional activation activity, fusion proteins with the Gal4(1–100) DNA binding domain were generated. Gal4(1–100) was selected to serve as an intramolecular binding surface to mask unproductive interactions in which the hydrophobic peptides might otherwise be tempted to participate. When presented as Gal4(1–100) fusion proteins, several of the peptides exhibited increased activity, a phenomenon that correlated with peptide hydrophobicity. Fluorescence polarization and mutagenesis were used to characterize the interactions between the artificial activators and Gal4(1–100). These studies yield clues for the design of future generations of small-molecule transcriptional activators that could exhibit enhanced cellular activity.



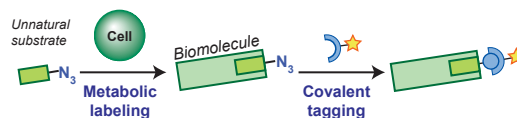
S-Nitrosation Stops Suicide

S-Nitrosation, a post-translational modification in which a cysteine thiol is converted to a nitrosothiol, is another mechanism that cells use to regulate signaling events. S-Nitrosation of the proapoptotic protease caspase-3 inhibits its activity and prevents apoptosis, but the mechanisms surrounding this process are a mystery. Mitchell *et al.* (p 659 and Point of View p 615) now report the synthesis and evaluation of a small-molecule transnitrosating agent designed to help elucidate the role of caspase-3 S-nitrosation.

The structure of the peptide-derived transnitrosation agent, termed PepSNO, is based on the recognition sequence of caspase-3 while incorporating a nitrosated thiol. A variety of methods, including a fluorometric activity assay, the biotin switch method, competitive inhibition experiments, mass spectrometry, and flow cytometry, were used to determine that when PepSNO is exposed to caspase-3, the nitric oxide group is specifically transferred to the active site cysteine and the enzyme is deactivated. Notably, treatment of Jurkat cells with PepSNO inhibited etoposide-stimulated apoptosis. Comparison of PepSNO to nitrosated glutathione, another transnitrosating agent, revealed that PepSNO was more selective for caspase-3 and more effective at preventing apoptosis. This study provides a starting point for the development of other caspase-selective transnitrosating agents.

Azide: The Chemical Reporter

Selective incorporation of an azide into a target biomolecule followed by



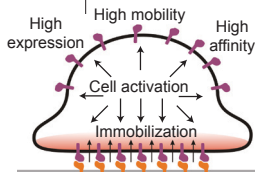
selective reaction of the azide with a detection reagent provides a compelling method for the detection of proteins, lipids, glycans, and other metabolites. Three bio-orthogonal chemical strategies have been used to tag azide-containing biomolecules: the Staudinger ligation, the Cu(I)-catalyzed azide–alkyne cycloaddition (“click chemistry”), and the strain-promoted [3 + 2] cycloaddition. However, it is not always obvious which of these three reactions is the best choice. Now, Agard *et al.* (p 644) compare the reactions in the labeling of purified proteins, protein mixtures, and live cell surfaces.

In experiments with purified proteins and protein mixtures, click chemistry was the clear winner, affording time- and concentration-dependent reactions with the highest sensitivity. In contrast, the different reaction environments of live cells called for different reaction conditions, because the reagents used in click chemistry proved toxic. Both the Staudinger ligation and the strain-promoted cycloaddition were effective reactions for labeling cell surface azide-containing glycoproteins. Notably, the structure of the azide predictably affected the reaction efficiency, providing clues for choosing the most suitable reaction in a given situation.

Tracking T Cell Adhesion

T cells are critical components of the immune system, patrolling the environment for intruders. The interaction between CD2, a receptor on the T cell surface, and CD58, a receptor on the antigen presenting cell (APC), modulates T cell–APC adhesion and facilitates the T cell’s ability to determine whether an immune response is needed. Zhu *et al.* (p 649) use several methods to characterize the molecular mechanisms that regulate this interaction.

Quantitative fluorescence imaging, fluorescence photobleaching recovery, and single particle tracking were used to examine live T cell adhesion to CD58. It was observed that T cell activation results in an increased density and rate of accumulation of CD58 at the contact area, an increase in the cell surface expression of CD2, a reduction of the lateral mobilities of both CD2 and CD58, and an enhancement of the affinity of the CD2–CD58 interaction. In addition, evidence implicating the cytoskeleton in selectively immobilizing CD2 at sites of CD58 contact was presented. The authors articulate a comprehensive model by which the CD2–CD58 interaction is regulated that incorporates the combined effects of coordinated changes in cell activation and receptor conformation, mobility, expression level, and surface density.



Introducing our AUTHORS



Jenifer Lum

Current position: University of Michigan, Ann Arbor, Medicinal Chemistry, Ph.D. candidate with Prof. Anna Mapp

Education: University of California, Los Angeles, B.S. in chemical engineering, 2001

Nonscientific interests: Basketball, running, mountain biking

Transcriptional activators are responsible for the recruitment of coactivators and RNA polymerase II to gene promoters. My research is focused on understanding the role of individual protein-protein contacts on transcriptional activator function by using short peptides as probes to disrupt or promote transcriptional activation. We have converted inactive Mediator-targeted peptide activation domains into robust activators by inclusion of a hydrophobic binding surface. Both interactions are necessary for activity and provide evidence that features beyond coactivator contact are necessary for natural activator function and future artificial activator design. (Read Lum's article on p 639.)



Douglas Mitchell

Current position: University of California, San Diego, Departments of Pharmacology, Cellular and Molecular Medicine, and Chemistry and Biochemistry, postdoctoral associate with Prof. Jack E. Dixon

Education: Carnegie Mellon University, B.S. in chemistry, 2002; University of California, Berkeley, Ph.D. in chemistry with Prof. Michael A. Marletta, 2006

Nonscientific interests: Enjoying life by whatever means necessary

Although it is well documented that signaling concentrations of NO inhibit apoptosis and that S-nitrosation of the caspase proteases plays an important role in apoptotic inhibition, the precise mechanisms and targets of S-nitrosation have yet to be fully elucidated. We have engineered a small molecule that is capable of selective transnitrosation with the active site of caspase-3/7. Expansion of our method to encompass other caspases provides an avenue for addressing specific questions about when, where, and how NO inhibits apoptosis. My future research interests include the design, synthesis, and characterization of new reagents to facilitate the study of signal transduction. (Read Mitchell's article on p 659 and Point of View on p 615.)



Nicholas Agard

Current position: University of California, Berkeley, Department of Chemistry, Chemical Biology Program, Ph.D. candidate with Prof. Carolyn R. Bertozzi

Education: Brown University, B.S. in chemistry, 2002

Nonscientific interests: Hiking and camping, wine tasting in the Napa and Sonoma valleys, and attending football games

In this work, we've compared strain-promoted chemistry and improvements to it with previously existing chemistries for modification of unnatural substrates. Use of these chemistries has been fundamental in exploring classes of biomolecules, particularly post-translationally modified proteins, by enabling their purification and imaging in complex mixtures. This paper provides a framework from which to choose an appropriate chemistry for tagging proteins modified by unnatural substrates and thus should serve as a guide to biologists. (Read Agard's article on p 644.)



Jeremy Baskin

Current position: University of California, Berkeley, Department of Chemistry, Ph.D. candidate with Prof. Carolyn R. Bertozzi

Education: Massachusetts Institute of Technology, S.B. in chemistry, 2004

Nonscientific interests: Classical music and playing piano

My research focuses on developing bioorthogonal chemistries for labeling biomolecules in living systems. The challenge here is taking the concepts of traditional organic synthesis and transplanting them from a round-bottom flask to a dish of cells or even a living animal. This paper discusses bioorthogonal chemistries for detecting the azide, a chemical reporter used to study aspects of cellular biochemistry, with a focus on strained alkyne probes. To me, the promise of bioorthogonal chemistries is the ability to visualize the molecular details of any physiological process, time-resolved, inside a living organism. (Read Baskin's article on p 644.)



Christopher W. Cairo

Current position: University of Alberta, Department of Chemistry, assistant professor

Education: State University of New York-Albany, B.Sc. in chemistry, 1996; University of Wisconsin-Madison, Ph.D. in chemistry with Prof. Laura L. Kiessling, 2002

Postdoctoral work: Harvard Medical School with Prof. David E. Golan, 2002-2006

Nonscientific interests: Reading, music, hiking, and herpetology

My research interests focus on using chemistry and biophysics to study the interaction and function of membrane receptors. We used non-ensemble measurements of receptor mobility in conjunction with fluorescence microscopy to study the T cell adhesion receptor CD2. By applying a novel affinity analysis, we were able to integrate measurements of receptor mobility, number, and affinity to understand the mechanisms T cells use to regulate adhesion. We infer that the cell uses these parameters, including location and conformation, to generate context-specific regulation of the receptor, thereby enhancing adhesion. (Read Cairo's article on p 649.)

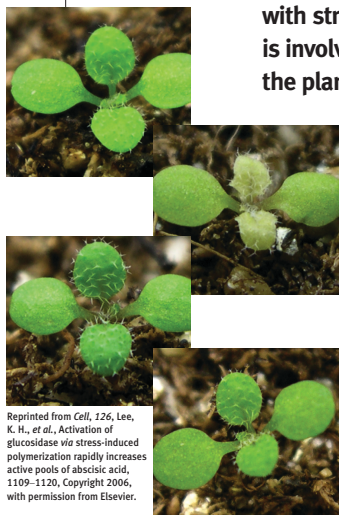
Spotlight

Plants Get Hormonal Too

Apparently, hormones play a role in how plants deal with stress too. The plant hormone abscisic acid (ABA) is involved in various physiological processes during the plant life cycle, including adapting to environmentally stressful conditions such as dehydration. Plants tweak their ABA levels in order to adjust to continually changing conditions, but the molecular mechanisms involved are not well understood. Now, Lee *et al.* (*Cell* 2006, 126, 1109–1120) demonstrate that the β -glucosidase AtBG1 is an important modulator of ABA levels and reveal regulatory mechanisms behind AtBG1 activity.

The observation that stress conditions or exposure to exogenous ABA induces the expression of *AtBG1* led to the discovery that loss of *AtBG1* results in defects in responses mediated by ABA. The use of wild-type and mutant proteins to investigate the activity of AtBG1 indicated that the enzyme specifically hydrolyzes the glucose ester of ABA (ABA-GE) to ABA. The presence of a pep-

tide sequence suspiciously similar to an endoplasmic reticulum (ER) retention signal suggested that AtBG1 resides in the ER. Indeed, ER-localized AtBG1 hydrolyzes ABA-GE, which appears to be imported into the ER by a membrane-localized transporter. Further investigations demonstrated that increased ABA levels in response to dehydration are correlated with AtBG1 levels, an indication that AtBG1 is activated under these conditions. Clues from previous studies suggesting that multimerization of β -glucosidases results in increased activity led to the discovery that dehydration causes polymerization of AtBG1, which results in higher enzymatic activity. The authors also demonstrated that the ABA produced by AtBG1 contributes to both intracellular and extracellular ABA signaling. Taken together, these data suggest that, in addition to *de novo* synthesis, an alternative regulatory mechanism for ABA exists. The activity of AtBG1 may facilitate rapid adjustment of ABA levels, which is required for adaptation to the ever-changing environment in the daily life of a plant. EG



Reprinted from *Cell*, 126, Lee, K. H., *et al.*, Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid, 1109–1120, Copyright 2006, with permission from Elsevier.

Membrane Manipulation

Phosphoinositide phospholipids are important signaling components of the plasma membrane (PM). Hydrolysis of phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) by phospholipase C (PLC) results in the closing of KCNQ channels, a family of potassium channels that regulates neuron excitability and is associated with certain inherited diseases, including epilepsy, cardiac ventricular arrhythmias, and deafness. However, it is not known whether depletion of PtdIns(4,5)P₂ alone is sufficient to close KCNQ channels or whether other signaling events also contribute to this event. Using a chemical dimerizer strategy, Suh *et al.* (*Science Express*, published online Sept 21, 2006, DOI: 10.1126/sci-

ence.1131163) present a method for investigating PtdIns(4,5)P₂ depletion without activating the PLC pathway.

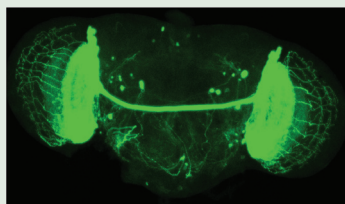
The chemical dimerizer strategy relies on the ability of the small molecule rapamycin to bring together two protein domains, FKBP (FK506 binding protein) and FRB (FKBP-rapamycin binding protein). Fusions of these proteins were generated to create a specific, non-invasive, inducible method for evaluating the cellular consequences of PLC-independent depletion of PtdIns(4,5)P₂. Inp54p, a phosphatase specific for the phosphate at the 5-position of PtdIns(4,5)P₂, was fused to a fluorescent derivative of FKBP (CF-Inp) and transfected into cells along with a membrane-anchored derivative of FRB, called Lyn11-FRB. In addition, a fluorescent pleck-

strin homology domain from PLC δ 1 was created as a PtdIns(4,5)P₂/InsP₃ biosensor. Addition of the rapamycin derivative iRap to cells transfected with these three constructs resulted in rapid translocation of CF-Inp to the PM, *in situ* depletion of PtdIns(4,5)P₂, and concomitant irreversible suppression of KCNQ current. Similar approaches were used to increase the levels of PtdIns(4,5)P₂, which augmented the current, and to induce the synthesis of PtdIns(3,4,5)P₃, which did not affect PtdIns(4,5)P₂ levels or the amplitude of the KCNQ current. These results further define the role of PtdIns(4,5)P₂ in KCNQ channel function and validate this method as a versatile approach for manipulating lipid composition of the PM. EG

An Axon Balancing Act

Axons, the long projections of nerve cells, have complex regulatory mechanisms that enable their function as the transmission lines of the nervous system. Mapping the networks involved in axon regulation will help contribute to our knowledge of brain development and could lead to new therapeutic strategies for nerve regeneration. Srahna *et al.* (*PLoS Biol.* 2006, 4, 2076–2090) use the visual system of the *Drosophila* brain to investigate the signals that regulate axon extension and retraction, putting forth a model describing the network that controls this remarkable process.

A candidate gene approach was employed to decipher the role of specific signaling pathways during axon extension and retraction in dorsal cluster neurons (DCNs), a group of ~40 neurons located in each brain hemisphere. Interestingly, all DCN axons initially extend toward the developing part of the fly optic lobe, termed the medulla, but as development continues only 11 or 12 of the 40 axons continue along defined paths, while the remaining ones retract. The authors found that blocking the Jun N-terminal kinase (JNK) pathway significantly decreased axon extension but that blocking fibroblast growth factor (FGF) receptor activity or the

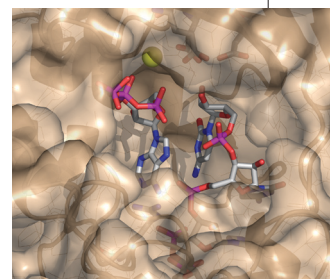


Reprinted from *PLoS Biol.*, 4, Srahna, M., *et al.*, A signaling network for patterning of neuronal connectivity in the *Drosophila* brain, 2076–2090.

ras-related C3 botulinum toxin substrate 1 (Rac1) GTPase promoted DCN axon extension. Further probing revealed that signaling through the Wnt pathway via Wnt5 and the Wnt signaling adaptor protein Dishevelled (Dsh) can attenuate Rac1 activity and thereby suppress Rac1 suppression of JNK. Exploration of how these pathways are intertwined indicated that JNK acts downstream of Rac1, which acts downstream of Dsh. It appears that a careful molecular balancing act between the Wnt and FGF signals governs the number of DCN axons that continue to extend versus the number that retract. Further investigation of the mechanisms that determine the identity of those axons that continue to extend will help connect the complex wiring of neuronal networks. **EG**

CCA Cinema

Transfer RNAs (tRNAs) are linked to their cognate amino acids just after an invariant terminal nucleotide stretch, CCA. This sequence is not encoded by the tRNA genes, but rather, it is appended later by a specialized enzyme, the CCA-adding RNA polymerase. Although such a tailing event appears reminiscent of the eukaryotic messenger RNAs and the poly-A tail, the story with tRNAs is far more complex. This polymerase must specifically recognize the shape of tRNAs, position the 3' terminus, and then sequentially add just three nucleotides onto the end. This phenomenon occurs without any RNA or DNA template guiding the reaction. Thus, a puzzling question has long remained: how does the enzyme achieve its exquisite specificity? Now, Tomita *et al.* (*Nature* 2006, 443, 956–960) add an impressive collection of X-ray crystal structures capturing an archaean CCA-adding enzyme at various points in the reaction pathway.



Reprinted by permission from Macmillan Publishers Ltd: *Nature*, Tomita, K, *et al.*, 443, 956–960, copyright 2006.

The authors determined the structure of the protein in complex with a number of small RNAs corresponding to intermediates along the pathway. In addition, they carefully looked for structural changes that might occur from the incoming nucleoside triphosphate. Using six different high-resolution structures, the authors postulate on the mechanism of the enzyme and even construct a movie that incorporates the X-ray snapshots along the path. The enzyme first stretches part of the tRNA to bring the 3' terminus into the active site. After addition of one C, the RNA snaps back by one nucleotide and repositions the new terminus into the active site. Once another cytidine triphosphate is provided, the enzyme changes conformation to a more closed state for addition of the next C to the tRNA. This open–closed switch is similar to how DNA polymerases clamp down on a template. Unlike traditional polymerases, the next addition of an A occurs in a closed and locked structure that cannot translocate. The enzyme clamps around the RNA helix and prevents further additions. This study displays the dynamic workings of an interesting and highly specialized polymerase found in all three kingdoms of life. **JU**

Making and Breaking Moenomycin

The enzymes involved in bacterial cell wall biosynthesis are excellent targets for antibiotic development because of their critical role in bacterial survival and the lack of analogous enzymes in humans. Transglycosylases catalyze formation of the glycan units of peptidoglycan, the major component of the cell wall. The natural product moenomycin A is the only known natural product inhibitor of transglycosylases, but its poor pharmacokinetic properties and complex structure have prohibited its use as an antibiotic in humans. Two recent papers (Adachi

et al., *J. Am. Chem. Soc.* 2006, 128, 14,012–14,013, and Taylor *et al.*, *J. Am. Chem. Soc.* published online Nov 4, 2006, DOI: 10.1021/ja065907x)

now describe flexible synthetic approaches to moenomycin and its analogues and report the inhibitory activity of moenomycin and a derivative against purified transglycosylases.

Moenomycin A is composed of a pentasaccharide attached to a 2-*O*-moenocinyl glycerate chain *via* a phosphodiester linkage. The total synthesis of moenomycin A, described in the paper by Taylor *et al.*, was designed to be efficient while facilitating generation of moenomycin analogues. The authors constructed the pentasaccharide unit by first synthesizing two disaccharide fragments, linking them together, and then attaching the fifth ring in the last steps. Efficient stereoselective glycosylations were accomplished with the sulfoxide glyco-

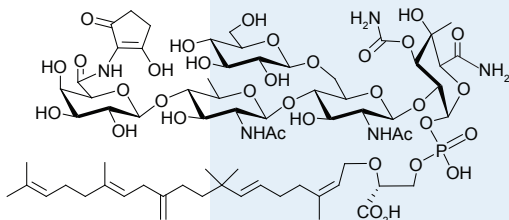
sylation reaction, and the reaction conditions were tweaked depending on the donor–acceptor reactivity profiles. Inverse addition and appropriate use of scavengers proved essential in order to suppress certain side reactions in some of the glycosylations.

The synthesis of the 2-*O*-moenocinyl glycerate piece and the generation of the phosphodiester linkage are described in the paper by Adachi *et al.* 2-*O*-Moenocinyl glycerate was created through conversion of the allyl alcohol functionality in moenocinol to an allyl ether, followed by protecting group shuffling and esterification. Conversion of the anomeric hydroxyl of the pentasaccharide to an H-phosphonate ester followed by reaction with 2-*O*-moenocinyl glycerate in the presence of 1-adamantanecarbonyl chloride, mild oxidation, and global deprotection afforded moenomycin A.

A method for degrading and reconstructing moenomycin was also developed to facilitate manipulation of the reducing end of the compound. This approach enables creation of moenomycin analogues with modified lipid chains without the need to synthesize the compounds from scratch. Successful implementation of this method required developing degradation conditions that left the pentasaccharide unit intact. The inherent lability of the allyl ether functionality that connects the pentasaccharide to the lipid chain enabled the researchers to come up with degradation conditions that were selective for the glycidyl ether linkage while leaving the glycosidic bonds untouched.

With these methods, moenomycin was successfully ripped apart and subsequently put

(continued on page 611)



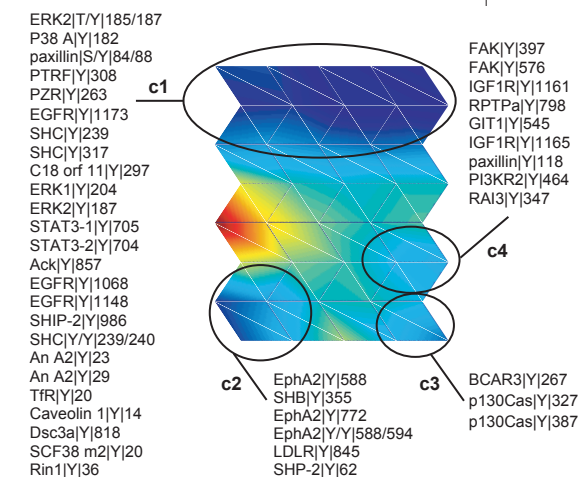
Reprinted with permission from Adachi, M., *et al.*, *J. Am. Chem. Soc.*, 128, 14012–14013. Copyright 2006 American Chemical Society.

Phosphorylation Phenotyping?

The receptor tyrosine kinase human epidermal growth factor receptor 2 (HER2) is a key component of a complex signaling network that regulates important cellular processes such as migration and proliferation. Overexpression of HER2 is notoriously associated with breast and other cancers, and drugs that selectively target HER2 have demonstrated effective anticancer activity in patients. In an effort to map the signaling network of HER2, Wolf-Yadlin *et al.* (*Mol. Syst. Biol.*, published online Oct 3, 2006, DOI: 10.1038/msb4100094) use quantitative mass spectrometry, biological response data, and computational analysis to correlate phosphorylation patterns with cell proliferation or with migration.

The cellular state of tyrosine phosphorylation was examined across 16 dimensions: four time points, two cell lines (one that did and one that did not overexpress HER2), and treatment with

either epidermal growth factor (EGF) or heregulin (HRG), growth factors that differentially stimulate HER2 heterodimers. Astonishingly, 332 phosphorylated peptides from 175 proteins were identified, 122 of which had not previously been described. Using the self-organizing map clustering algorithm, which enables the identification of clusters of tyrosine-phosphorylated peptides with similar temporal dynamics, the authors readily identified four clusters that reveal connectivity in the data. In order to correlate this signaling data with a phenotypic effect, they also measured cell migration and proliferation under the same conditions. In general, HER2 overexpressing cells exhibited enhanced cell migration. Moreover, phosphorylation patterns of cells stimulated with EGF versus HRG pointed to the network connections behind the increased migratory ability of HER2 overexpressing cells and elucidated distinct pathways by which these



Reprinted by permission from Macmillan Publishers Ltd: *Mol. Syst. Biol.*, advance online publication, Oct 3, 2006, DOI: 10.1038/msb4100094.

growth factors promote cell migration. In contrast, HER2 overexpression had a minimal effect on cell proliferation; rather, EGF treatment emerged as the primary driver of cell growth. A model using partial-least-squares regression was constructed to quantitatively correlate phosphorylation patterns with cell migration or proliferation, establishing a powerful approach for exploring the relationship between protein phosphorylation and cellular processes. **EG**

Making and Breaking Moenomycin, *continued from page 610*

back together. In addition, a moenomycin derivative containing a 10-carbon neryl chain in place of the much longer natural polyprenyl unit in moenomycin was also constructed. The compounds were tested for their ability to inhibit purified transglycosylases from two clinically relevant bacterial species, *Staphylococcus aureus* and *Enterococcus faecalis*. Notably, both compounds inhibited the purified proteins comparably, but moenomycin was a much more potent

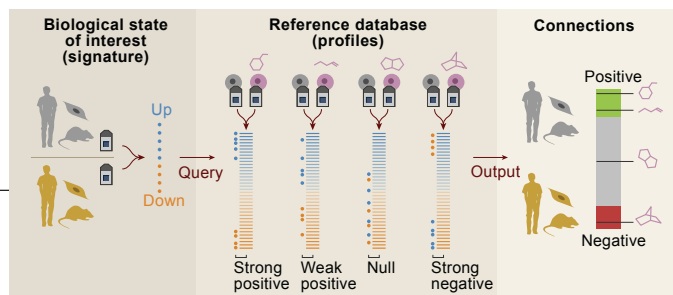
inhibitor of bacterial growth, an indication of the importance of the lipid chain in the context of a biological system.

These studies provide access to new synthetic methods for creating moenomycin analogues, facilitating investigations into the mechanism of inhibition of moenomycin, the biological role of transglycosylases, and the development of moenomycin-based antibiotics. **EG**

Getting Connected!

Establishing connections among physiological and pathological processes and genetic and small-molecule perturbations can lead to unanticipated links that could help decipher the incredibly complex web that defines a biological state. In an attempt to establish a systematic method for exploring these relationships, Lamb *et al.* (*Science* 2006, 313, 1929–1935) present the Connectivity Map, a resource in which gene-expression profiles of cells exposed to small molecules are assembled into a public database for which data-mining tools are available to detect noteworthy relationships among the profiles.

Data from the expression profiles of breast cancer cells exposed to 164 distinct bioactive small molecules were used to create a first-generation Connectivity Map. A query signature, or list of genes whose expression is correlated with a biological state of interest, could then be scanned in the Connectivity Map in the search for prominent relationships. A range of query signatures from both



From Lamb, J., *et al.*, *Science*, 2006, Sept 29, 2006 DOI: 10.1126/science.1132939. Reprinted with permission from AAAS.

internal and external studies were collected and evaluated. The data included the effects on gene expression of small molecules, such as histone deacetylase inhibitors, estrogens, and phenothiazines, and of disease states, such as diet-induced obesity, Alzheimer's disease, and dexamethasone resistance in acute lymphoblastic leukemia. Remarkably, the Connectivity Map revealed both positive and negative connectivity relationships that correctly predicted several known relationships, pointed to the previously unknown mechanism of action of a small molecule, and identified several molecules with potential therapeutic utility. On the basis of these encouraging preliminary results, the authors propose that an expanded Connectivity Map should be generated as a community resource project. Depending on the Map's utility, the exciting prospect of further expansion toward the ultimate goal of creating a comprehensive description of all biological states in the context of genomic signatures could be realized. EG

Linking Lipids to Life

Some pathogens exert their destructive behavior by producing pore-forming toxins, which essentially poke holes through cell membranes and potentially lead to cell death. The molecular processes that govern this pathway, however, are not well understood. Gurcel *et al.* (*Cell* 2006, 126, 1135–1145) explore the cellular response to aerolysin, a pore-forming toxin produced by certain bacteria, and demonstrate evidence for a chain of events that helps explain the cell's ability to repair its membrane and survive.

The authors initially observed that exposure of mammalian cells to aerolysin resulted in activation of the sterol regulatory element binding proteins (SREBPs), transcription factors that regulate chole-

sterol and fatty acid biosynthesis. Further investigations revealed that activation of the SREBPs was caused by loss of potassium through the toxin pores. Interestingly, they noted that potassium efflux had previously been linked to caspase-1 activation, and indeed, caspase-1 was activated in response to aerolysin exposure. It was also known that activation of caspase-1 is dependent on the assembly of large multi-protein complexes called inflammasomes, and they further demonstrated that aerolysin exposure triggers formation of inflammasomes. Moreover, it was demonstrated that prevention of caspase-1 or inflammasome activation blocked aerolysin-induced SREBP-2 activation and that caspase-1 activation induced SREBP activation through

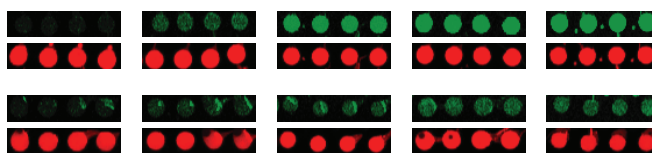
a well-established pathway involving the escort protein SCAP (SREBP cleavage-activating protein) and two transmembrane proteases, S1P and S2P. These data undeniably link caspase-1 and SREBP activation to a common pathway and indicate that caspase-1 activation is upstream of SREBP activation. Finally, they showed that blocking the caspase-1 or SREBP pathways after exposure of primary cells to aerolysin or infection of cells with aerolysin-producing bacteria increases cell death, an indication that activation of these pathways promotes cell survival. Taken together, these results connect intracellular ion levels, caspases, SREBPs, and lipid metabolism as part of the survival mechanism that cells employ to fight pore-forming toxins. EG

Quality of Protein Microarrays

The assembly and deciphering of protein interaction networks promise to reveal valuable information about how organisms function. The accuracy of commonly used methods for accessing protein–protein interaction data suffers from difficulties in normalizing the behavior of proteins that by nature vary widely in their physical properties. Gordus *et al.* (*J. Am. Chem. Soc.* 2006, 128, 13,668–13,669) propose a method for minimizing the effects of variations in concentration, surface density, and activity of proteins used in microarrays.

To maximize the chances of working with structures that behave similarly, the authors chose protein domains, rather than whole proteins, to systematically investigate protein–protein interactions in a microarray format. Seven Src homology 2 domains labeled with a fluorescent tag were printed on a microarray surface. The fraction of surface area covered by each protein was evaluated and was found to vary considerably. Next, they used a labeled phosphopeptide known to interact with five of the domains to evaluate

the amount of active protein on the surface; it was found to vary substantially, to the extent that spot intensity did not accurately reflect interaction affinity. However, when saturation binding curves were obtained and normalized with respect to the amount of active protein on the surface, the data did manifest the correct affinities of the interactions. A strength of microarrays is their ability to control protein concentrations, so obtaining this type of



Reprinted with permission from Gordus, A., and MacBeath, G., *J. Am. Chem. Soc.*, 128, 13668–13669. Copyright 2006 American Chemical Society.

quantitative information should dramatically improve the quality and quantitative integrity of protein microarray data. The authors suggest that because the dependencies of concentration and activity of proteins also affect data obtained in other protein interaction assays, such as yeast two-hybrid systems and affinity purification of protein complexes, more diligent efforts should be made to obtain quantitative information when defining protein interaction networks. **EG**

UPCOMING CONFERENCES

American Society for Cell Biology Annual Meeting

December 9–13, 2006
San Diego, CA

Biophysical Society Annual Meeting

March 3–7, 2007
Baltimore, MD

Metals in Biology, GRC

January 28–February 2, 2007
Ventura, CA

Glycobiology, GRC

March 4–9, 2007
Ventura, CA

MicroRNAs and siRNAs: Biological Functions and Mechanisms

January 28–February 2, 2007
Keystone, CO

2007 ACS Spring National Meeting

March 25–29, 2007
Chicago, IL

Regulation and Specificity of S-Nitrosylation and Denitrosylation

Steven R. Tannenbaum^{†,*} and Forest M. White[†]

[†]Division of Biological Engineering, and ^{*}Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The biological chemistry of endogenously formed nitric oxide (NO) is exceedingly complex, and the effect of NO produced by neighboring cells on target cells is even less well-understood (Figure 1) (1, 2). Signaling from NO can occur in a variety of modes, including activation of guanylate cyclase, interaction with the mitochondrion (3), oxidation and nitrosative stress (4), DNA and protein damage (5), and S-nitrosylation (SNO) of proteins (PSNO) (6–9). The terms nitrosation and nitrosylation are used interchangeably here to denote formation of SNO. The difference in terminology refers only to the mechanism of formation. Thus, cells exposed to NO, either endogenously or exogenously, experience layer upon layer of signaling potential, almost certainly dependent upon NO concentration and location (e.g., subcellular, cell type, and tissue). The systems biology of NO has hardly been explored.

The ready nitrosation of sulfhydryl (SH) groups *via* acidified nitrite or NO/O₂ has been known for many years (10, 11), and the potential for protein regulation has been intuited from biological experiments (12); however, only after the development of the biotin switch assay (13) did the full potential of this reaction begin to unfold. At the time of this writing, hundreds of PSNOs have been characterized *via* the biotin switch assay, and many others have been discovered through more targeted investigations (e.g., 3, 14, 15). The regulatory mechanism

of the PSNO system has not been fully characterized, but it is clear that for NO to act as a signaling molecule, it would not make sense to have indiscriminate nitrosylation of any SH group exposed on a protein. Therefore, there must be a mechanism of selectivity, both in formation of PSNO and in denitrosylation to the original state. The need for specificity would seem to eliminate the possibility that NO, N₂O₃, or even S-nitroso-glutathione (GSNO) would be universal donors for conversion of any SH group of the protein (PSH) to a PSNO. Mitchell *et al.* (16) on p 659 of this issue have now demonstrated the possibility of rigorous control of PSNO *via* specific protein–protein interactions. In an earlier paper (17), they demonstrated that nitrosylated thioredoxin (Trx-SNO) transnitrosates active caspase-3 and, therefore, could be a regulator of caspase-3 activity. In the current paper, they now demonstrate that a specific S-nitrosylated peptide (AcDEhCD) can be a substrate mimic for caspase-3 and also modulate caspase-3 activity. Furthermore, the SNO peptide caused much more rapid deactivation of caspase-3 than GSNO, but less than Trx-SNO or AcDEVD-chloromethylketone, an indication that “the rate-determining step was the transnitrosation reaction, not peptide binding.” These experiments now set the table for further research into the SNO regulatory system. For control, one would have to postulate a cascade for conversion of NO to a thermodynamically specific NO donor

ABSTRACT Signal transduction from nitric oxide (NO) is important in physiological processes such as smooth muscle relaxation and neurotransmission. NO signaling occurs through a variety of mechanisms, including S-nitrosylation (SNO) of sulfhydryl groups in protein. A paper in this issue of the journal dissects the regulatory system involved in SNO of caspase-3 and highlights the need for more direct investigations of the complex biological regulation by NO signaling.

*Corresponding author,
srt@mit.edu.

Published online November 17, 2006

10.1021/cb600439h CCC: \$33.50

© 2006 by American Chemical Society

SNO-based modulation of the signaling network potentially provides another level of cellular control.

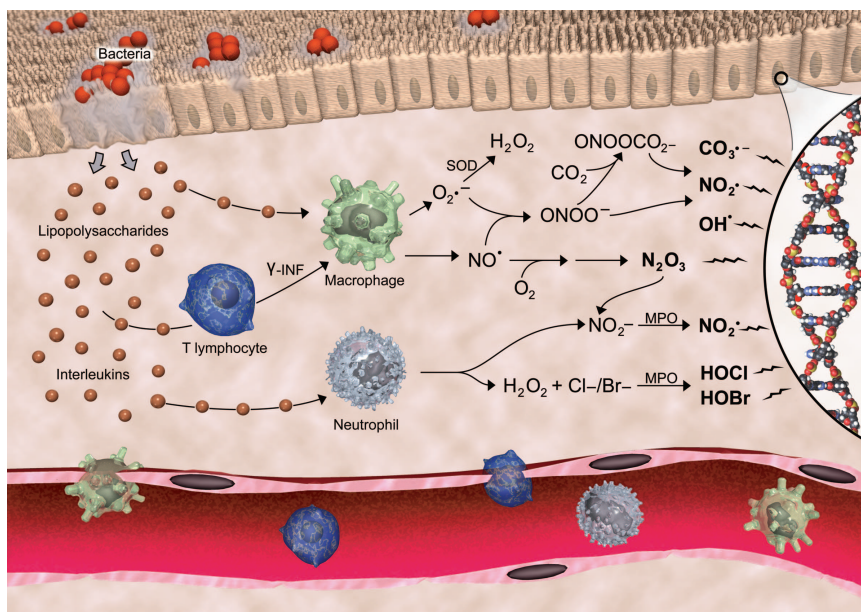


Figure 1. Origins and chemistry of NO from the inducible NOS. INF = γ -interferon, MPO = myeloperoxidase, SOD = superoxide dismutase. Image designed by Jeff Dixon and reprinted with permission from Peter C. Dedon.

and, ultimately, a specific mechanism for protein denitrosation.

Two models are proposed for tight regulation of PSNO (Figure 2). In model i, NO produced by an NO synthase (NOS) interacts with the thiol radical of glutathione (GSH) to form GSNO (1), which several groups have reported at micromolar concentration in cells. GSNO reacts with thioredoxin to give Trx-SNO (18). Trx-SNO forms a complex with a protein (PSH) that then undergoes internal NO transfer, followed by release of the complex. Finally, protein disulfide isomerase (PDI), which contains a thioredoxin

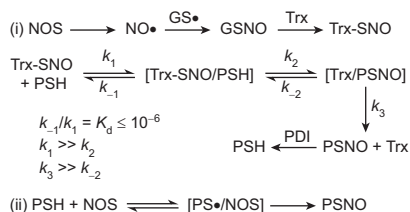


Figure 2. Models for controlled SNO. Details of the models are discussed in the text.

motif, is proposed as the regulatory partner for denitrosylation (15, 19, 20). In model ii, the close association between an NOS and a target protein through a protein–protein interaction is probably facilitated by localization to a membrane or through a scaffold (7, 21). In model i, the very strong interaction energy characteristic of proteins binding to receptors ($K_d < 10^{-6}$) (22) drives the relatively slow transfer of NO between species in the encounter complex (k_2 in Figure 2) (23, 24) in a manner reminiscent of the reaction of electrophiles with proteins (25). Following completion of the reaction, the proteins separate. Denitrosylation follows a similar path involving the third protein (PDI in Figure 2). In model ii, a microregion is implied in the interaction of the partner proteins that allows the NO released by the NOS to be efficiently trapped by the target protein. This mechanism has been postulated for several proteins, including nitrosylation of the *N*-methyl-D-aspartic acid receptor and Dexras1, both of which are found in a complex with neuronal NOS

(26, 27). Because NO is only reactive with other radicals, the target cysteine in the target protein must be a thiyl radical or be capable of radical formation in a concerted reaction.

The Marletta laboratory has focused on caspase-3 as a regulatory target for NO (16, 17). These papers clearly indicate that targeting of SNO is key to regulating biological activity. The implication is that recent efforts to characterize the SNO proteome through treatment of cells, tissues, or proteins with nonspecific nitrosating agents will identify many proteins that are capable of being nitrosylated *in vitro*; however, they may not identify which proteins are actually nitrosylated *in vivo* because of the absence of targeting proteins and subcellular localization in these experiments. Nitrosylated proteins identified in these studies will contain target cysteines that are almost certainly on the surface of the protein, or have a low pK_a , which makes them intrinsically reactive to any electrophile. An equivalent experiment would be chemical phosphorylation of proteins to identify kinase targets. As an alternative to these studies, a current goal should be to identify Trx-SNO target proteins, and PDI- or Trx-PSNO targets for denitrosylation. If model i is correct, it should be possible to chemically tag and purify the [Trx-SNO-PSH] complex and thus identify proteins in the act of being nitrosylated. There will almost certainly be other components of the regulatory system, but this is a reasonable first approach.

It will also be important to validate the role of a PSNO in determining the phenotype of a cell. This has clearly been done for some of the procaspases and caspases, but many other components of the extrinsic and intrinsic apoptotic pathways exist, giving layer upon layer of overlapping signaling networks (28). For instance, in addition to direct regulation of caspase activity, SNO-based modulation of the signaling network potentially provides another level of cellular control, amplifying phosphorylation-based

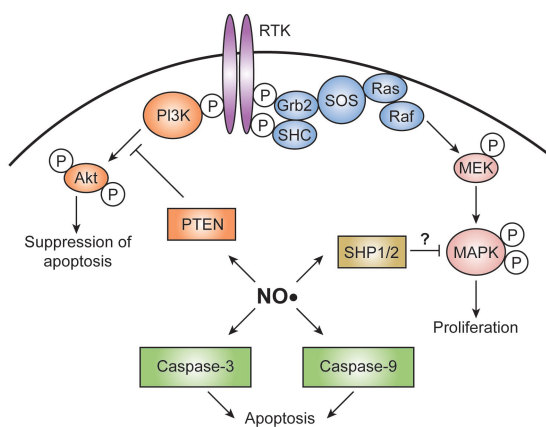


Figure 3. NO may play a central role in regulating phosphorylation-dependent signaling networks in the cell. Grb2 = growth factor receptor-bound protein 2, MEK = MAPK/Erk kinase, PI3K = phosphoinositide 3 kinase, RTK = receptor tyrosine kinase, SHC = Src homologous and collagen-like, SOS = son of sevenless.

signaling networks by modulating kinase and phosphatase activity. Specifically, inhibition of SH2-containing protein tyrosine phosphatases 1 and 2 (SHP1/2) through SNO of the active site cysteine has been demonstrated to lead to increased activation of extracellular signal-regulated protein kinases 1 and 2 (Erk1/2) (29), and NO has been shown to regulate phosphatase and tensin homologue on chromosome 10 (PTEN) activity (30), further enhancing the pro-survival signal (Figure 3). Interestingly, a threshold has been demonstrated for the concentration of NO required to induce apoptosis in cells in culture (31). Therefore, a balance exists between NO-induced processes that send cells into apoptosis or that rescue cells from apoptosis. However, NO is just one piece of the puzzle, and exogenous stimuli can override NO signaling; for example, the ability of insulin to protect cells treated with tumor necrosis factor- α through kinase activation of serine/threonine protein kinase B (Akt) (32).

Many parallels can be drawn between NO and effectors of other signaling pathways, including phosphorylation, ubiquitination,

glycosylation, acetylation, and methylation (7). With the exception of NO, multiple enzymes have been identified for the formation and removal of each of these post-translational modifications. For instance, acetylation is regulated by histone acetyl transferases and histone deacetylases; activity and localization of these proteins have been correlated with protein phosphorylation levels, highlighting the interplay among multiple signaling networks. Interestingly, correlations exist between protein phosphorylation and SNO occurring through either model i or

model ii. In model i, NO transfer is predicated by a protein-protein interaction regulated by thioredoxin domain recognition. Similarly, the interaction of SH2 domains on the substrate with tyrosine phosphorylation sites on the activated kinase leads to a complex that facilitates phosphorylation transfer from kinase to substrate. In model ii, NO transfer is dependent on proximity and co-localization of NOS and protein substrates, akin to the mitogen-activated protein kinase (MAPK) cascade in which specificity and rapid response are enabled by the linkage of kinases and substrates to a common scaffold protein (e.g., kinase suppressor of Ras). Last, in the protein phosphorylation field, recent quantitative analyses of protein tyrosine phosphorylation levels across multiple time points and conditions have highlighted the role of individual phosphorylation sites regulating biological activities such as proliferation and migration (33). Given the tools that have been developed to enable this study, it should now be possible to extend a similar effort to identify and quantify *in vivo* PSNOs

from multiple cell states and, thus, link these proteins to biological phenotype.

Our understanding of the NO signaling field is still in its infancy. However, the Mitchell and Marletta papers (16, 17) describe a potential mechanism for targeted NO transfer and highlight the need for more directed studies to uncover additional components in this signaling network. Only through these efforts will we begin to understand the complexities underlying biological regulation by NO signaling.

Acknowledgments: The research in this article is supported by the National Cancer Institute (CA 26731) and the National Institute of Allergy and Infectious Disease (AI 65354).

REFERENCES

- Lancaster, J. R., Jr. (2006) Nitroxidative, nitrosative, and nitrate stress: kinetic predictions of reactive nitrogen species chemistry under biological conditions, *Chem. Res. Toxicol.* 19, 1160–1174.
- Nalwaya, N., and Deen, W. M. (2005) Nitric oxide, oxygen, and superoxide formation and consumption in macrophage cultures, *Chem. Res. Toxicol.* 18, 486–493.
- Clementi, E., Brown, G. C., Feelisch, M., and Moncada, S. (1998) Persistent inhibition of cell respiration by nitric oxide: crucial role of S-nitrosylation of mitochondrial complex I and protective action of glutathione, *Proc. Natl. Acad. Sci. U.S.A.* 95, 7631–7636.
- Forrester, M. T., Benhar, M., and Stamler, J. S. (2006) Nitrosative stress in the ER: a new role for S-nitrosylation in neurodegenerative diseases, *Chem. Biol.* 1, 355–358.
- Dedon, P. C., and Tannenbaum, S. R. (2004) Reactive nitrogen species in the chemical biology of inflammation, *Arch. Biochem. Biophys.* 423, 12–22.
- Jaffrey, S. R., Erdjument-Bromage, H., Ferris, C. D., Tempst, P., and Snyder, S. H. (2001) Protein S-nitrosylation: a physiological signal for neuronal nitric oxide, *Nat. Cell Biol.* 3, 193–197.
- Hess, D. T., Matsumoto, A., Kim, S. O., Marshall, H. E., and Stamler, J. S. (2005) Protein S-nitrosylation: purview and parameters, *Nat. Rev. Mol. Cell Biol.* 6, 150–166.
- Hao, G., Derakhshan, B., Shi, L., Campagne, F., and Gross, S. S. (2005) SNOSID, a proteomic method for identification of cysteine S-nitrosylation sites in complex protein mixtures, *Proc. Natl. Acad. Sci. U.S.A.* 103, 1012–1017.
- Greco, T. M., Hodara, R., Parastatidis, I., Heijnen, H. F., Dennehy, M. K., Liebler, D. C., and Ischiropoulos, H. (2006) Identification of S-nitrosylation motifs by site-specific mapping of the S-nitrosocysteine proteome in human vascular smooth muscle cells, *Proc. Natl. Acad. Sci. U.S.A.* 103, 7420–7425.

10. Keshive, M., Singh, S., Wishnok, J. S., Tannenbaum, S. R., and Deen, W. M. (1996) Kinetics of S-nitrosation of thiols in nitric oxide solutions, *Chem. Res. Toxicol.* **9**, 988–993.
11. Wink, D. A., Nims, R. W., Darbyshire, J. F., Christodoulou, D., Hanbauer, I., Cox, G. W., Laval, F., Laval, J., Cook, J. A., and Krishna, M. C. (1994) Reaction kinetics for nitrosation of cysteine and glutathione in aerobic nitric oxide solutions at neutral pH. Insights into the fate and physiological effects of intermediates generated in the NO/O₂ reaction, *Chem. Res. Toxicol.* **7**, 519–525.
12. Stamler, J. S., Lamas, S., and Fang, F. C. (2001) Nitrosylation. The prototypic redox-based signaling mechanism, *Cell* **106**, 675–683.
13. Jaffrey, S. R., and Snyder, S. H. (2001) The biotin switch method for the detection of S-nitrosylated proteins, *Sci. STKE* **2001**, PL1.
14. de Gonzalez, O., Martinez-Ruiz, A., Villanueva, L., Lopez-Ferrer, D., Higuera, M. A., Tarin, C., Rodriguez-Crespo, I., Vazquez, J., and Lamas, S. (2005) S-nitrosylation of Hsp90 promotes the inhibition of its ATPase and endothelial nitric oxide synthase regulatory activities, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8525–8530.
15. Zai, A., Rudd, M. A., Scribner, A. W., and Loscalzo, J. (1999) Cell-surface protein disulfide isomerase catalyzes transnitrosation and regulates intracellular transfer of nitric oxide, *J. Clin. Invest.* **103**, 393–399.
16. Mitchell, D. A., and Marletta, M. A. (2006) Design and characterization of an active site selective Caspase-3 transnitrosating agent, *ACS Chem. Biol.* **1**, 659–665.
17. Mitchell, D. A., Morton, S. U., and Marletta, M. A. (2005) Thioredoxin catalyzes the S-nitrosation of the caspase-3 active site cysteine, *Nat. Chem. Biol.* **1**, 154–158.
18. Nikitovic, D., and Holmgren, A. (1996) S-nitroso-glutathione is cleaved by the thioredoxin system with liberation of glutathione and redox regulating nitric oxide, *J. Biol. Chem.* **271**, 19180–19185.
19. Sliskovic, I., Raturi, A., and Mutus, B. (2005) Characterization of the S-denitrosation activity of protein disulfide isomerase, *J. Biol. Chem.* **280**, 8733–8741.
20. Uehara, T., Nakamura, T., Yao, D., Shi, Z. Q., Gu, Z., Ma, Y., Masliyah, E., Nomura, Y., and Lipton, S. A. (2006) S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration, *Nature* **441**, 513–517.
21. Matsumoto, A., Comatas, K. E., Liu, L., and Stamler, J. S. (2003) Screening for nitric oxide-dependent protein-protein interactions, *Science* **301**, 657–661.
22. Lauffenburger, D. A., and Linderman, J. J. (1993) *Receptors Models for Binding, Trafficking, and Signaling*, Oxford University Press, New York.
23. Meyer, D. J., Kramer, H., Ozer, N., Coles, B., and Ketterer, B. (1994) Kinetics and equilibria of S-nitrosothiol-thiol exchange between glutathione, cysteine, penicillamines and serum albumin, *FEBS Lett.* **345**, 177–180.
24. Singh, S. P., Wishnok, J. S., Keshive, M., Deen, W. M., and Tannenbaum, S. R. (1996) The chemistry of the S-nitrosoglutathione/glutathione system, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14428–14433.
25. Skipper, P. L. (1996) Influence of tertiary structure on nucleophilic substitution reactions of proteins, *Chem. Res. Toxicol.* **9**, 918–923.
26. Fang, M., Jaffrey, S. R., Sawa, A., Ye, K., Luo, X., and Snyder, S. H. (2000) Dexas1: a G protein specifically coupled to neuronal nitric oxide synthase via CAPON, *Neuron* **28**, 183–193.
27. Lane, P., Hao, G., Gross, S. S. (2001) S-nitrosylation is emerging as a specific and fundamental post-translational protein modification: head-to-head comparison with O-phosphorylation, *Sci. STKE* **2001**, RE1.
28. Porter, A. G. (2006) Flipping the safety catch of procaspase-3, *Nat. Chem. Biol.* **2**, 509–510.
29. Barrett, D. M., Black, S. M., Todor, H., Schmidt-Ullrich, R. K., Dawson, K. S., and Mikkelsen, R. B. (2005) Inhibition of protein-tyrosine phosphatases by mild oxidative stresses is dependent on S-nitrosylation, *J. Biol. Chem.* **280**, 14453–14461.
30. Yu, C. X., Li, S., and Whorton, A. R. (2005) Redox regulation of PTEN by S-nitrosothiols, *Mol. Pharmacol.* **68**, 847–854.
31. Wang, C., Trudel, L. J., Wogan, G. N., and Deen, W. M. (2003) Thresholds of nitric oxide-mediated toxicity in human lymphoblastoid cells, *Chem. Res. Toxicol.* **16**, 1004–1013.
32. Kim, J. E., and Tannenbaum, S. R. (2004) Insulin regulates cleavage of procaspase-9 via binding of X chromosome-linked inhibitor of apoptosis protein in HT-29 cells, *Cancer Res.* **64**, 9070–9075.
33. Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H. D., Grantcharova, V., Lauffenburger, D. A., and White, F. M. (2006) Effects of HER2 overexpression on cell signaling networks governing proliferation and migration, *Mol. Syst. Biol.* **2**, 54.

Transcription Factor TFII-I Conducts a Cytoplasmic Orchestra

Ananda L. Roy*

Department of Pathology, Programs in Genetics and Immunology, Tufts University School of Medicine, 150 Harrison Avenue, Boston, Massachusetts 02111

Calcium signaling is involved in virtually all cellular processes and regulates intracellular enzymes located in the cytoplasm, nucleus, and organelles. The intracellular calcium levels (100–200 nM) are 10,000-fold lower than the extracellular levels (1–2 mM) (1). Although the precise reason is unknown, a likely explanation is that calcium can precipitate phosphate, “the currency of life” (1). Thus, it is not surprising that nature has evolved multiple ways to regulate the flow and use of intracellular calcium. As a consequence, dynamic alteration in intracellular Ca^{2+} results in activation of a wide range of signaling pathways and biological responses (2, 3). The transient receptor potential channel 3/6/7 (TRPC 3/6/7) subfamily of cation channels are emerging as important components that regulate phospholipase C (PLC)-dependent intracellular Ca^{2+} entry (4, 5). Although it has been known for some time that Ca^{2+} signaling *via* ion channels can regulate nuclear gene expression, a recent paper by Caraveo *et al.* (6) demonstrates a novel cytoplasmic function of a transcription factor, TFII-I, that regulates Ca^{2+} entry by TRPC3.

Interaction of TRPC3 with PLC- γ *via* the Split PH Domain. Ligand-dependent activation of receptor tyrosine kinases is associated with Ca^{2+} influx, resulting in production of inositol 1,4,5-trisphosphate (IP3) and diacylglycerol from phosphatidylinositol 4,5-bisphosphate through the lipase action of PLC- γ (1–3). For instance, ligation of the

growth factor receptors (GFRs) or immune receptors results in activation of PLC- γ and transient Ca^{2+} influx (1–3). However, recent studies demonstrated a novel lipase-independent function of PLC- γ that regulates intracellular Ca^{2+} entry *via* TRPC3 (7, 8; Figure 1, panel a). A pleckstrin homology (PH) domain has been identified in a variety of signaling intermediates, and scientists generally believed that PH domains mediated membrane proximal lipid binding (9). However, a gestalt algorithm search revealed that a PH-like domain or a split PH domain is present in a variety of proteins, including PLC- γ and TRPC3 (8). Indeed, PLC- γ interacts with TRPC3 *via* its split PH domain. Thus, interaction of the two halves of the split PH domain from two proteins results in a structural complementation and enhanced surface expression of TRPC3, both of which are required for agonist-induced Ca^{2+} entry (ACE) (8). The recent study by Caraveo *et al.* (6) throws in another unexpected twist to this story.

Double Duty Transcription Factor. The ubiquitously expressed transcription factor TFII-I was identified and characterized on the basis of its ability to bind DNA sequence elements and activate transcription (10, 11). Later, researchers demonstrated that it is actually a signal-induced multifunctional transcription factor that is tyrosine-phosphorylated in response to B cell receptor (BCR) and growth factor signaling (12). Induced tyrosine phosphorylation of TFII-I is necessary for its transcriptional function. Surprisingly, TFII-I

ABSTRACT In response to extracellular ligands, surface receptor tyrosine kinases and G-protein-coupled receptors activate isoforms of phospholipase C (PLC) and initiate calcium signaling. PLC can activate expression of surface transient receptor potential channels (TRPC) such as TRPC3, which modulate calcium entry through the plasma membrane. A recent paper shows that competitive binding of cytoplasmic TFII-I, a transcription factor, to PLC- γ results in inhibition of TRPC3-mediated agonist-induced Ca^{2+} entry. These results establish a novel cytoplasmic function for TFII-I.

*Corresponding author,
ananda.roy@tufts.edu.

Published online November 17, 2006

10.1021/cb6004323 CCC: \$33.50

© 2006 by American Chemical Society

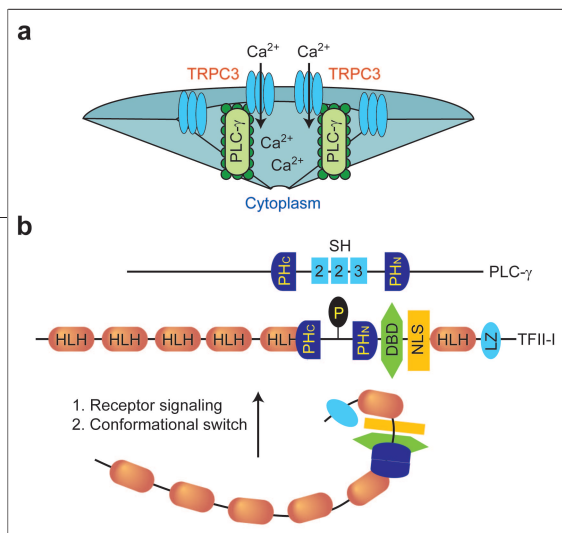


Figure 1. Interactions with a PH domain. a) PLC- γ interacts with TRPC3 via its split pH domain. This interaction is necessary for enhanced surface expression of TRPC3 and Ca²⁺ influx. b) Receptor signaling results in tyrosine phosphorylation and a conformational switch in transcription factor TFII-I, which is found in the cytoplasm in resting cells. TFII-I possesses multiple helix-loop-helix (HLH) interaction domains. In the absence of signals, the nuclear localization signal (NLS) of TFII-I is probably masked by limited intramolecular interaction via these HLH domains. The DNA binding domain (DBD) of TFII-I overlaps with the split PH domain. Thus, a receptor-mediated tyrosine phosphorylation may not only expose the split PH domain but also unmask both the DBD and the NLS in TFII-I for its subsequent nuclear action. Signaling also likely uncovers the leucine zipper in TFII-I. Signal-induced TFII-I interacts with the SH-2 domain of PLC- γ via both its phosphotyrosine moiety and the split PH domain.

was shown to physically and functionally interact with Bruton's tyrosine kinase (Btk), which is activated upon BCR engagement (13). However, consistent with its ubiquitous expression pattern, TFII-I is also tyrosine-phosphorylated by Src family kinases upon growth factor activation (14). Together, these observations led to the demonstration that TFII-I is latently found in the cytoplasm; however, upon activation (tyrosine phosphorylation) by signaling through surface receptors, it translocates to the nucleus for transcriptional activation (14–16). Using a phosphopeptide that includes the Btk phosphorylation site in TFII-I (17) as bait, Caraveo *et al.* (6) “fished-out” PLC- γ as an interacting partner. Tyrosine-phosphorylated TFII-I interacts with the Src-homology (SH)-2 domain of PLC- γ (Figure 1, panel b). The most exciting part of this study further reveals that TFII-I has a split PH domain that mimics that of TRPC3 and is involved in PLC- γ interaction (6). Hence, TFII-I can compete with TRPC3 for PLC- γ binding. The

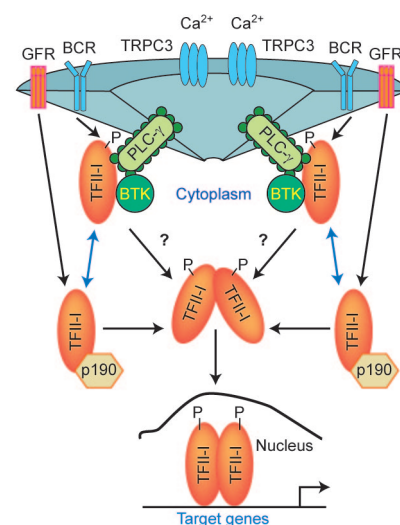
split PH domain in TFII-I is in a “closed conformation” in its basal state but opens up upon tyrosine phosphorylation. Open TFII-I conformation results in its interaction with PLC- γ via both the phosphotyrosine and split PH domains (Figure 1, panel b). Tethering away, PLC- γ lowers the surface expression of TRPC3 and reduces Ca²⁺ channeling activity (Figure 2). Interaction of TFII-I with PLC- γ and consequently inhibition of ACE via TRPC3 do not require either the nuclear localization or transcriptional function of TFII-I because a nuclear localization signal deficient mutant (Δ NLS) of TFII-I was competent in PLC- γ binding (6). Surprisingly, en-

dogenous TFII-I was constitutively associated with PLC- γ prior to signaling, and BCR signaling did not enhance or diminish this interaction (6). However, because the cells were cultured in serum, which could trigger

tyrosine phosphorylation of TFII-I in the absence of BCR signaling (15), the observed constitutive tyrosine phosphorylation and interaction with PLC- γ may not be totally unexpected.

Using a loss-of-function (post-transcriptional gene silencing) experiment, Caraveo *et al.* (6) show that TFII-I is critical for TRPC3 function in multiple cell types because silencing of TFII-I in PC12 neuronal cells or human embryonic kidney cell lines led to an increase in surface expression of TRPC3 and, consequently, enhanced Ca²⁺ influx (6). Importantly, the increase in surface TRPC3 does not appear to be at the level of transcription, and this suggests a post-transcriptional effect of TFII-I on TRPC3 expression. Reconstitution experiments convincingly demonstrated further that the nuclear function of TFII-I is not necessary for this effect. Ectopic expression of a transcriptionally incompetent Δ NLS-TFII-I in these cells restored normal surface TRPC3 expression and Ca²⁺ influx, consistent with the requirement of transcription-independent cytoplasmic function of TFII-I in regulating Ca²⁺ influx (6).

Figure 2. TFII interactions in the cytoplasm. Upon BCR (or GFR) signaling, tyrosine-phosphorylated TFII-I interacts with PLC- γ , titrating it away from TRPC3; this results in a reduction of both ion channel activity and Ca²⁺ influx. Btk is presumably in the same complex. Distinct pools of cytoplasmic TFII-I are probably associated with distinct cytoplasmic tethers and thus capable of receiving distinct signal via different cell surface receptors. The pool associated with PLC- γ and Btk might receive signal from BCR, whereas the pool associated with p190RhoGAP might receive signal from GFR. GFR-mediated tyrosine phosphorylation of TFII-I results in its dissociation from the cytoplasmic complex, dimerization, and subsequent nuclear translocation for activation of target genes. A dynamic equilibrium probably occurs between the different pools of TFII-I.



How Can a Factor Be in Two Locations at the Same Time?

That tyrosine-phosphorylated TFII-I is found associated with PLC- γ in the cytoplasm (6) and is also required for transcription in the nucleus (14–17) presents a conundrum (Figure 2). Several explanations could exist. First, it is possible that there are two (or perhaps more) distinct pools of cytoplasmic TFII-I. In this scheme, different pools of TFII-I might be associated with distinct cytoplasmic tethers and thus be capable of receiving distinct signal *via* different cell surface receptors (Figure 2). For instance, the pool of TFII-I associated with PLC- γ (and Btk) might receive signal from BCR, whereas the pool associated with p190RhoGAP might receive signal from GFRs (16). This latter pool would dissociate from the complex upon signaling and translocate to the nucleus for gene activation (16). Perhaps a dynamic equilibrium occurs between these pools. A second alternative is that only one pool of cytoplasmic TFII-I exists but that different signals (and corresponding kinases) target distinct sites in TFII-I. Depending on the signal and the particular phosphotyrosine moiety, TFII-I might interact with different partners and thus activate distinct downstream pathways. Yet a third possibility is that the different interactions (and thus functions) of TFII-I are kinetically distinct. In this model, the immediate early signaling events (2–5 min) would lead to tyrosine phosphorylation of TFII-I and its subsequent membrane proximal interactions with PLC- γ (and Btk and/or other cytoplasmic tethers). These events might in turn result in Ca²⁺ entry *via* regulation of TRPC3 (6). Because the nuclear translocation of TFII-I occurs later (12–20 min) (14, 16), a “kinetic switch” might result in the release of tyrosine-phosphorylated TFII-I from the cytoplasmic tether and subsequent dimerization and nuclear translocation. A fourth alternative is also plausible. In this scenario, while a subset of the events occur near the plasma membrane, others occur near the nuclear membrane. Analysis of whole

cell lysates might have obfuscated such a distinction. Given that phosphoinositols control chromatin dynamics and that PLC- γ plays an important role in chromatin remodeling and gene regulation (18), TFII-I might connect phospholipid metabolism to signal-induced gene regulation.

Biological Implications. Regardless of how TFII-I mediates signaling events in the cytoplasm and activates transcription in the nucleus, that it interacts with PLC- γ and regulates Ca²⁺ influx *via* TRPC3 is an exciting development in the field. Moreover, this discovery has potentially important physiological ramifications. The genes that encode members of the TFII-I family occupy adjacent locations in human chromosome 7, a portion of which is deleted in a haploinsufficient manner in Williams–Beuren syndrome (WBS) (19). WBS is a rare developmental disorder that is caused by a hemizygous microdeletion of ~1.5 megabases, spanning 17 genes at chromosomal location 7q11.23 (20). Although this multi-system dysfunction with unusual craniofacial, behavioral, and cognitive features is probably caused by haplo-insufficiency of several genes, a good genotype–phenotype correlation may be present. Consistent with this notion, murine models with targeted deletions in TFII-I family genes are providing useful information about potential WBS phenotypes (21). That neuronal calcium signaling requires a large number of low-affinity IP3 receptors in Purkinje cells of the cerebellum, where they are not very effective at calcium release, suggests perhaps a “conformational coupling” between IP3 receptors and TRPC3 in these cells (1, 22). Given the unusually high expression pattern of TFII-I family genes in normal neuronal tissues and its extensive cytoplasmic expression in Purkinje cells (11, 23), it is likely that alterations in the levels of TFII-I observed in WBS regulate such coupling and thereby regulate neurocognitive processes.

Acknowledgment: The work done in my lab is supported in part by NIH grants AI45150 and HD046034. I thank Peter Brodeur for critically reading the manuscript.

REFERENCES

- Clapham, D. E. (1995) Calcium signaling, *Cell* 80, 259–268.
- Berridge, M. J. (2000) The versatility and universality of calcium signaling, *Nat. Rev. Mol. Biol.* 1, 11–21.
- Patterson, R. L., van Rossum, D. B., Nikolaidis, N., Gill, D. L., and Snyder, S. H. (2005) Phospholipase C- γ : diverse roles in receptor-mediated calcium signaling, *Trends Biochem. Sci.* 30, 688–697.
- Ramsey, I. S., Delling, M., and Clapham, D. E. (2006) An introduction to TRP channels, *Annu. Rev. Physiol.* 68, 619–647.
- Trebak, M., Vazquez, G., Bird, G., and Putney, J. W. (2003) The TRPC3/6/7 subfamily of cation channels, *Cell Calcium* 33, 451–461.
- Caraveo, G., van Rossum, D. B., Patterson, R. L., Snyder, S. H., and Desiderio, S. (2006) Action of TFII-I outside the nucleus as an inhibitor of agonist-induced calcium entry, *Science* 314, 122–125.
- Patterson, R. L., van Rossum, D. B., Ford, D. L., Hurt, K. J., Bae, S. S., Suh, P. G., Kuroski, T., Snyder, S. H., and Gill, D. L. (2002) Phospholipase C- γ is required for agonist-induced Ca²⁺ entry, *Cell* 111, 529–541.
- van Rossum, D. B., Patterson, R. L., Sharma, S., Barrow, R. K., Kornberg, M., Gill, D. L., and Snyder, S. H. (2005) Phospholipase C γ 1 controls surface expression of TRPC3 through an intermolecular PH domain, *Nature* 434, 99–104.
- Lemmon, M. A. (2005) Pleckstrin homology domains: Two halves make a hole? *Cell* 120, 574–576.
- Roy, A. L., Meisterernst, M., Pogoniec, P., and Roeder, R. G. (1991) Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF, *Nature* 354, 245–248.
- Roy, A. L., Du, H., Gregor, P. D., Novina, C. D., Martinez, E., and Roeder, R. G. (1997) Cloning of an Inr and E-box binding protein TFII-I that interacts physically and functionally with USF, *EMBO J.* 16, 7091–7104.
- Roy, A. L. (2001) Biology and biochemistry of the signal induced transcription factor TFII-I, *Gene* 274, 1–13.
- Yang, W., and Desiderio, S. (1997) BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement, *Proc. Natl. Acad. Sci. U.S.A.* 94, 604–609.
- Cheriyath, V., Desgranges, Z., and Roy, A. L. (2002) c-Src dependent transcriptional activation of TFII-I, *J. Biol. Chem.* 277, 22798–22805.
- Novina, C. D., Kumar, S., Bajpai, U., Cheriyath, V., Zhang, K., Pillai, S., Wortis, H. H., and Roy, A. L. (1999) Regulation of nuclear localization and transcriptional activity of TFII-I by Bruton's tyrosine kinase, *Mol. Cell. Biol.* 19, 5014–5024.
- Jiang, W., Sordella, R., Chen, G. C., Hakre, S., Roy, A. L., and Settleman, J. (2005) An FF domain-dependent protein interaction mediates a signaling pathway for growth factor-induced gene expression, *Mol. Cell* 17, 23–35.

17. Egloff, A. M., and Desiderio, S. (2001) Identification of phosphorylation sites for Bruton's tyrosine kinase within the transcriptional regulator BAP/TFII-I, *J. Biol. Chem.* *276*, 27806–27815.
18. Rando, O. J., Chi, T. H., and Crabtree, G. R. (2003) Second messenger control of chromatin remodeling, *Nat. Struct. Biol.* *10*, 81–83.
19. Perez Jurado, L. A., Wang, Y. K., Peoples, R., Coloma, A., Cruces, J., and Francke, U. (1998) A duplicated gene in the breakpoint regions of the 7q11.23 Williams-Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK, *Hum. Mol. Genet.* *7*, 325–334.
20. Hirota, H., Matsuoka, R., Chen, X. N., Salandanan, L. S., Lincoln, A., Rose, F. E., Sunahara, M., Osawa, M., Bellugi, U., and Korenberg, J. R. (2003) Williams syndrome deficits in visual spatial processing linked to GTF2IRD1 and GTF2I on chromosome 7q11.23, *Genet. Med.* *5*, 311–321.
21. Tassabehji, M., Hammond, P., Karmiloff-Smith, A., Thompson, P., Thorgeirsson, S. S., Durkin, M. E., Popescu, N. C., Hutton, T., Metcalfe, K., Rucka, A., Stewart, H., Read, A. P., Maconochie, M., and Donnai, D. (2005) GTF2IRD1 in craniofacial development of humans and mice, *Science* *310*, 1184–1187.
22. Sharp, A. H., McPherson, P. S., Dawson, T. M., Aoki, C., Campbell, K. P., and Snyder, S. H. (1993) Differential immunohistochemical localization of inositol 1,4,5-trisphosphate- and ryanodine-sensitive Ca²⁺ release channels in rat brain, *J. Neurosci.* *13*, 3051–3063.
23. Danoff, S. K., Taylor, H. E., Blackshaw, S., and Desiderio, S. (2004) TFII-I, a candidate gene for Williams syndrome cognitive profile: parallels between regional expression in mouse brain and human phenotype, *Neuroscience* *123*, 931–938.

Desensitization at the Interface

James R. Howe*

Department of Pharmacology, Yale University School of Medicine, SHM B-251, 333 Cedar Street, New Haven, Connecticut 06520

Neurons make specialized contacts with their target cells called synapses, sites where small-molecule neurotransmitters are released in response to electrical activity. Once released, the neurotransmitters bind to the extracellular domains of receptor proteins in the cell membrane of the postsynaptic cell. At most excitatory synapses in the brain, glutamate is the neurotransmitter, and rapid neuron-to-neuron communication is mediated by α -amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) receptors, a subtype of ionotropic glutamate receptor (iGluR). The binding of glutamate, AMPA, or other agonists to these receptors initiates a series of conformational changes (“gating”) that allow cations to flow through the channels down their electrochemical gradients. The resultant transmembrane (TM) current initiates a regenerative electrical signal called an action potential, but the channels quickly become unresponsive during sustained exposure to glutamate. A recent paper by Armstrong *et al.* (1) provides the first high-resolution views of receptor conformations adopted during desensitization.

Both structural and functional data indicate that iGluRs are tetrameric assemblies, organized as dimers of dimers, with each subunit containing a binding site for glutamate (2). Individual subunits are composed of four domains, with the extracellular ligand binding domain (LBD) and the pore domain (composed of membrane helices) being sufficient for channel gating. X-ray crystallographic structures of the isolated LBD show that it is a clamshell-like structure and that

glutamate and other ligands bind within a deep cleft between the two halves of the clamshell (LBD domains 1 and 2). In agonist-bound structures, domain 2 is displaced $\sim 20^\circ$ relative to its position in *apo* structures, closing the clamshell and trapping agonists (3, 4). Domain 2 is connected to the extracellular ends of the first two TM helices, and the movement of domain 2 is thought to result in rearrangements of the helices that open the channel. The available evidence, although circumstantial, supports the conclusion that closing of the LBD is the initial large-scale conformational change that triggers subsequent gating rearrangements at the level of the pore (3, 5).

Most receptors desensitize upon continued exposure to neurotransmitter, and desensitization is especially prominent for AMPA receptors. The current generated when channels formed from the AMPA receptor subunit GluR2 were rapidly exposed to a high concentration of glutamate is illustrated (Figure 1, panel a). The diffusion-limited binding of glutamate results in the near-simultaneous activation of hundreds of receptors in the patch, generating a rapidly developing inward current, which then decays to 1–2% of its peak value in 15–20 ms. The decay of the current results from entry of the channels into desensitized states, and the accumulation of channels in these unresponsive states limits the size of the current evoked by a second application of glutamate (which progressively increases as channels recover from desensitization; Figure 1, panel a). Desensitization predominates at low micro-

ABSTRACT Normal brain function requires the faithful transmission and integration of information on a timescale of milliseconds, and this rapid signaling is mediated by cell membrane receptors that are ligand-gated ion channels. Fast excitatory transmission occurs when synaptically released glutamate opens channels in neighboring neurons, but these channels desensitize rapidly during sustained high-frequency firing. Recent structural data have begun to provide important insights into the molecular mechanisms that underlie channel activation and desensitization.

*Corresponding author,
james.howe@yale.edu.

Published online November 17, 2006

10.1021/cb6004382 CCC: \$33.50

© 2006 by American Chemical Society

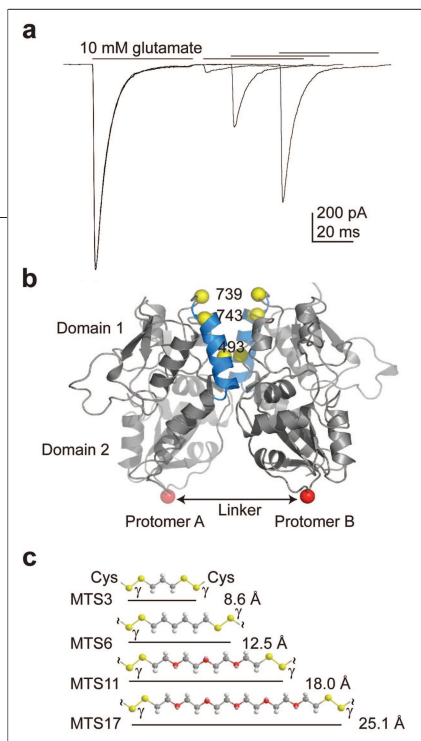


Figure 1. Probing the GluR2 channel. a) Inward currents evoked by the rapid application (bars) of 10 mM glutamate in a patch of membrane that contains hundreds of recombinant GluR2 channels. The first application of glutamate causes a rapidly activating current that quickly fades as the channels desensitize. The responses to a second application of glutamate (made at three different intervals) are reduced. b) Ribbon diagram of the wild-type GluR2-S1S2 structure. The LBD crystallizes as a dimer in which contacts are made between domain 1 residues in each monomer across the two-fold axis of symmetry. This dimer interface is in blue, and the yellow balls show positions at which cysteines were introduced by site-directed mutagenesis. The red spheres at the bottom of domain 2 indicate the position where the LBD would be connected to the first and second TM helices in the full-length receptor. c) Reagents used to cross-link non-native cysteines across the dimer interface. Adapted with permission from Elsevier, copyright 2006 (1) and the Biophysical Society, copyright 2006 (11).

molar concentrations of glutamate and appears to have evolved as a mechanism to protect neurons from excitotoxic damage when ambient levels of glutamate are elevated (e.g., during cerebral ischemia). This protection comes, however, with a cost. Although at many synapses glutamate clearance mechanisms remove glutamate before most channels desensitize during a single synaptic event, high-frequency trains of impulses can lead to the accumulation of

channels in desensitized states and ultimately result in the failure of transmission. Because information processing in the brain is frequency-encoded, the rate of desensitization (and recovery from it) significantly impact synaptic signaling.

Previous work that combined crystallographic and biochemical studies of the isolated GluR2 LBD with mutagenesis and electrophysiology on full-length receptors led to the conclusion that desensitization occurred when monomer–monomer contacts along the two-fold axis of symmetry in each receptor dimer (the “dimer interface”) slipped, rendering closure of the LBD less effective at promoting conformations of the membrane helices that allow ion flux (6). Building on this previous work, Armstrong and colleagues (1) introduced cysteines into positions along the dimer interface that the prior GluR2 crystal structures predicted would be largely inaccessible to bulk solvent in resting and activated channel states (Figure 1, panel b). Using voltage-clamp recording of currents through recombinant channels expressed in oocytes, they then measured the rate at which these cysteines were modified by sodium (2-sulfonatoethyl)methane thiosulfonate (MTSES) under conditions that favored resting, active, or desensitized channel states. (To minimize potentially confounding MTSES modification of native cysteines, the authors employed a recombinant GluR2 construct from which the amino terminal domain had been removed and in which most native cysteines were replaced by conservative mutation.) They found that MTSES reaction rates were ~2 orders of magnitude faster for desensitized channels, the first direct evidence that desensitization is indeed accompanied by rearrangement of the dimer interface at the level of the LBD.

To begin to quantify the extent of the rearrangement, the authors used bifunctional thiol-directed reagents that cross-link nearby cysteine residues (Figure 1, panel c). Cysteines were introduced at various depths along the interface, and four linkers that

ranged in length from 8.6 to 25.1 Å were used to cross-link cysteines located on opposing sides of the dimer interface. Each reagent would be predicted to cross-link cysteine pairs in resting and activated states, whereas cross-linkers whose fully extended length was less than the separation of the residues following rearrangement of the interface would be expected to reduce desensitization. Results with these molecular rulers indicated that during desensitization the two monomers separate at the top and bottom of the V-shaped interface by 16.2 and 12.4 Å, respectively.

While characterizing the various cysteine mutants, the authors identified one that gave very small steady-state currents in normal conditions but much larger currents after the channels were exposed to the reducing agent dithiothreitol (DTT). Western blotting demonstrated that the mutant channels formed spontaneous inter-subunit disulfide bonds, although previous crystal structures suggested that the non-native cysteines were much too far apart to do so in conformations adopted by non-desensitized channels. This led Armstrong and co-workers to hypothesize that the disulfide bond was formed when the LBD adopted a conformation not observed in previous crystal structures, perhaps one preferentially associated with desensitization. Studies on the DTT sensitivity of additional cysteine mutants highlighted the importance of a region on the surface of domain 2 (below the interface) that contains a loop pointing across the two-fold axis to residues on the adjacent monomer. Binding studies on the isolated LBD, as well as functional measurements made with a domain 2 histidine mutant (G725) in the absence and presence of Zn^{2+} , supported the view that tethering the relevant domain 2 residues across the two-fold axis promoted desensitization (rather than stabilizing a resting closed state).

Armed with this additional information, the authors introduced the S729C mutation into the GluR2-S1S2 LBD and solved the

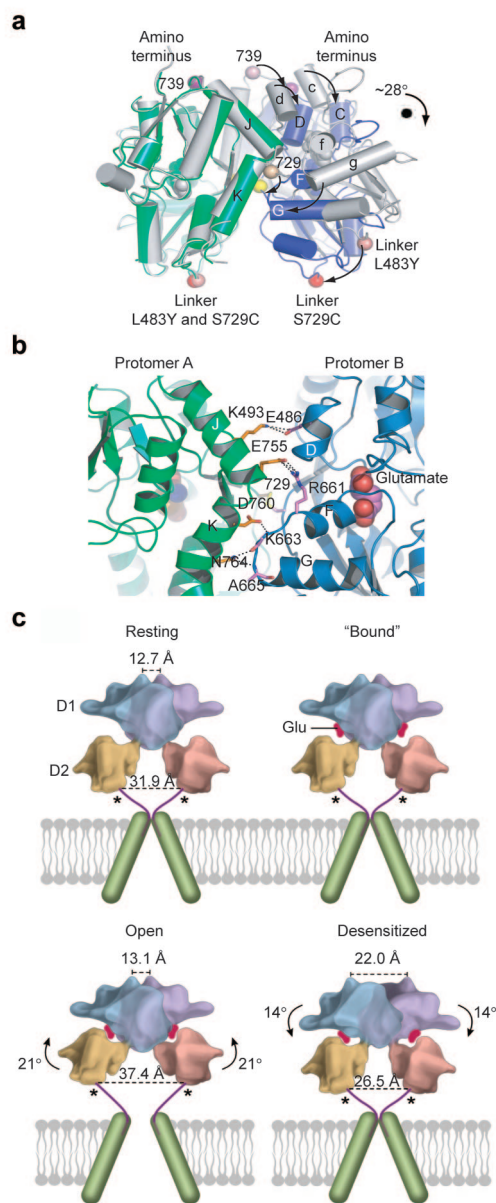


Figure 2. Structural changes in the desensitized GluR2 LBD structure. **a)** Superposition of GluR2 LBD structures putatively corresponding to conformations adopted by active (L483Y, gray) and desensitized (S729C, blue and green) channels. The view is roughly down the two-fold axis with the linkers that replace residues connected to the channel pore shown as pink or red spheres (bottom). The arrows indicate displacements of the linkers and residues along the dimer interface that characterize the transition from active to desensitized conformations. Domain 1 residues along the dimer interface separate, whereas separation of the linkers decreases. **b)** Close-up view of the dimer interface in the S729C (desensitized) structure reveals new interactions formed between amino acids comprising the loop between helices F and G (R661, K663, and A665) and residues in helix K of the adjacent monomer. **c)** Cartoon showing the LBD and TM helices of an AMPA receptor dimer in resting, bound, open, and desensitized states. Glutamate binds to the open LBD. Movement of domain 2 in each monomer closes the LBD, creating what is likely an unstable transition state in which both the LBD and channel pore are closed (not shown). Movements of the TM helices partially relieve this instability and open the channel (on average 2–3 times). Eventually, however, separation of the dimer interface dissipates the work initially done by closure of the LBD, and the channel adopts a stable state in which glutamate binding and activation gating are uncoupled. Adapted with permission from Elsevier, copyright 2006 (1).

dence that during recovery from desensitization the first two bound glutamates (one in each dimer) dissociate rapidly, whereas the last two glutamates dissociate much more slowly (7). The structure also suggests that the asymmetry of the amino terminal domains observed previously in cryo-electron microscopy images (8) results from asymmetric rearrangement of the LBD dimer interface.

The X-ray crystallographic data, as well as the new functional results, speak directly to stable conformations adopted under equilibrium conditions but do not provide information on receptor kinetics. In this sense, the available crystal structures are somewhat like a set of vacation snapshots, which reveal where you were but not how you got there (or how long it took). However, although the crystallographic data do not discriminate between alternative kinetic mechanisms, they do provide a clear roadmap for additional structure–function work on full-length receptors. For example, kinetic studies done before structural data were obtained suggested that the rate at which channels recover from desensitization was primarily determined by the rate at which agonists dissociated from desensitized channels (9, 10), a proposal that nicely accounts for the well-known correlation between agonist affinity and the rate of recovery (11). In contrast, more recent studies concluded that the rate-determining steps during recovery are sequential reassembly of the two dimer interfaces (12). Interestingly, the new LBD structure shows that separation of the dimer interface results in additional interactions across the binding cleft that would be predicted to stabilize the closed-cleft conformation, whereas the new inter-subunit interactions between domain 2 residues are relatively minimal. The newly identified interactions in these two regions are obvious targets for kinetic studies that might further illuminate the sequence of events during recovery.

Ultimately, structural data on receptor constructs that include both the LBD and the

crystal structure at 2.3 Å resolution. Like previous structures of the isolated GluR2 binding core, the S729C mutant crystallized as a dimer (Figure 2). However, the dimer interface is separated substantially, as suggested by the oocyte experiments on functional receptor mutants. The distances characterizing this rearrangement in the crystal

structure agree well with those estimated in the cross-linking experiments; this supports the conclusion that the structure is similar to the conformation of the LBD adopted by desensitized channels.

Unlike previous GluR2-S1S2 structures, the two LBDs are asymmetric, a possible explanation for electrophysiological evi-

channel pore are required if the mechanisms of activation and desensitization are to be understood, likely in combination with high-resolution electrophysiology and spectroscopic measurements to provide real-time correlations between ion flux and protein movements (13). Nevertheless, the recent work of Armstrong and colleagues represents another significant advance in unraveling how a major class of synaptic proteins works.

13. Blunck, R., Cordero-Morales, J. F., Cuello, L. G., Perozo, E., and Bezanilla, F. (2006) Detection of the opening of the bundle crossing in KcsA with fluorescence lifetime spectroscopy reveals the existence of two gates for ion conduction, *J. Gen. Physiol.* 128, 569–581.

REFERENCES

1. Armstrong, N., Jasti, J., Beich-Frandsen, M., and Gouaux, E. (2006) Measurement of conformational changes accompanying desensitization in an ionotropic glutamate receptor, *Cell* 127, 85–97.
2. Mayer, M. L. (2006) Glutamate receptors at atomic resolution, *Nature* 440, 456–462.
3. Armstrong, N., and Gouaux, E. (2000) Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core, *Neuron* 28, 165–181.
4. Jin, R., Homing, M., Mayer, M. L., and Gouaux, E. (2002) Mechanism of activation and selectivity in a ligand-gated ion channel: structural and functional studies of GluR2 and quisqualate, *Biochemistry* 41, 15635–15643.
5. Jin, R., Banke, T. G., Mayer, M. L., Traynelis, S. F., and Gouaux, E. (2003) Structural basis for partial agonist action at ionotropic glutamate receptors, *Nat. Neurosci.* 6, 803–810.
6. Sun, Y., Olson, R., Homing, M., Armstrong, N., Mayer, M., and Gouaux, E. (2002) Mechanism of glutamate receptor desensitization, *Nature* 417, 245–253.
7. Robert, A., and Howe, J. R. (2003) How AMPA receptor desensitization depends on receptor occupancy, *J. Neurosci.* 23, 847–858.
8. Nakagawa, T., Cheng, Y., Ramm, E., Sheng, M., and Walz, T. (2005) Structure and different conformational states of native AMPA receptor complexes, *Nature* 433, 545–549.
9. Partin, K. M., Fleck, M. W., and Mayer, M. L. (1996) AMPA receptor flip/flop mutants affecting deactivation, desensitization, and modulation by cyclothiazide, aniracetam, and thiocyanate, *J. Neurosci.* 16, 6634–6647.
10. Raman, I. M., and Trussell, L. O. (1995) The mechanism of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionate receptor desensitization after removal of glutamate, *Biophys. J.* 68, 137–146.
11. Zhang, W., Robert, A., Vogensen, S. B., and Howe, J. R. (2006) The relationship between agonist potency and AMPA receptor kinetics, *Biophys. J.* 91, 1336–1346.
12. Robert, A., Armstrong, N., Gouaux, J. E., and Howe, J. R. (2005) AMPA receptor binding cleft mutations that alter affinity, efficacy, and recovery from desensitization, *J. Neurosci.* 25, 3752–3762.

Intracellular Trafficking of Porphyrins

Iqbal Hamza*

Department of Animal and Avian Sciences and Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742

With a few exceptions, heme, a metalloporphyrin, is synthesized via a multistep biosynthetic pathway with well-defined intermediates that are highly conserved throughout evolution. Depending upon the organelle and cell type, heme pathway intermediates are utilized for the synthesis of other tetrapyrrole compounds, including bilins, chlorophylls, and corrins (1–3). Despite our extensive knowledge of heme biosynthesis, the intracellular trafficking of heme and porphyrins are not well understood (Figure 1). In a recent issue of *Nature*, Krishnamurthy *et al.* (4) report on the identification of ABCB6 (ATP-binding cassette, subfamily B, member 6), a mitochondrial outer membrane (OM) transporter, which was shown by different experimental approaches to be able to transport porphyrins. ABCB6 is proposed to translocate coproporphyrinogen III (CPgenIII), a heme precursor, from the cytoplasm to the mitochondria for synthesis of heme. Their findings therefore represent a significant advancement in our understanding of intraorganellar porphyrin transport in mammalian cells.

Porphyrins are heterocyclic organic rings made from four pyrrole subunits linked via methine bridges. The name porphyrin is derived from the Greek word *prophura* for purple because the extensive conjugation of these tetrapyrroles gives them their violet-red hues. Heme is an iron-containing porphyrin and serves as a prosthetic group for many biological processes, including oxidative metabolism, xenobiotic detoxification, the synthesis and sensing of diatomic

gases, cellular differentiation, gene regulation at the level of transcription, protein translation and targeting, and protein stability.

Heme synthesis culminates in the mitochondrial matrix, but the eight sequential enzymatic steps are spatially separated between the cytoplasm and the mitochondria. CPgenIII is a product of the fifth enzyme in the heme pathway, uroporphyrinogen III decarboxylase, and enters the mitochondria to undergo three additional enzymatic reactions to generate heme (Figure 1, panel a) (5). The pathway culminates when ferrous iron is catalytically inserted into the protoporphyrin IX (PPIX) ring by ferrochelatase (FECH), a mitochondrial inner-membrane (IM)-associated enzyme (6, 7). Importantly, iron and porphyrins are toxic to cells; iron generates hydroxyl radicals from Fenton-based reactions, and PPIX catalyzes light-dependent generation of oxygen radicals. Moreover, the end product of the synthetic pathway is heme, a cytotoxic macrocycle with peroxidase activity. Consequently, cellular heme synthesis is coupled with iron availability and apo-protein synthesis to prevent uncoordinated accumulation of iron, porphyrin intermediates, and heme (8, 9).

Two unanswered questions are pertinent to heme synthesis in eukaryotic cells: how does δ -aminolevulinic acid (ALA), the first intermediate of the pathway, exit from the mitochondrial matrix into the cytoplasm, and how does CPgenIII translocate from the cytoplasm into the intermembrane mitochondrial space (IMS)? The results from Krishnamurthy *et al.* (4) provide evidence that ABCB6 may be responsible for

ABSTRACT Hemes are porphyrins that play a critical role in diverse biological processes. Heme synthesis culminates in the mitochondrial matrix, but the eight-step biosynthetic pathway is spatially shared between the mitochondria and cytoplasm. A recent paper describes the nature of the transporter that translocates the heme precursor coproporphyrinogen III into the mitochondria for heme synthesis. The identification of ABCB6 (ATP-binding cassette, subfamily B, member 6) and future studies aimed at precisely delineating the mechanism and the physiological nature of its ligand(s) will further enhance our current understanding of the intracellular movement of tetrapyrroles in eukaryotes.

*Corresponding author,
hamza@umd.edu.

Published online November 17, 2006

10.1021/cb600442b CCC: \$33.50

© 2006 by American Chemical Society

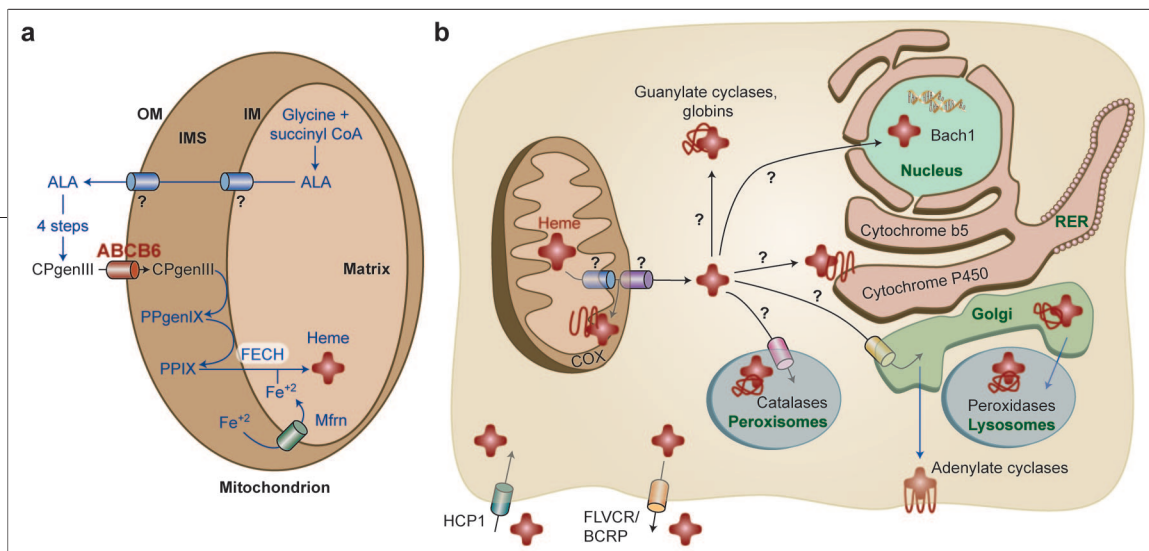


Figure 1. Schematic model of mammalian heme homeostasis. Presumptive heme pathways that are currently unknown are marked with a question mark. **a)** In eukaryotic cells, heme is synthesized *via* a multistep pathway that is spatially separated between the cytosol and the mitochondria. The mechanism of export of ALA from the mitochondria is unknown. ALA is converted to CPgenIII *via* four enzymatic reactions, and CPgenIII is imported into the mitochondria, presumably by ABCB6, an OM transporter, or by OGC. CPgenIII is converted to protoporphyrinogen (PPGenIX) and PPIX on the IMS side of the IM. The final step of heme synthesis occurs by the enzymatic chelation of ferrous iron (Fe^{+2}) into PPIX catalyzed by FECH located on the matrix side of the IM. Mitoferrin (Mfrn)/Mrs 3/4 is predicted to import Fe^{+2} into the mitochondria. CoA = coenzyme A. **b)** The nascent heme moiety is somehow transported through two mitochondrial membranes and incorporated into a multitude of hemoproteins, presumably by hemochaperones, in different cellular compartments. Recent studies have identified HCP1 as the heme importer, primarily in intestinal cells, whereas FLVCR and BCRP are predicted to export heme in erythrocytes. COX = cytochrome oxidase, RER = rough endoplasmic reticulum.

CPgenIII transport. However, the recent work by Kabe *et al.* (10) suggests that the 2-oxoglutarate carrier (OGC) may also perform the same function. Thus, additional work is required to firmly resolve the identity of the transporter and the biological nature of its ligand. Both studies use the commercially available oxidized planar conjugated macrocycle, coproporphyrin III (CPIII), rather than the physiologically relevant substrate CPgenIII, a reduced nonplanar porphyrin.

The findings that ABCB6 expression is coordinated with heme synthesis and cellular function provide further proof of the importance of orchestrated timing of biological processes during development. As erythrocytes mature, the primary function of these oxygen-carrying cells is hemoglobin synthesis, which is underscored by a dramatic and highly coordinated up-regulation of heme and globin production. Using biochemical and cell culture assays with wild-type and mutant forms of ABCB6, the authors provide evidence that *Abcb6* messenger RNA (mRNA) and protein are up-regulated in erythroleukemia and in G1ER cells, which are an immature red cell precursor line (4). They show that ABCB6 binds porphyrins, including heme, but

binding and competition assays suggest that the substrate is likely to be CPIII. ABCB6 activity was found to be highly regulated by cellular heme levels. Inhibition of heme synthesis by succinyl acetone resulted in down-regulation of *Abcb6*, whereas ectopic expression of *Abcb6* resulted in increased PPIX accumulation and up-regulation of mRNA for several heme biosynthesis genes. Notably, knock-down of *Abcb6* by RNA-mediated interference resulted in diminished heme synthesis, further proof that *Abcb6* expression is coordinated with heme synthesis. Furthermore, genetic ablation of a single copy of *Abcb6* in mouse embryonic stem cells revealed a haploid insufficiency phenotype; *Abcb6*^{+/-} heterozygous cells had accumulated half the levels of PPIX compared with wild-type cells when treated with the heme precursor, ALA. Altogether, the paper provides evidence that ABCB6 plays an important role in regulating heme synthesis either by directly channeling CPgenIII to the mitochondria (Figure 1, panel a) or by indirectly regulating another step in the pathway.

The molecules and the mechanisms involved in heme transport across biological membranes to various cellular destinations

are poorly understood, partly because of the use of *in vitro* biochemical approaches, static microscopic techniques, and inappropriate genetic model systems. Given the well-established paradigm for intracellular copper trafficking pathways (11), it is likely that specific pathways also exist for transport, trafficking, sequestration, and egress of heme in cells (Figure 1, panel b). Although several parallels exist between copper and heme, the most noteworthy is that both are essential cofactors that participate in electron transfer reactions but are toxic compounds when found in excess. A major difference, however, is that nutritional copper is acquired exogenously, whereas heme is produced endogenously by a defined and regulated pathway. Although the pathway and intermediates for heme synthesis have been well defined, the handling of heme from its point of synthesis in the mitochondria to its insertion into hemoproteins remains poorly understood (Figure 1, panel b). Heme is a hydrophobic molecule and is insoluble in the aqueous cellular milieu. Free heme is toxic to biological macromolecules. How then is heme transported through the mitochondrial IMs to specific hemoproteins that reside in the cytoplasm,

peroxisomes, mitochondrial IMS, secretory pathway, and nucleus? What are the mechanisms for incorporating heme into apohemoproteins? Are these mechanisms specific to certain target proteins or to the milieu of a subcellular compartment? Humans have intracellular hemoproteins, such as hemo-, myo-, neuro-, and cytoglobins, as well as heme enzymes, including cytochrome P450s, adenylate cyclases, soluble guanylate cyclases, peroxidases, catalases, and respiratory cytochromes. These enzymes are located in different cellular organelles, and they perform diverse biological functions that depend upon heme as a cofactor. Thus, in principle, specific intracellular pathways are likely to also exist for the safe, efficient, and accurate transfer of heme from the mitochondrial IM to distinct hemoproteins that are present in various subcellular compartments (Figure 1, panel b).

Researchers have recently demonstrated that the breast cancer resistance protein (BCRP) and the feline leukemia virus subgroup C receptor (FLVCR) are potential heme exporters in developing erythroid cells and that the heme carrier protein 1 (HCP1) is the intestinal heme importer in mammals (12–14) (Figure 1, panel b). Although it is unclear why heme export would be necessary in red blood cells given the overwhelming requirement for heme in hemoglobin synthesis, these studies underscore the necessity for translocation of heme between membrane compartments.

A conceptual setback in identifying heme trafficking pathways has been the difficulty in dissociating biosynthesis from downstream trafficking events for three main reasons: (i) organisms normally make endogenous heme *via* a highly regulated pathway, (ii) defects in the heme synthesis pathway are usually lethal or have pleiotropic effects, and (iii) exogenous heme/porphyrins are poorly utilized by organisms that normally make heme. Although hemes are found in all phyla, certain prokaryotic organisms neither make heme nor contain hemo-

proteins, and the protozoa *Leishmania spp.* appear to lack seven of the eight enzymes of the heme biosynthetic pathway (2, 15). An exception is *Caenorhabditis elegans*, a free-living nematode that does not synthesize heme but ingests dietary heme to fulfill its heme auxotrophy. *C. elegans* has the repertoire of hemoproteins that humans have (16). It represents a unique genetic model system for dissecting the cellular and molecular determinants of heme homeostasis because it has a clean genetic background devoid of endogenous heme. Worms will therefore permit external control over the flux of heme and intracellular trafficking pathways, an advantage not attainable in other model systems.

The practical implications of discoveries in porphyrin transport are far-reaching. Identification of mammalian heme transporters, including HCP1, will allow the design of more bioavailable forms of iron or porphyrin-based “nutraceuticals” to deliver iron more effectively to iron-deficient populations. Identification of the mechanisms by which enzymes such as the cytochrome P450s and guanylate cyclases acquire heme will provide novel insights into modulating biologic responses to pharmaceuticals, xenobiotics, and gases such as nitric oxide. Finally, characterization of how heme is transported in organisms may lead to the discovery of parallel trafficking pathways for other tetrapyrroles such as vitamin B₁₂.

REFERENCES

1. Ponka, P. (1999) Cell biology of heme, *Am. J. Med. Sci.* 318, 241–256.
2. Panek, H., and O'Brian, M. R. (2002) A whole genome view of prokaryotic haem biosynthesis, *Microbiology* 148, 2273–2282.
3. Papenbrock, J., and Grimm, B. (2001) Regulatory network of tetrapyrrole biosynthesis—studies of intracellular signalling involved in metabolic and developmental control of plastids, *Planta* 213, 667–681.
4. Krishnamurthy, P. C., Du, G., Fukuda, Y., Sun, D., Sampath, J., Mercer, K. E., Wang, J., Sosa-Pineda, B., Murti, K. G., and Schuetz, J. D. (2006) Identification of a mammalian mitochondrial porphyrin transporter, *Nature* 443, 586–589.
5. Ferreira, G. C., Andrew, T. L., Karr, S. W., and Dailey, H. A. (1988) Organization of the terminal two enzymes of the heme biosynthetic pathway. Orientation of protoporphyrinogen oxidase and evidence for a membrane complex, *J. Biol. Chem.* 263, 3835–3839.
6. Dailey, H. A. (2002) Terminal steps of haem biosynthesis, *Biochem. Soc. Trans.* 30, 590–595.
7. Shaw, G. C., Cope, J. J., Li, L., Corson, K., Hersey, C., Ackermann, G. E., Gwynn, B., Lambert, A. J., Wingert, R. A., Traver, D., Trede, N. S., Barut, B. A., Zhou, Y., Minet, E., Donovan, A., Brownlie, A., Balzan, R., Weiss, M. J., Peters, L. L., Kaplan, J., Zon, L. I., and Paw, B. H. (2006) Mitoferrin is essential for erythroid iron assimilation, *Nature* 440, 96–100.
8. Hamza, I., Chauhan, S., Hassett, R., and O'Brian, M. R. (1998) The bacterial *irr* protein is required for coordination of heme biosynthesis with iron availability, *J. Biol. Chem.* 273, 21669–21674.
9. Rouault, T., and Klausner, R. (1997) Regulation of iron metabolism in eukaryotes, *Curr. Top. Cell Regul.* 35, 1–19.
10. Kabe, Y., Ohmori, M., Shinouchi, K., Tsuboi, Y., Hira, S., Azuma, M., Watanabe, H., Okura, I., and Handa, H. (2006) Porphyrin accumulation in mitochondria is mediated by 2-oxoglutarate carrier, *J. Biol. Chem.* 281, 31729–31735.
11. Finney, L. A., and O'Halloran, T. V. (2003) Transition metal speciation in the cell: insights from the chemistry of metal ion receptors, *Science* 300, 931–936.
12. Krishnamurthy, P., Ross, D. D., Nakanishi, T., Bailey-Dell, K., Zhou, S., Mercer, K. E., Sarkadi, B., Sorrentino, B. P., and Schuetz, J. D. (2004) The stem cell marker *Bcrp/ABCG2* enhances hypoxic cell survival through interactions with heme, *J. Biol. Chem.* 279, 24218–24225.
13. Quigley, J. G., Yang, Z., Worthington, M. T., Phillips, J. D., Sabo, K. M., Sabbath, D. E., Berg, C. L., Sassa, S., Wood, B. L., and Abkowitz, J. L. (2004) Identification of a human heme exporter that is essential for erythropoiesis, *Cell* 118, 757–766.
14. Shayeghi, M., Latunde-Dada, G. O., Oakhill, J. S., Laftah, A. H., Takeuchi, K., Halliday, N., Khan, Y., Warley, A., McCann, F. E., Hider, R. C., Frazer, D. M., Anderson, G. J., Vulpe, C. D., Simpson, R. J., and McKie, A. T. (2005) Identification of an intestinal heme transporter, *Cell* 122, 789–801.
15. Sah, J. F., Ito, H., Kolli, B. K., Peterson, D. A., Sassa, S., and Chang, K. P. (2002) Genetic rescue of *Leishmania* deficiency in porphyrin biosynthesis creates mutants suitable for analysis of cellular events in uroporphyrin and for photodynamic therapy, *J. Biol. Chem.* 277, 14902–14909.
16. Rao, A. U., Carta, L. K., Lesuisse, E., and Hamza, I. (2005) Lack of heme synthesis in a free-living eukaryote, *Proc. Natl. Acad. Sci. U.S.A.* 102, 4270–4275.

G-Protein-Coupled Receptor Trafficking: Understanding the Chemical Basis of Health and Disease

Alfredo Ulloa-Aguirre^{†,‡}, Jo Ann Janovick[†], Alfredo Leños Miranda[‡], and P. Michael Conn^{†,‡,§,*}

[†]Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, Oregon 97006, [‡]Research Unit in Reproductive Medicine, Instituto Mexicano del Seguro Social, México D. F., Mexico, [§]Departments of Physiology and Pharmacology, and Cell and Developmental Biology, Oregon Health & Science University, Portland, Oregon 97239

G-protein-coupled receptors (GPCRs) are a large and functionally diverse superfamily of plasma membrane (PM) receptors that consist of single-polypeptide chains that traverse the lipid bilayer seven times, forming characteristic transmembrane (TM) helices. After binding ligand, GPCRs interact with G proteins and transduce signals from outside the cell to the intracellular environment (1). This Perspective summarizes the emerging role of GPCR trafficking as a post-translational control mechanism, using the gonadotropin-releasing hormone (GnRH) receptor (GnRHR) as a model for evolved inefficiency in trafficking and PM expression of a receptor.

The GnRHR Is a Key Regulator in Reproduction. The GnRHR is a GPCR located in the anterior pituitary (Figure 1) that mediates responses to its ligand GnRH, a decapeptide produced by hypothalamic neurons (2, 3). The GnRHR integrates the neural system of the hypothalamus with the anterior pituitary endocrine system by producing peripheral signaling that is mediated by the pituitary gonadotropins, glycoprotein hormones that enter the peripheral circulation and regulate gonadal function (Figure 1). The GnRHR is a therapeutic target for the regulation of fertility and the treatment of several reproductive disorders (2).

Diversity in Reproductive Patterns. Although the GnRHR recognizes the same endogenous ligand in most mammals and many non-mammalian vertebrates, it also serves different regulatory patterns of reproduction. Among animals with particular specializations are dogs (which show pseudopregnancy), mice and rats (short-cycling animals), opossums (non-placental mammals that give birth from an external pouch), and guinea pigs

ABSTRACT The primary function of cell surface receptors is to recognize specific chemical signals from other substances and produce a biological response. Point mutations in cell surface receptors may result in production of misfolded proteins that are translated but do not reach their proper functional destination in the cell. Also, for some G-protein-coupled receptors, large amounts of wild-type receptor may be destroyed without arriving at the plasma membrane (PM). For the human gonadotropin-releasing hormone receptor, this “inefficiency” has resulted from strong and convergent evolutionary pressure, producing receptor molecules that are sensitive to single changes in chemical charge and are delicately balanced between expression at the PM or retention/degradation in the endoplasmic reticulum. This Perspective focuses on the evolved mechanisms that control PM expression of this receptor at this post-translational level.

*Corresponding author,
connm@ohsu.edu.

Received for review August 18, 2006
and accepted October 12, 2006.

Published online November 17, 2006

10.1021/cb600360h CCC: \$33.50

© 2006 by American Chemical Society

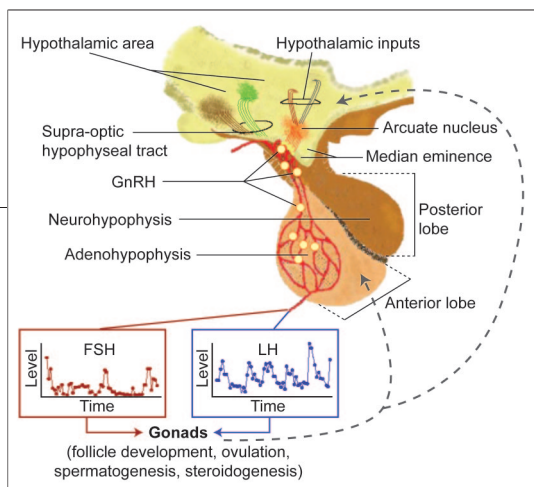


Figure 1. Functional relations of the hypothalamic–pituitary axis. GnRH is synthesized and secreted by specialized neurons located mainly in the arcuate nucleus of the medial basal hypothalamus and the preoptic area of the anterior hypothalamus. In contrast to other hypothalamic neurons that project to the posterior lobe of the pituitary (e.g., the supra-optic hypophyseal tract), GnRH-producing neurons project to many sites within the brain and also to the median eminence, terminating in an extensive plexus of boutons on the primary portal vessel, which delivers GnRH to its target cell, the gonadotrope of the adenohypophysis. The secretion and interaction of GnRH with its cognate receptor occur in a pulsatile and intermittent manner; such an episodic signaling allows the occurrence of distinct rates and patterns of synthesis and pulsatile secretion of

luteinizing hormone (LH) and follicle-stimulating hormone (FSH), whose coordinated release allows for an extremely precise control of gonadal function. These trophic hormones are responsible for stimulating the synthesis and secretion of gonadal hormones and for effecting the process of gametogenesis. The characteristics of the pulsatile release of GnRH, LH, and FSH appear to be regulated by several hypothalamic neurotransmitters (e.g., adrenergic and opioidergic regulation), as well as by the gonadal hormone environment.

(rodents with a long luteal phase). In primates, the tight regulation of reproduction is imposed by the need for efficiency in a process of long gestations, mono-ovulatory cycles, and single-offspring births.

In a broader evolutionary view of reproduction, it is interesting to compare the length of time, metabolic cost, and number of offspring in humans to fish or reptiles. In the latter, small numbers of eggs survive to hatch and a still smaller number actually survive to reproduce themselves. The metabolic investment in each egg is quite modest. Clearly, the management of the human reproductive process must be more closely regulated to protect the greater investment of metabolic energy and time commitment by making the process “work” effectively a greater proportion of the time. It is therefore interesting to consider how nature has accommodated such changes without modifying the basic components of the system.

The mammalian GnRHR type I (hereafter referred to only as GnRHR) is among the smallest members of the GPCR superfamily and bears unique structural features, including the lack of a carboxyl-terminal intracellular extension (3, 4) (Figure 2). Fish, reptiles, birds, and the primate type II GnRHR (3) do possess this carboxyl tail, whose presence is associated with differential physiological receptor regulation; when added to the mammalian GnRHR, it dramatically increases PM expression levels of this receptor (4).

Unique Chemical Features of GnRHR from Different Species Influence Cellular Trafficking. To enable agonist binding, the GnRHR needs to be at the PM. In the case of the human (h) GnRHR [and, potentially, other GPCRs (5–7)], nature has chemically modified the receptor so that the translation product is delicately balanced between routing to the PM or retention/degradation in the endoplasmic reticulum (ER) (8). This receptor is

extremely sensitive to mutations; a single change in charge can affect the balance (9).

A particular feature of primate GnRHRs is the presence of a lysine residue at position 191, which is located in the extracellular loop 2 (EL2) (Figure 3) and restricts the GnRHR PM expression (10). Non-primate mammals utilize a less-effective Glu191 in this position (or Gly191 in the opossum, all 328 amino acids), whereas rats and mice do not have this insertion at all [327 amino acids (3)] and a higher proportion of translation product of both rodent receptors is expressed at the PM (Figure 3).

Protein (Mis)folding and Intracellular Trafficking. The GnRHR is scrutinized by the quality-control system of the cell. This system employs recognition mechanisms to examine newly synthesized proteins and ensure that correctly folded proteins are placed into function (8, 11). Among the components of the “conformation-screening” mechanisms are protein chaperones of the ER; these assist in folding of nascent proteins, allowing them to achieve their “proper” conformation and preventing conformationally defective proteins from accumulating within the cell and causing disease (8, 9, 11). Because the quality-control system monitors the structural form of the protein rather than the biological function, misfolded/misrouted mutant proteins can be functionally competent proteins. Therapeutic approaches that redirect misrouted proteins to their correct destination within the cell may restore them to proper functions (12). That a misfolded protein may still behave as a functionally competent protein upon rescue will depend on how the structural defect disrupts the structure of the protein. In the case of the hGnRHR, insertion or removal of prolines or the insertion or removal of cysteines (that leads to loss or gain of cysteine bridges) may result in an uncorrectable misfolded/misrouted protein (9, 13, 14). Conversely, agents that stabilize and/or reshape the defective structure caused by mutations may still allow

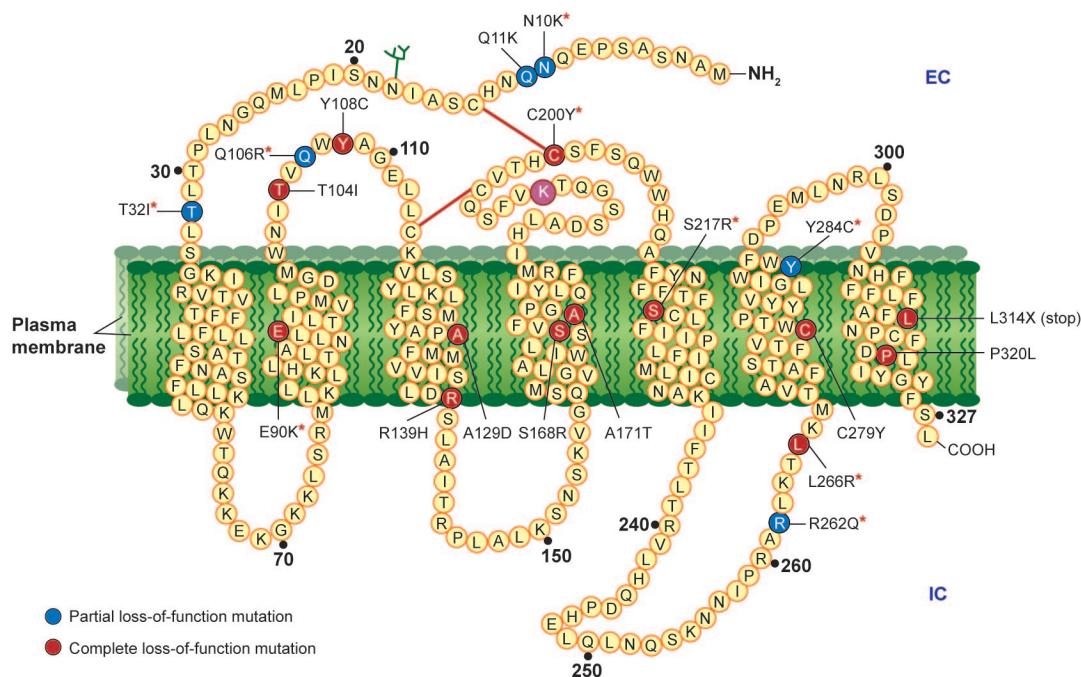


Figure 2. Sequence of the hGnRHR and location of the inactivating (loss-of-function) mutations identified to date. The GnRHR consists of a single polypeptide chain that traverses the cell surface membrane seven times, forming characteristic TM helices interconnected by alternating extracellular (EC) and intracellular (IC) loops. The circle corresponding to the lysine residue at position 191 in the EL2 is enlarged and rose-colored. Mutants with an asterisk (N10K, T32I, E90K, Q106R, C200Y, S217R, R262Q, L266R, and Y284C) were reconstructed in the corresponding mouse and rat receptor sequences and tested for function (see text).

the receptor to “escape” from proteasomal degradation (8, 12, 14–16). In this case, the rescued misfolded competent receptor may be expressed at the PM and affect function (9).

Pharmacological Chaperones (Pharmacoperones) and Misfolded GnRHR Mutant Rescue. Early studies demonstrated that the function of a mutant hGnRHR (E90K), which causes severe hypogonadotropic hypogonadism (HH) in humans (17) (Figure 2), was completely restored when the primate-specific amino acid Lys191 was deleted (10). HH is a disease characterized by decreased release of gonadotropins, leading to impaired gonadal function (9). Lys191 is absent in wild-type (WT) mouse GnRHR (mGnRHR) and rat GnRHR (rGnRHR), which route with higher efficiency to the PM (8, 18) than the human counterpart. Studies (15, 16) employing a number of pharmacoperones (low-molecular-weight molecules that enter cells and are templates for correcting misfolded mutants and restor-

ing function) revealed that E90K and most other naturally occurring mutant receptors, thought to bear defects in domains involved in receptor function, were actually misfolded proteins whose routing to the PM was compromised (9, 14–16). Two particular hGnRHR mutants involving changes in charge, S168R and S217R, were recalcitrant to rescue by different classes of pharmacoperones (9).

These findings initially suggested that these sites were unrescuable because they may involve domains important for receptor function. However, further studies demonstrated that this was not the case and that residues at positions 168 and 217 in the hGnRHR were important in regulating the position of the EL2 and the intimacy of residues Cys14 and Cys200 necessary to form a disulfide bond and a properly folded receptor in the hGnRHR (Figure 4) (20). Because of charge considerations, the unfavorable exchange of serine and arginine at the 4 and 5 TM helices likely moved the EL2 into a

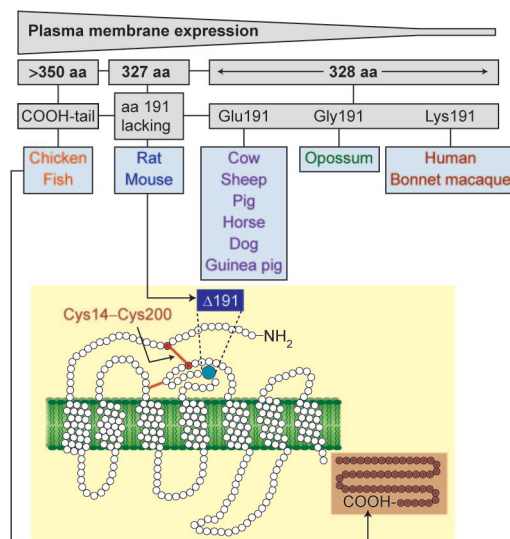


Figure 3. Evolutionary changes of the GnRHR. The GnRHR is larger in birds, fish, and reptiles; in these species, the receptor bears an intracellular carboxyl-terminal extension (dark-brown circles within the light-brown square). In mammals, the GnRHR is shorter and the carboxyl tail is absent; the blue circle represents the amino acid residue in position 191, which is frequently glutamic acid or glycine but is replaced by lysine in primates. In rat and mice GnRHR, this amino acid is absent. In most mammals, an association between cysteine residues 14 and 200 (red-filled circles) must form for proper routing of the receptor from the ER to the PM; in primates this association is formalized by a covalent bond (red line). Glutamic acid or lysine at position 191 destabilizes the association between these two cysteines. The disulfide bridge between the first and second extracellular loops (formed by cysteine residues 114 and 195 or 196 in most mammals; lower red line) is a structural feature present in almost all GPCRs known and is associated with the fundamental stability of the structure containing seven TM helices.

position from which the formation of a cysteine bridge is unlikely and the mutant with the unformed cysteine bridge cannot pass the quality-control system even in the presence of pharmacoperones.

Studies with pharmacoperones also showed that PM expression of the WT hGnRHR but not the WT rGnRHR and WT mGnRHR increased substantially upon exposure to these agents (18, 21, 22). This indicates that a large portion of the hGnRHR is normally inefficiently trafficked to the PM, retained by the quality-control system, and likely degraded.

The observation that the hGnRHR was so susceptible to alterations of single charges in the receptor structure supported the view that the human receptor is precariously balanced between retention in the ER and routing to its final destination; this is not seen in rats or mice, animals that routed the GnRHRs to the PM with higher efficiency.

Because deletion of Lys191 from the hGnRHR led to almost complete PM expression of the WT and E90K mutant receptors, we concluded that the presence of this residue was associated with routing regulation. For reasons that were buried in the physically diffuse amino acid differences between the rodent and human sequences, the simple addition of Lys191 to WT rodent GnRHRs did not decrease routing (20, 22). As discussed below, removing Lys191 from the hGnRHR obviates the need for the Cys14–Cys200 bond (20).

Mutations Leading to Misfolded hGnRHRs Have Less Impact on Trafficking in Rat and Mouse WT GnRHRs.

Reconstruction of nine naturally occurring mutations leading to misfolded GnRHRs in humans (Figure 2) in the corresponding mouse and rat receptor sequences revealed that only the E90K, S217R (G216R and S216R in the mouse and rat GnRHR, respectively), L266R, and Y284C (L265R and Y283C in both rodent GnRHR sequences) substitutions significantly impact rat and mouse GnRHR function (20, 22). Notably, pharmacoperone treatment restored function of all but the G216R and S216R mutants (which corresponded to the pharmacoperone-insensitive S217R mutant in the hGnRHR), an indication that the effect of these mutations was to cause severe protein misfolding and loss of the ability to translocate the protein to the PM (22). In

KEYWORDS

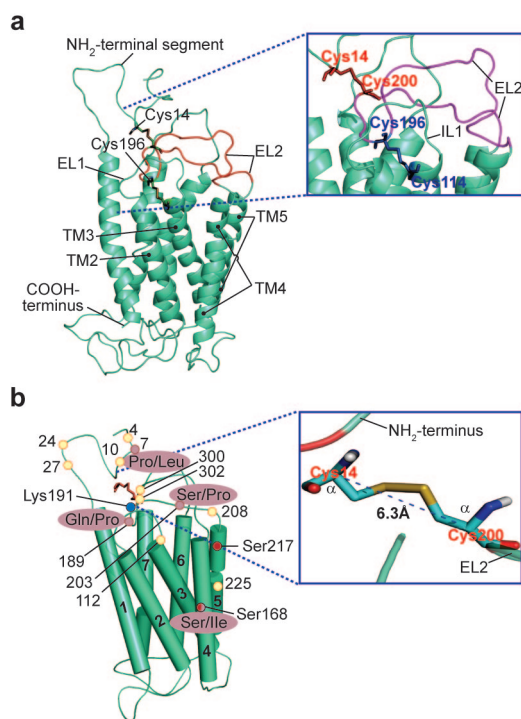
Disulfide bridge: A covalent linkage formed between two sulfhydryl groups on cysteines.

Evolution: A change in the properties of populations of organisms that transcend the lifetime of a single individual.

Gonadotropin-releasing hormone receptor: The receptor for the hypothalamic peptide hormone gonadotropin-releasing hormone that stimulates the release of gonadotropins.

G-protein-coupled receptors: A large and functionally diverse superfamily of plasma membrane receptors that consist of single-polypeptide chains that traverse the lipid bilayer seven times, forming characteristic transmembrane helices.

Figure 4. Structures of hGnRHR a) Predicted structure of the hGnRHR based on homology modeling with the structure of bovine rhodopsin (19). The coiled structures represent the antiparallel α -helices of transmembrane domains 1–7 connected by the extra- and intracellular loops of the receptor (curved cords). Intermolecular associations between Cys14 (at the NH₂-terminal segment) and Cys200 (at the EL2, colored in red) and between Cys114 (at the COOH-terminal end of the EL1) and Cys196 (at the EL2) are shown (colored sticks). Top right: magnification of the upper-third portion of the receptor showing the Cys14–Cys200 (red sticks) and Cys114–Cys196 (blue sticks) disulfide bridges. b) A model of the hGnRHR showing the seven TM helices displayed as rods. hGnRHR mutants at Ser168 and Ser217 (red circles) cannot be rescued by pharmacoperone treatment. The approximate location of amino acid residues that differ between the human and rat GnRHR and that represent thermodynamically unfavored substitutions are represented by small lavender circles (human/rat in enlarged roseate circles) (amino acids 7, 168, and 203; 202 in rat and mouse); amino acid 189 is included for its physical proximity to the Cys14–Cys200 bridge. Amino acids 7, 168, and 189 frequently coevolved with the appearance of the “extra” amino acid in position 191 (lysine in primates, glutamic acid or glycine in all non-rat/mouse mammalian species sequenced to date). The locations of some other residues that impact on the net levels of receptor PM expression and allow the effect of Lys191 are shown in yellow. Positions of these residues are simply pictorial. Residues located in the NH₂ extracellular segment and in the EL2 as well as sequences flanking this loop (*i.e.*, within TM helices 4 and 5) and those that abut on that area (ELs 1 and 3) presumably control the destabilizing role of Lys191 (blue circle) on the formation of the Cys14–Cys200 bridge as disclosed by mutagenesis experiments (21). This bridge is magnified in the right figure; the covalent bond between the two sulfur atoms is shown by gold sticks. The non-optimized models shown in this figure were generated with the molecular visualization program PyMol (DeLano Scientific, San Francisco, CA) and kindly provided by Angel Piñeiro and Eduardo Jardón-Valadez from the Faculty of Chemistry of the National University of Mexico, Mexico City, Mexico.



addition, the mGnRHR appeared to be less forgiving than the rGnRHR to substitutions, because the mouse mutations more greatly reduced PM expression of the mutant receptors compared with the rat mutant counterparts (22). These observations indicated that these rodent GnRHRs appear to be more tolerant of mutation than is the hGnRHR and that the small number of semi- and non-conservative differences between the mouse and rat GnRHRs (at positions 11, 24, 160, and 216) might have a physiological influence on receptor function. In fact, creation of “rat-like” mGnRHR misfolded mutants indicated that a modest amino acid change, that is, substitution of glycine in position 216 in the mGnRHR with serine (present in the rGnRHR), resulted in markedly increased PM expression of the mGnRHR-E90K sequence; modification of the corresponding sequence in the rGnRHR, to make it more “mouse-like”, was associated with a 2-fold decrease of expression (22). Further substitutions exchanging proline with glycine in position 11 and isoleucine with threonine in position 24 or 160 resulted in a substantial decrease in mGnRHR-E90K function, which was more prominent

whenever Gly216 was present. Ser216 in the rGnRHR sequence increased the efficiency of routing to the PM of rat mutants that are otherwise misrouted as disclosed by pharmacoperone treatment.

Mutant GnRHRs Exert Dominant-Negative Effects Due to an Influence on Trafficking. Receptor oligomerization and interactions with accessory proteins are well-documented features of GPCR function (23, 24). In some cases, intracellular association of receptors as homo- or heterodimers led to either cell surface targeting (a dominant-positive effect) or the intracellular retention of the complex (a dominant-negative effect) (6, 9, 24). In the case of the hGnRHR, misfolded hGnRHR mutants may form intracellular complexes at the ER, affecting proper delivery and trafficking of the receptor to the cell surface (8, 21, 25). Cotransfection experiments revealed that this dominant-negative effect of mutants on their WT counterpart resulted from a physical interaction between these molecules (21).

Despite the high homology between rat, mouse, and human GnRHRs, pharmacoperone-sensitive rat and mouse misfolded receptors exhibited different

dominant-negative actions on their corresponding receptors: (i) The WT rGnRHR escapes the dominant-negative effects of mutant rGnRHRs (e.g., rGnRHR-E90K) but not those from human and mouse mutants, thus indicating that this receptor retains the ability to oligomerize. (ii) Mutant mGnRHRs are more effective in dominant-negative actions on both rodent receptors. (iii) Mutant hGnRHRs are more effective, as dominant-negative regulators, on the human WT receptor than on rat and mouse WT receptors, which lack Lys191 (22). The effect of Ser216 in the rGnRHR appeared to be important for the loss of dominant-negative effects of mutant rGnRHRs on the WT counterpart; experiments showed that the Gly216 (mouse) → Ser216 (rat) substitution in the mGnRHR mutants sequence resulted in loss of dominant-negative effect of mouse mutants on the rat receptor species (22). In the hGnRHR, substitution of Ser217 with glycine led to a modest reduction in receptor function and concomitantly to loss of the dominant-negative effect of the mutant on the WT receptor, whereas removal of Lys191 from the S217G mutant allowed functional recovery. These findings implicate the absence of Lys191 and presence of Ser216/217 as determinants for PM expression and loss of the dominant-negative effect of mutant GnRHRs. In this scenario, the counteracting effect of Ser217 on the dominant-negative action of mutant hGnRHRs may be mitigated by the insertion of Lys191, because deletion of this residue from the hGnRHR sequence led to loss of the dominant-negative effect of misfolded mutants on WT receptor expression (18, 22).

Cys14–Cys200 Bridge Formation and Tolerance to Mutations. The amino acid sequence of the GnRHR predicts two disulfide bridges at extracellular regions (Cys14–Cys195/196 and Cys14–Cys199/200;

Figure 4). The first bridge is a structural feature present in almost all known GPCRs and is associated with the stability of the heptahelical structure (1); mutation of the cysteine residue at either end of this bridge in the GnRHR resulted in a complete loss of activity, which could not be rescued by pharmacoperones (20). The functional significance of the Cys14–

Cys199/200 bridge differed depending on the receptor species: (i) It was not essential for optimal function of the rGnRHR because replacement of the cysteine residues at either end of the bridge does not affect agonist-stimulated intracellular signaling. (ii) In the mGnRHR, breakage of this bridge results in ~50% decrease in receptor function. (iii) For the hGnRHR, formation of this bridge was an absolute requirement for correct routing and PM expression of the receptor because bridge-breaking mutants exhibit either no activity or marginal activity (14, 22). Further, exposure of these mutants to pharmacoperones normalizes receptor function, an indication that the absence of the bridge resulted in a misfolded protein.

What are the chemical determinants that lead to the requirement of the Cys14–Cys200 bridge in folding of the hGnRHR? Removal of Lys191 from misfolded hGnRHR mutants (including Cys14 and Cys200 mutants) led to functional rescue of the altered receptors, an indication that the association between Cys14 and Cys200 may be potentially disrupted or diminished by the presence of Lys191 (10, 20). In contrast with the majority of mutations resulting in HH in humans (Figure 2), the effect of Lys191 is not wholly an effect of charge because replacement by alanine, glutamic acid, or glutamine also produced inefficient PM expression of human sequence compared with the hGnRHR-des-Lys191 variant. The observation that removal of Lys191 obviates the need for this bridge in the hGnRHR and that inserting Lys191 alone into the rat or mouse sequence (>88% homologous with the human sequence) did not impact the requirement for this bridge between the NH₂-terminal and the EL2 led to the search for additional components of the requisite motif.

This problem was approached by locating the thermodynamically unfavored changes among GnRHR sequences for various animal species and identifying those amino acid residues that frequently coevolved with the appearance of the “extra” amino acid in position 191 (lysine in primates, glutamic acid or glycine in all non-rat/mouse mammalian species sequenced to date; Figure 3) or that were proximal to it and to the Cys14–Cys200 bridge (Figure 4). When the molecule as a whole is considered, the modifications associated with orientation of EL2 (positions 7, 168, 189, and 202/203) all involve the gain or loss of proline or serine residues. Proline formed a five-membered nitrogen-containing ring, a feature found in very tight turns in

KEYWORDS

Intracellular trafficking: The passage of newly synthesized proteins from the ER to their final destination within the cell.

Mutant receptors: Protein molecules with an altered amino acid sequence due to changes in the nucleotide sequence of their corresponding gene.

Pharmacological chaperones: Low-molecular-weight molecules that enter cells and are templates for correcting misfolded mutants.

Pharmacoperone: A pharmacological chaperone.

Protein folding: The process whereby a protein molecule assumes its 3D shape.

protein structures. Serine, with a slightly polar nature, small size, and propensity of the side-chain hydroxyl oxygen to H-bond with the protein backbone, also was found in association with tight turns of the protein structure. The rest of the motifs were identified by making guesses based on the physical relation between amino acids in the 3D state of the receptor molecule. With this information, human receptors that were more “rat-like” were constructed and tested; these expressed at the higher levels associated with rat receptor and lacked the requirement for the Cys14–Cys200 bridge, which is another structural feature of the rGnRHR (20). The spatial alignment was, therefore, quite important, because these two residues must be sufficiently close to each other to allow bridge formation. When the bridge formed, the hGnRHR was recognized by the quality-control system as correctly folded. On the contrary, when it did not form, it was viewed as defective and was retained in the ER. Thus, the presence of Lys191 limited the number of hGnRHR molecules that were exported from the ER to the PM.

Inefficient WT hGnRHR Trafficking: A Post-Translational Regulatory Mechanism? It seems that the cell is exploiting the insertion of Lys191 in the hGnRHR for controlling routing in normally functioning cells. Why waste half the amount of newly synthesized receptor molecules? Several possibilities exist. First, having a pool of receptors that can be recruited during very demanding conditions (*e.g.*, during the mid-follicular phase of the menstrual cycle) may be very important; the increased synthesis of the receptor at these times would allow the receptor to escape from proteosomal

degradation. Second, by regulating the amount of receptors expressed, the pituitary may more precisely modulate gonadotropin release and maturation of a single oocyte during each ovulatory cycle. In this scenario, the complexity of reproduction appears to have evolved as the investment in creating a single offspring increased. A look at GnRHR in animals whose sequences appeared odd revealed that several species display marked differences in their reproductive patterns when compared with their evolutionarily close relatives. Among rodents (animals with large litters), only the guinea pig is known to have added an amino acid (glutamic acid) at position 191 of the GnRHR; this particular rodent, a hystricomorph that diverged very early in rodent evolution, exhibits a long luteal phase, which is a primate characteristic. Most non-rodent mammals also contain glutamic acid at this position; this suggests that loss of an amino acid in the homologous position may represent a specialization associated with very short reproductive cycles like those present in rats and mice. In the opossum (a non-placental mammal that places fetuses in a marsupium), the presence of the uncharged glycine at position 191 may reflect the divergence of this group and the specialization needed for this unique form of reproduction.

The large number of biological systems that appear to rely on assessment of fidelity of protein structure (26) and folding suggests that regulation at this level may prove to be more common than presently appreciated.

Acknowledgment: This work is supported by CONACYT grant 45991-M; NIH grants HD-19899, RR-00163, HD-18185, TW/HD-00668.

REFERENCES

1. Ulloa-Aguirre, A., and Conn, P. M. (1998) G protein-coupled receptors and the G protein family, in *Handbook of Physiology-Endocrinology*, (Conn P. M., Ed.) Section 7, Cellular Endocrinology, pp 87–124, Oxford University Press: New York.
2. Conn, P. M., and Crowley, F. (1991) Gonadotropin-releasing hormone and its analogues, *N. Engl. J. Med.* 324, 93–103.
3. Millar, R. P., Lu, Z.-L., Pawson, A. J., Flanagan, C. A., Morgan, K., and Maudsley, S. R. (2004) Gonadotropin-releasing hormone receptors, *Endocr. Rev.* 25, 235–275.
4. Lin, X., Janovick, J. A., Brothers, P. S., Blömenrohr, M., Bogerd, J., and Conn, P. M. (1998) Addition of catfish gonadotropin-releasing hormone (GnRH) receptor intracellular carboxyl-terminal tail to rat GnRH receptor alters receptor expression and regulation, *Mol. Endocrinol.* 12, 161–171.
5. Petäjä-Repo, U. E., Hogue, M., Laperrière, A., Walkers, P., and Bouverier, M. (2000) Export from the endoplasmic reticulum represents the limiting step in the maturation and cell surface expression of the human δ opioid receptor, *J. Biol. Chem.* 275, 13727–13736.
6. Pfeiffer, M., Koch, T., Schroder, H., Klutzny, M., Kirscht, S., Kreienkamp, H.-J., Höllt, V., and Schulz, S. (2001) Homo- and heterodimerization of somatostatin receptor subtypes. Inactivation of ss_{t_3} receptor function by heterodimerization with $ss_{t_{2A}}$, *J. Biol. Chem.* 276, 14027–14036.
7. Pietilä, E. M., Tuusa, J. T., Apaja, P. M., Aatsinko, J. T., Hakalahti, A. E., Rajaniemi, H. J., and Petäjä-Repo, U. E. (2005) Inefficient maturation of the rat luteinizing hormone receptor. A putative way to regulate receptor numbers at the cell surface, *J. Biol. Chem.* 280, 26622–26629.
8. Ulloa-Aguirre, A., Janovick, J. A., Brothers, S. P., and Conn, P. M. (2004) Pharmacological rescue of conformationally-defective proteins. Implications for the treatment of human disease, *Traffic* 5, 821–837.
9. Ulloa-Aguirre, A., Janovick, J. A., Leañes-Miranda, A., and Conn, P. M. (2004) Misrouted cell surface GnRH receptors as a disease aetiology for congenital isolated hypogonadotropic hypogonadism, *Hum. Reprod. Update* 10, 177–192.

10. Maya-Núñez, G., Janovick, J. A., Ulloa-Aguirre, A., Söderlund, D., Conn, P. M., and Méndez, J. P. (2002) Molecular basis of hypogonadotropic hypogonadism: restoration of mutant (E³⁹K) GnRH receptor function by a deletion at a distant site, *J. Clin. Endocrinol. Metab.* **87**, 2144–2149.
11. Ellgaard, L. (1999) Setting the standards: quality control in the secretory pathway, *Science* **286**, 1882–1888.
12. Bemier, V., Lagacé, M., Bichet, D. G., and Bouvier, M. (2004) Pharmacological chaperones: potential treatment for conformational diseases, *Trends Endocrinol. Metab.* **15**, 222–228.
13. Pask, A. J., Kanasaki, H., Kaiser, U. B., Conn, P. M., Janovick, J. A., Stockton, D. W., Hess, D. L., Justice, M. J., and Behringer, R. R. (2005) A novel mouse model of hypogonadotropic hypogonadism: *N*-ethyl-*N*-nitrosourea-induced gonadotropin-releasing hormone receptor gene mutation, *Mol. Endocrinol.* **19**, 972–981.
14. Janovick, J. A., Maya-Nunez, G., and Conn, P. M. (2002) Rescue of hypogonadotropic hypogonadism-causing and manufactured GnRH receptor mutants by a specific protein-folding template: misrouted protein as novel disease etiology and therapeutics target, *J. Clin. Endocrinol. Metab.* **87**, 3255–3262.
15. Leaños-Miranda, A., Janovick, J. A., and Conn, P. M. (2002) Receptor-misrouting: an unexpectedly prevalent and rescuable etiology in GnRHR-mediated hypogonadotropic hypogonadism, *J. Clin. Endocrinol. Metab.* **87**, 4825–4828.
16. Janovick, J. A., Goulet, M., Bush, E., Greer, J., Wettlaufer, D. G., and Conn, P. M. (2003) Structure-activity relations of successful pharmacologic chaperones for rescue of naturally occurring and manufactured mutants of the gonadotropin-releasing hormone receptor, *J. Pharmacol. Exp. Ther.* **305**, 608–614.
17. Söderlund, D., Canto, P., de la Chesnaye, E., Ulloa-Aguirre, A., and Méndez, J. P. (2001) A novel homozygous mutation in the second transmembrane domain of the gonadotrophin releasing hormone receptor gene, *Clin. Endocrinol. (Oxford)* **54**, 493–498.
18. Janovick, J. A., Ulloa-Aguirre, A., and Conn, P. M. (2003) Evolved regulation of gonadotropin-releasing hormone receptor cell surface expression, *Endocrine* **22**, 317–327.
19. Palczewski, K., Kumasaka, T., Hori, T., Nehnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: a G protein-coupled receptor, *Science* **289**, 739–745.
20. Janovick, J. A., Knollman, P. E., Brothers, S. P., Ayala-Yáñez, R., Aziz, A. S., and Conn, P. M. (2006) Regulation of G protein-receptor trafficking by inefficient plasma membrane expression. Molecular basis of an evolved strategy, *J. Biol. Chem.* **281**, 8417–8425.
21. Brothers, S. P., Comea, A., Janovick, J. A., and Conn, P. M. (2004) Human 'loss-of-function' GnRH receptor mutants retain wild type receptors in the endoplasmic reticulum: molecular basis of the dominant-negative effect, *Mol. Endocrinol.* **18**, 1787–1797.
22. Knollman, P. E., Janovick, J. A., Brothers, S. P., and Conn, P. M. (2005) Parallel regulation of membrane trafficking and dominant-negative effects by misrouted gonadotropin-releasing hormone receptor mutants, *J. Biol. Chem.* **280**, 24506–24514.
23. Comea, A., Janovick, J. A., Maya-Núñez, G., and Conn, P. M. (2001) Gonadotropin-releasing hormone receptor microaggregation. Rate of monitored by fluorescence resonance energy transfer, *J. Biol. Chem.* **276**, 2153–2158.
24. Angers, S., Salahpour, A., and Bouvier, M. (2002) Dimerization: an emerging concept for G protein-coupled receptor ontogeny and function, *Annu. Rev. Pharmacol. Toxicol.* **42**, 409–435.
25. Leaños-Miranda, A., Ulloa-Aguirre, A., Ji, T. H., Janovick, J. A., and Conn, P. M. (2003) Dominant-negative action of disease-causing gonadotropin-releasing hormone receptor (GnRHR) mutants: a trait that potentially coevolved with decreased plasma membrane expression of GnRHR in humans, *J. Clin. Endocrinol. Metab.* **88**, 3360–3367.
26. Castro-Fernandez, C., Maya-Nunez, G., and Conn, P. M. (2005) Beyond the signal sequence: protein routing in health and disease, *Endocr. Rev.* **26**, 479–503.

Converting Inactive Peptides into Potent Transcriptional Activators

Jenifer K. Lum[†], Chinmay Y. Majmudar[‡], Aseem Z. Ansari^{§,*}, and Anna K. Mapp^{†,*}

Departments of [†]Medicinal Chemistry and [‡]Chemistry, University of Michigan, 930 N. University Avenue, Ann Arbor, Michigan 48109-1055, and [§]Department of Biochemistry and the Genome Center, University of Wisconsin, Madison, Wisconsin 53706

Transcriptional activators fulfill a critical function in gene regulation, assisting in the recruitment of coactivators and RNA polymerase II proximal to genes in preparation for activated transcription (Figure 1) (1). An increasing number of malfunctioning transcription factors have been linked to aberrant gene expression and disease pathogenesis (2). Therefore, significant efforts have been directed toward the development of artificial activators that control gene expression in a specific and potent manner equivalent to that of the natural protein counterparts (3, 4). It is the latter characteristic, potency, that has proven to be the greatest challenge in the construction of these designer molecules (3, 4). The potency of an activator is primarily dictated by the activation domain (AD), the module responsible for coactivator binding and transcriptional machinery assembly. Only rarely are artificial ADs able to rival the cellular potency of their natural counterparts; this is in part due to our incomplete understanding of the mechanistic details of AD function (4).

Natural ADs are unstructured in solution and have a characteristic “stickiness” that results in a promiscuous binding profile; they have been shown *in vitro* to interact with both transcriptionally relevant and unrelated proteins with moderate affinity (5, 6). As part of a complex signaling network, there are masking proteins (mp) that shield activation domains from non-productive

binding interactions until they are required for gene expression (Figure 1) (7, 8). An excellent example of this is the mouse homolog of double minute 2 (hDM2), a masking protein that regulates p53 function by discriminating p53-coactivator interactions and playing a role in p53 turnover (9). In contrast, artificial ADs generally operate outside of the endogenous regulatory pathways and lack masking partners or interactions. Therefore, artificial ADs are typically unable to utilize all of the resources of the cellular machinery that contribute to functional potency (7–11). Recently we identified an exception to this trend in artificial AD XL_γ; this hydrophobic peptide functions as a robust transcriptional activator when a masking interaction is available but poorly when it is not (12).

From a previously described screen to identify ligands for the transcriptional machinery protein Med15(Gal11), we identified a number of eight-residue peptides that did not function as activation domains in cells despite exhibiting micromolar dissociation constants for the transcriptional machinery (13, 14). Several of these inactive peptides bore an excess of hydrophobic amino acids in the region that had been randomized in the screen (Figure 2, panel a). Given the probable unstructured nature of these short peptides in the absence of a target binding protein, it appeared likely that non-productive binding interactions and proteolysis were contributing to the poor activity levels (11, 15, 16). Because of the

ABSTRACT Significant efforts have been devoted to the development of artificial transcriptional activators for use as mechanistic tools, as therapeutic agents, and for biomanufacturing applications. One of the primary challenges has been the development of artificial activators that exhibit potency in cells comparable to that of endogenous activators; the vast majority function only moderately in the cellular context. Here we demonstrate that the superimposition of two distinct binding modes, a masking interaction and an interaction with the transcriptional machinery, has a profoundly positive effect on the cellular activity of artificial activators, with up to 600-fold enhancement observed. Incorporation of this feature into future generations of small molecule transcriptional activators should increase their nuclear uptake and facilitate their accessibility to their target proteins, thus significantly augmenting both their activity and utility.

*Corresponding authors,
ansari@biochem.wisc.edu,
amapp@umich.edu

Received for review August 21, 2006
and accepted October 9, 2006.

Published online November 10, 2006

10.1021/cb600363n CCC: \$33.50

© 2006 by American Chemical Society

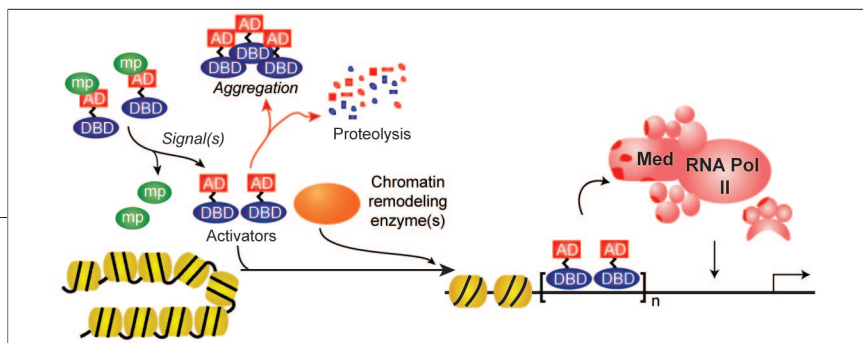


Figure 1. Cellular mechanisms that control transcriptional activator function. Activators have two essential domains: a DBD and an AD (1). Factors that contribute to the pool of activator proteins that are available to regulate transcription include mp interactions, aggregation, target/coactivator binding, and proteolysis (8–11, 33).

importance of masking interactions to the function of endogenous transcriptional activators and to XL_n, we hypothesized that the incorporation of an intramolecular binding interaction to provide structure and/or mask the hydrophobic surfaces would thus increase the potency of the peptides as activation domains.

To investigate this hypothesis, we chose Gal4(1–100) as the DNA binding domain (DBD) because it has at least three binding surfaces known to interact with hydrophobic partners (12, 17–19). Initially all Gal4(1–100) + ligand combinations were screened for activity *via* a β -galactosidase plate assay in yeast, and ligands 23, 31, 34, 35, 36, and 37 functioned as activation domains in this context. A more quantitative liquid β -galactosidase assay revealed that the activity of several of the ligands increased significantly (Figure 2, panel b). Ligand 36 (ALWFFPSE), for example, showed only 1.2-fold activity when attached to LexA but upon attachment to Gal4(1–100) was transformed into a robust transcriptional activator with \sim 600-fold levels of up-regulation observed. Ligand hydrophobicity appears to be a key determinant of the functional increase. The four ADs demonstrating the greatest increase in activity (23, 31, 36, and 37) also have the largest mean calculated hydrophobicity (Figure 2, panel a) (20).

Direct attachment of the ligands to the dimerization domain of Gal4 is also important to maintain the increased activity observed with 23, 31, 36 and 37; the use of Gal4(1–147), for example, as a DBD produces non-functional transcriptional regulators (0.9–1.2-fold activity; Figure 2, panel b). In this case, the structure of Gal4(1–147) may differ from Gal4(1–100), with a portion of the hydrophobic binding surfaces in the

Gal4(52–100) region obscured (21). The difference in activation between the Gal4(1–100) and the Gal4(1–147) constructs is not due to lower expression levels of the Gal4(1–147) proteins (Figure 2, panel c). Western blots of the Gal4(1–147) ligand fusions (Figure 2, panel c) demonstrate that all are present at similar levels regardless of the activation domain attached. In contrast, Westerns blots of the Gal4(1–100) fusion proteins are more variable with the hydrophobic ligands (23, 31, 36, and 37) present at significantly lower levels compared to the control activation domains 28 (weak activator) and VP2 (strong activator), neither of which show an increase in activity when attached to Gal4(1–100). In addition, a time course experiment with ligand 36 revealed a degradation profile that varies with cell density (Supplementary Figure 1). Taken together, these results indicate that Gal4(1–100) has an effect on both the activity and stability of the most potent peptides.

To directly probe the interaction between the ligands and Gal4(1–100), the affinities for DNA-bound Gal4(1–100) of both transcriptionally active and inactive ligands were assessed using fluorescence polarization. For this purpose, ligands were individually synthesized and labeled with fluorescein (see Methods). Ligands 36 and 23 show the greatest activity enhancement when attached to Gal4(1–100) and also interact most tightly with that protein fragment (K_D 's of 2.6 ± 0.1 and $3.1 \pm 0.1 \mu\text{M}$, respectively), with ligand 31 exhibiting a 2- to 3-fold higher K_D corresponding to its smaller activity enhancement ($7.3 \pm 0.2 \mu\text{M}$). Ligand 37, which shows the smallest activity enhancement of the four, also binds, but as a result of the propensity of Gal4(1–100) to aggregate at concentrations above 15–

20 μM , a complete binding curve could not be obtained. In contrast, ligand 28 and VP2, neither of which exhibit an activity enhancement when part of a Gal4(1–100) fusion, do not detectably bind (Supplementary Figure 2). Similar binding affinity trends were observed with the peptides that DNA binding is not required for the Gal4(1–100) interaction. Taken together, the results are consistent with a Gal4(1–100) binding interaction playing a role in enhancing transcription mediated by the ligands.

The binding surface within Gal4(1–100) was characterized through mutagenesis experiments in which amino acid residues in the dimerization region of Gal4(1–100) (Gal4_{dd}, residues 52–100) were replaced with alanine and the activities of the resulting constructs were assessed in quantitative β -galactosidase assays. Substitution at three residues (76, 77, and 84) resulted in a dramatic decrease ($\geq 80\%$) in activity (Figure 3, panel a). Additional residues (81, 82, and 90) had a more moderate effect on ALWFFPSE (ligand 36) function (Supplementary Figure 3). The residues that have the greatest impact on ALWFFPSE function define a binding interface at the second helix and first loop of the dimerization domain, a conformationally mobile and solvent-exposed region of the structure. In addition, these same mutations in Gal4_{dd} had no effect on the activity of a control AD, VP2, that shows no enhanced activity when attached to Gal4(1–100) (Supplementary Figure 4). Substitution at residues 69, 70, and 75 increased activity 3- to 9-fold (Figure 3, panel a); given their position relative to the binding site, these alanine substitutions likely increase the accessibility of the binding site. Substitution of residues 97 and 89 similarly produced an increase in activity. In this instance, attenuated dimerization of helix 3 may facilitate interaction of the activation peptide with the loop 1 binding site. In addition, the NMR structure of Gal4_{dd} reveals that residues 97–100 are

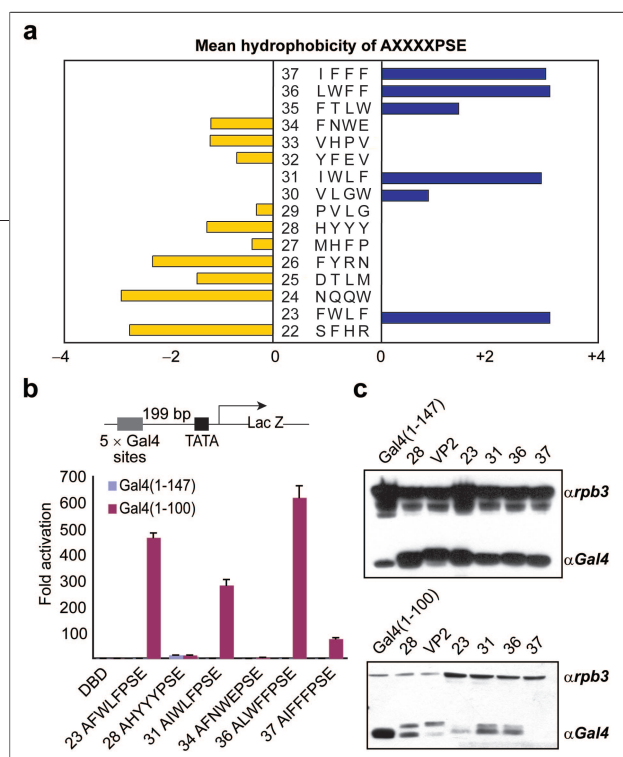


Figure 2. The DBD affects the potency of transcriptional machinery ligands. **a**) In a previously reported screen, ligands for the transcriptional machinery component Med15(Gal11) were identified from libraries of 8-residue peptides containing four randomized positions and four constant residues (alanine, proline/leucine, serine, glutamic acid) (13). The mean hydrophobicity of each ligand was calculated using a hydrophobicity calculator (CCS scale) (20). **b**) Plasmids encoding the ligands listed in panel a fused to either the Gal4(1–100) or Gal4(1–147) DBD were prepared and the resulting constructs tested for their ability to activate transcription in *Saccharomyces cerevisiae* using a β -galactosidase assay. Activity is reported as average fold activation relative to DBD alone from at least quadruplicate measurements. The error is reported as standard deviation of the mean (SDOM). **c**) Extracts of yeast cells expressing the various activator fusions shown in panel b were subjected to Western blot analysis using an antibody towards Gal4. Rpb3 was used as a loading control to facilitate lane to lane comparisons. See Supporting Information for more details.

largely unstructured and thus may provide needed flexibility such that key residues of the activation peptide can reach the binding site (17). Leu77 is ~ 25 Å from Phe97, and when extended, Gal4(97–100)+SS+ALWFFPESE is estimated to be ~ 35 Å long, within range of the proposed binding site.

In addition to the dimerization region of Gal4(1–100), the effects of alanine substitution on the activation domain (ligand 36) were also assessed (Figure 3, panel b). As noted earlier, the four most active ligands (23, 31, 36, and 37) have hydrophobic residues at the four positions randomized in the original library. A tryptophan in position 3 is present in the ligands that have the highest activity (36, 23, and 31), suggesting that a

acid) are present in all of the peptide ligands as a result of the design of the original combinatorial peptide library. Of these residues, replacement of proline with alanine had the most dramatic effect, with a 40% decrease in activity (Figure 3, panel c). The primary role of these residues is thus as a linker or spacer, facilitating interaction of the core hydrophobic amino acids with Gal4_{dd}.

In addition to the ligands described above, many endogenous and designer activation domains also contain a preponderance of hydrophobic residues and are unstructured in the absence of a protein binding partner (25). We thus investigated whether our design strategy was generalizable such that cellular activity enhance-

ments similar to those observed with the earlier ligands could be obtained by the use of a DBD containing a hydrophobic binding interface. We first examined several relatively short, non-natural peptide ADs (Figure 4, panel a). KBP1.66 and KBP2.2 were isolated from a phage display screen against the mammalian co-activator CREB binding protein (22). AH is designed to generally mimic the amphipathic helix motif found in most natural activators (23). These ligands show no direct sequence overlap with natural activators and thus are unlikely to interact with endogenous masking proteins that regulate stability and activity. Correspondingly, this peptide bound Gal4(1–100) more tightly than 37 (K_D 3.9 ± 0.1 μ M). However, switching the asparagine and tryptophan residues at the third and fourth positions of ligand 34 (AFNWEPESE) did not increase activity, again indicating that hydrophobicity in the core residues is critical for the Gal4_{dd} interaction. Four residues (alanine, proline, serine, and glutamic

acid) are present in all of the peptide ligands as a result of the design of the original combinatorial peptide library. Of these residues, replacement of proline with alanine had the most dramatic effect, with a 40% decrease in activity (Figure 3, panel c). The primary role of these residues is thus as a linker or spacer, facilitating interaction of the core hydrophobic amino acids with Gal4_{dd}. In addition to the ligands described above, many endogenous and designer activation domains also contain a preponderance of hydrophobic residues and are unstructured in the absence of a protein binding partner (25). We thus investigated whether our design strategy was generalizable such that cellular activity enhance-

ments similar to those observed with the earlier ligands could be obtained by the use of a DBD containing a hydrophobic binding interface. We first examined several relatively short, non-natural peptide ADs (Figure 4, panel a). KBP1.66 and KBP2.2 were isolated from a phage display screen against the mammalian co-activator CREB binding protein (22). AH is designed to generally mimic the amphipathic helix motif found in most natural activators (23). These ligands show no direct sequence overlap with natural activators and thus are unlikely to interact with endogenous masking proteins that regulate stability and activity. Correspondingly, this peptide bound Gal4(1–100) more tightly than 37 (K_D 3.9 ± 0.1 μ M). However, switching the asparagine and tryptophan residues at the third and fourth positions of ligand 34 (AFNWEPESE) did not increase activity, again indicating that hydrophobicity in the core residues is critical for the Gal4_{dd} interaction. Four residues (alanine, proline, serine, and glutamic

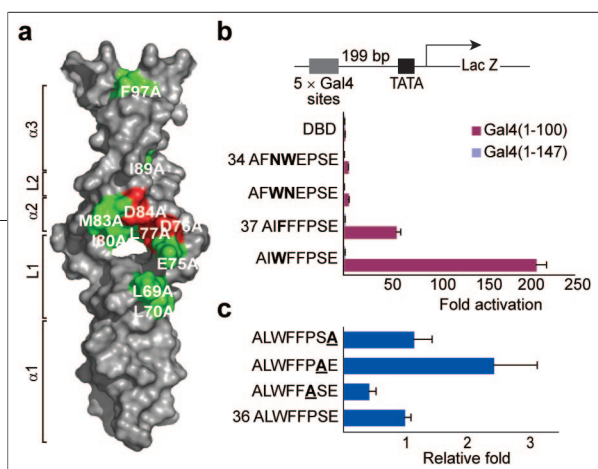


Figure 3. Identification of Gal4 and ligand residues contributing to the activity enhancement. **a)** Surface rendering of the NMR structure of the Gal4(50–100) dimer with key structural elements denoted (17). The residues in Gal4_{dd} that upon replacement with alanine increase the activity of ALWFFPSE (36) are shaded in green, while the residues that significantly decrease the activity are indicated in red. The mutagenesis experiments were carried out with fusions containing LexA as a DBD attached to Gal4(40–100)+ligand 36. (Supplementary Figure 3). **b)** Substitution at positions 3 and 4 was carried out to assess the relative importance of hydrophobic residues at those positions. Phenylalanine in the third position of ligand 37

(AIWFFPSE) was changed to tryptophan to generate peptide AIWFFPSE (37m). The asparagine in ligand 34 AFWNEPSE was switched with the adjacent tryptophan to generate peptide AFWNEPSE (34m). The altered sequences were fused to either Gal4(1–100) or Gal4(1–147) and tested in the reporter assay as in Figure 2. The last three residues of peptide 36 ALWFFPSE were mutated to alanine and tested as described above. The activity is reported as average relative fold activation generated by normalizing the activity of each mutant to ligand 36. The results in all panels are derived from quadruplicate measurements with the indicated error (SDOM).

part of a Gal4(1–100) fusion protein; this is significant given the frequency with which VP16-derived ADs are used to construct artificial activators for a wide range of applications (3, 4). Similar to VP2, Gal4 did not increase in activity upon attachment to Gal4(1–100). However, both Gcn4 and p53 functioned more robustly by >1 order of magnitude. p53 is a mammalian activator and in yeast lacks the natural masking

partner hDM2 that regulates its stability and activity. It appears that attachment to Gal4(1–100) results in a restoration of a significant portion of the masking function.

Typically the potency of a transcriptional activator is ascribed to the strength of its interaction with the transcriptional machinery and/or the location of that binding interaction (6, 13, 28, 29). Here we have demonstrated that interaction(s) outside of the

transcriptional machinery have a profoundly positive effect on transcription function. Micro-molar ligands for the transcriptional machinery, for example, only function as robust transcription activators when such a secondary binding interaction is available. Given the hydrophobic nature of these ligands, it is likely that the secondary interaction alters the binding and stability profile of the ligands, decreasing non-specific binding and premature proteolysis that would reduce the effective concentration of the ligands available for transcriptional activation. This strategy of incorporating a secondary “masking” interaction appears general; several natural and artificial activator peptides had their activity significantly enhanced *in vivo* via this mechanism. These results thus address a long-standing challenge in artificial transcriptional activator development, robust cellular activity. Although these studies were carried out with peptidic ADs, incorporation of a similar feature into future generations of small-molecule transcriptional activators (30, 31) should significantly augment both their activity and utility.

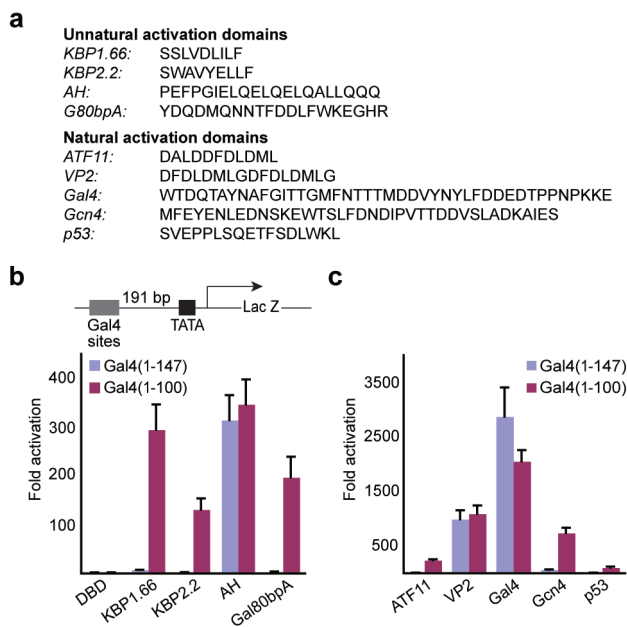


Figure 4. Activity enhancement of natural activation domains. **a)** The sequences of activation domains that contain hydrophobic residues important for activity used to evaluate the transferability of the Gal4_{dd} interaction. **b** and **c)** Activation domains listed in **a)** were fused to the Gal4(1–100) or Gal4(1–147) DBD and tested in the β-galactosidase reporter assay described in Figure 2. The activity and error are as reported in Figure 3, panel c.

transcriptional machinery, for example, only function as robust transcription activators when such a secondary binding interaction is available. Given the hydrophobic nature of these ligands, it is likely that the secondary interaction alters the binding and stability profile of the ligands, decreasing non-specific binding and premature proteolysis that would reduce the effective

METHODS

Yeast Strains and Plasmids. The yeast strain used for determining transcriptional activity was ZZY41 MATα his3Δ200 leu2-1 Δtrp1 ura3-52 Δlys2, gal4, gal80 URA3::pZZ41 (32). This strain contains a GAL1-LacZ reporter with five GAL4 DNA-binding sites 199 bp upstream of the TATA-box integrated at the URA3 locus. The Gal4(1–100) or Gal4(1–147) constructs were on an ARS/CEN plasmid under the control of a β-actin promoter with a HIS⁺ selection marker. Plasmids encoding for ADs fused *via* a SerSer linker to Gal4(1–100) and Gal4(1–147) were generated by incorporating oligos for each AD sequence into pJKL5 and pJKL11, respectively, as previously described (14). The strain used to test the Gal4_{dd} mutants was EGY48 containing the plasmid reporter pSH18-34 (Invitrogen). Alanine scanning mutagenesis was performed on pJKL4, an ARS/CEN plasmid that contains LexA(1-87)-Gal4(40–100) driven by a β-actin promoter with a HIS⁺ selection marker. To generate the LexA+Gal4(40–100) alanine mutants, site-directed mutagenesis was performed on pJKL4 (wild type), pJKL4-8 (peptide 36), and pJKL4-20 (VP2) in accordance with manufacturer instructions (Stratagene). The plasmid for bacterial expression of Gal4(1–100) was generated by polymerase chain reaction from yeast

genomic DNA and cloned into pGEX6p-2 (GE Healthcare).

Expression and Purification of Gal4(1–100).

glutathione S-transferase (GST) tagged Gal4(1–100) was isolated from Rosetta2(DE3) pLysS *Escherichia coli* (Novagen) induced at 16 °C with 0.1 mM isopropyl-beta-D-thiogalactopyranoside and purified using glutathione Sepharose beads (GE Healthcare). Cleavage of the GST tag was accomplished overnight at 4 °C using PreScission Protease (GE Healthcare) in cleavage buffer (50 mM Tris pH 7.0, 0.2% w/v NP-40, 1 mM DTT). The solution was loaded onto a gel filtration column (Superdex 75, GE Healthcare) to remove residual GST and to exchange the buffer to storage buffer (20 mM Hepes pH 6.8, 200 mM NaCl, 1 mM DTT, 10% v/v glycerol, 0.1 mM EDTA and 0.01% w/v NP-40). Fractions containing Gal4(1–100) were pooled and concentrated using a Vivascience 10K centrifugal filter device.

Quantitative β -Galactosidase Assays. As previously described (14).

Peptide Synthesis and Fluorescent Labeling. Peptides were synthesized on Rink amide resin using fmoc-protected amino acids, labeled at the amino terminus with 5/6-carboxyfluorescein succinimidyl ester (Pierce), and characterized by electrospray mass spectrometry using standard protocols (13). A SerSer linker was incorporated at the amino terminus of each sequence analogous to the Gal4(1–100) constructs used to assess cellular activity.

Dissociation Constant Measurements. Dissociation constant measurements were carried out at RT on a Beacon 2000. Prior to each experiment, a 13 μ M solution of Gal4(1–100) (dimer) in storage buffer was added to an equal concentration of duplex DNA containing an optimal Gal4 binding site (TCCGGAGGACTGTCTCCGG and its complement) for 30 min at RT. Next, fluorescein-labeled peptide was added to the Gal4(1–100)-DNA solution to give a final concentration of 50 nM labeled peptide and measurements were taken as previously described (12).

Acknowledgment: A.Z.A. is grateful for the support of the March of Dimes Foundation, the Steenbock and Industrial and Economic Development Research funds from University of Wisconsin–Madison and Wisconsin Alumni Research Foundation. A.K.M acknowledges support for this work from the National Institutes of Health (GM65330), the Alfred P. Sloan Foundation, the Burroughs Wellcome Fund, and the March of Dimes foundation. J.K.L. was supported by the Michigan C.B.I. Training Program (NIH GM09597-07). We thank C. Luty and the Saper laboratory for technical assistance.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Ptashne, M., and Gann, A. (2001) *Genes & Signals*, Cold Spring Harbor Laboratory, New York.
- Damell, J. E., Jr. (2002) Transcription factors as targets for cancer therapy, *Nat. Rev. Cancer* 2, 740–749.
- Ansari, A. Z., and Mapp, A. K. (2002) Modular design of artificial transcription factors, *Curr. Opin. Chem. Biol.* 6, 765–772.
- Lum, J. K., and Mapp, A. K. (2005) Artificial transcriptional activation domains, *ChemBioChem* 6, 1311–1315.
- Green, M. R. (2005) Eukaryotic transcription activation: right on target, *Mol. Cell* 18, 399–402.
- Melcher, K. (2000) The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins, *J. Mol. Biol.* 301, 1097–1112.
- Lohr, D., Venkov, P., and Zlatanova, J. (1995) Transcriptional regulation in the yeast GAL gene family: a complex genetic network, *FASEB J.* 9, 777–787.
- des Etages, S. A., Falvey, D. A., Reece, R. J., and Brandriss, M. C. (1996) Functional analysis of the PUT3 transcriptional activator of the proline utilization pathway in *Saccharomyces cerevisiae*, *Genetics* 142, 1069–1082.
- Brooks, C. L., and Gu, W. (2006) p53 ubiquitination: Mdm2 and beyond, *Mol. Cell* 21, 307–315.
- Khidekel, N., and Hsieh-Wilson, L. C. (2004) A ‘molecular switchboard’—covalent modifications to proteins and their impact on transcription, *Org. Biomol. Chem.* 2, 1–7.
- Muratani, M., and Tansey, W. P. (2003) How the ubiquitin-proteasome system controls transcription, *Nat. Rev. Mol. Cell Biol.* 4, 192–201.
- Lu, Z., Rowe, S. P., Brennan, B. B., Davis, S. E., Metzler, R. E., Nau, J. J., Majumdar, C. Y., Mapp, A. K., and Ansari, A. Z. (2005) Unraveling the mechanism of a potent transcriptional activator, *J. Biol. Chem.* 280, 29689–29698.
- Wu, Z., Belanger, G., Brennan, B. B., Lum, J. K., Minter, A. R., Rowe, S. P., Plachetka, A., Majumdar, C. Y., and Mapp, A. K. (2003) Targeting the transcriptional machinery with unique artificial transcriptional activators, *J. Am. Chem. Soc.* 125, 12390–12391.
- Majumdar, C. Y., Lum, J. K., Prasov, L., and Mapp, A. K. (2005) Functional specificity of artificial transcriptional activators, *Chem. Biol.* 12, 313–321.
- Mayer, M. P., and Bukau, B. (2005) Hsp70 chaperones: cellular functions and molecular mechanism, *Cell. Mol. Life Sci.* 62, 670–684.
- Horwich, A. L., Weber-Ban, E. U., and Finley, D. (1999) Chaperone rings in protein folding and degradation, *Proc. Natl. Acad. Sci. U.S.A.* 96, 11033–11040.
- Hidalgo, P., Ansari, A. Z., Schmidt, P., Hare, B., Simkovich, N., Farrell, S., Shin, E. J., Ptashne, M., and Wagner, G. (2001) Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain, *Genes Dev.* 15, 1007–1020.
- Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S. C. (1992) DNA recognition by GAL4: structure of a protein-DNA complex, *Nature* 356, 408–414.
- Himmelfarb, H. J., Pearlberg, J., Last, D. H., and Ptashne, M. (1990) GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators, *Cell* 63, 1299–1309.
- Tossi, A., and Sandri, L. (2002) Molecular diversity in gene-encoded, cationic antimicrobial polypeptides, *Curr. Pharm. Des.* 8, 743–761.
- Chakshumathi, G., Mondal, K., Lakshmi, G. S., Singh, G., Roy, A., Ch, R. B., Madhusudhanan, S., and Varadarajan, R. (2004) Design of temperature-sensitive mutants solely from amino acid sequence, *Proc. Natl. Acad. Sci. U.S.A.* 101, 7925–7930.
- Frangioni, J. V., LaRiccia, L. M., Cantley, L. C., and Montminy, M. R. (2000) Minimal activators that bind to the KIX domain of p300/CBP identified by phage display screening, *Nat. Biotechnol.* 18, 1080–1085.
- Giniger, E., and Ptashne, M. (1987) Transcription in yeast activated by a putative amphipathic alpha-helix linked to a DNA-binding unit, *Nature* 330, 670–672.
- Han, Y., and Kodadek, T. (2000) Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast, *J. Biol. Chem.* 275, 14979–14984.
- Ferreira, M. E., Hermann, S., Prochasson, P., Workman, J. L., Berndt, K. D., and Wright, A. P. (2005) Mechanism of transcription factor recruitment by acidic activators, *J. Biol. Chem.* 280, 21779–21784.
- Stanojevic, D., and Young, R. A. (2002) A highly potent artificial transcription factor, *Biochemistry* 41, 7209–7216.
- Tanaka, M. (1996) Modulation of promoter occupancy by cooperative DNA binding and activation-domain function is a major determinant of transcriptional regulation by activators *in vivo*, *Proc. Natl. Acad. Sci. U.S.A.* 93, 4311–4315.
- Wu, Y. B., Reece, R. J., and Ptashne, M. (1996) Quantitation of putative activator-target affinities predicts transcriptional activating potentials, *EMBO J.* 15, 3951–3963.
- Volkman, H. M., Rutledge, S. E., and Schepartz, A. (2005) Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain, *J. Am. Chem. Soc.* 127, 4649–4658.
- Minter, A. R., Brennan, B. B., and Mapp, A. K. (2004) A small molecule transcriptional activation domain, *J. Am. Chem. Soc.* 126, 10504–10505.
- Kwon, Y., Arndt, H. D., Mao, Q., Choi, Y., Kawazoe, Y., Devan, P. B., and Uesugi, M. (2004) Small molecule transcription factor mimic, *J. Am. Chem. Soc.* 126, 15940–15941.
- Gaudreau, L., Keaveney, M., Nevado, J., Zaman, Z., Bryant, G. O., Struhl, K., and Ptashne, M. (1999) Transcriptional activation by artificial recruitment in yeast is influenced by promoter architecture and downstream sequences, *Proc. Natl. Acad. Sci. U.S.A.* 96, 2668–2673.
- Chrivia, J. C., Kwok, R. P., Lamb, N., Hagiwara, M., Montminy, M. R., and Goodman, R. H. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP, *Nature* 365, 855–859.

A Comparative Study of Bioorthogonal Reactions with Azides

Nicholas J. Agard^{†,††}, Jeremy M. Baskin^{†,††}, Jennifer A. Prescher[†], Anderson Lo[†], and Carolyn R. Bertozzi^{†,‡,§,¶,||,*}

Departments of [†]Chemistry and ^{*†}Molecular and Cell Biology and [§]Howard Hughes Medical Institute, University of California, Berkeley, California 94720, and [¶]Molecular Foundry, Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720^{††}these authors contributed equally to this work.

ABSTRACT

Detection of metabolites and post-translational modifications can be achieved using the azide as a bioorthogonal chemical reporter. Once introduced into target biomolecules, either metabolically or through chemical modification, the azide can be tagged with probes using one of three highly selective reactions: the Staudinger ligation, the Cu(I)-catalyzed azide-alkyne cycloaddition, or the strain-promoted [3 + 2] cycloaddition. Here, we compared these chemistries in the context of various biological applications, including labeling of biomolecules in complex lysates and on live cell surfaces. The Cu(I)-catalyzed reaction was found to be most efficient for detecting azides in protein samples but was not compatible with live cells due to the toxicity of the reagents. Both the Staudinger ligation and the strain-promoted [3 + 2] cycloaddition using optimized cyclooctynes were effective for tagging azides on live cells. The best reagent for this application was dependent upon the specific structure of the azide. These results provide a guide for biologists in choosing a suitable ligation chemistry.

The discovery of the green fluorescent protein launched a new era in cellular biochemistry in which proteins could be monitored in complex living systems. Fluorescent protein fusions, in addition to other genetically encoded epitope tags, have been artfully employed to identify the subcellular localization, trafficking, and interaction patterns of thousands of proteins, while also facilitating their purification for molecular analysis (1). Meanwhile, biomolecules that are not directly encoded in the genome, such as glycans, lipids, and other metabolites, are not amenable to these conventional tagging technologies. Their importance as post-translational modifications and signaling molecules has placed some urgency on filling this void.

The bioorthogonal chemical reporter strategy provides a means to tag biomolecules without the requirement of direct genetic encoding (2). In this approach, a functional group (the chemical reporter) that does not interact with any biological functionality (*i.e.*, bioorthogonal) is incorporated into the target biomolecule using the cell's metabolic machinery (3–5) or delivered to a protein by virtue of its enzymatic activity (6, 7). Subsequently, the reporter is covalently tagged with an exogenous probe using a highly selective chemical reaction (8, 9). This two-step procedure has been used to modify cell surface glycans and proteins (5, 10, 11), profile protein glycosylation (12–14) and farnesylation (3), and identify proteins with a specific catalytic mechanism

using activity-based probes (6, 7). With the expansion of the bioorthogonal chemical reporter strategy into the realm of living animals (6, 7, 13, 15), the profiling and noninvasive imaging of biomarkers associated with disease progression appears imminent.

Success of the bioorthogonal chemical reporter strategy rests on the choice of the appropriate functional group and labeling reaction. The chemical reporter must be tolerated by the cellular machinery and sufficiently robust to avoid unwanted chemical or metabolic side reactions. In addition, the labeling reaction must proceed rapidly and selectively at physiological pH and temperatures. For applications using live cells or organisms, the reagents must be nontoxic.

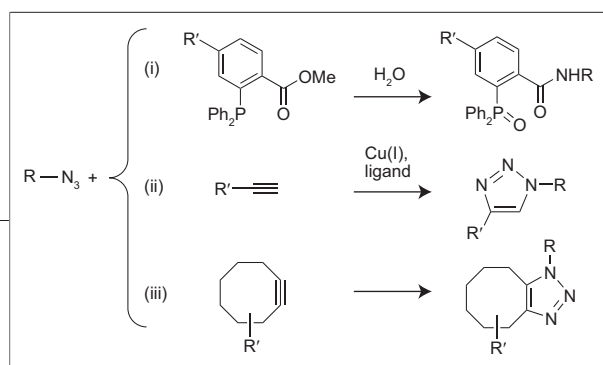
The azide is the most versatile bioorthogonal chemical reporter. Its small size and stability in physiological settings have enabled azide-functionalized metabolic precursors to hijack the biosynthetic pathways for numerous biomolecules, including glycans (16), proteins (4, 17), lipids (3), and nucleic acid-derived cofactors (18). Three reactions have been reported for tagging azide-labeled biomolecules (Scheme 1). One of these, the Staudinger ligation (Scheme 1, reaction i), capitalizes on the selective reactivity of phosphines and azides to form an amide bond (5, 15, 19). The other two involve the reaction of azides with alkynes to give triazoles, a process that is typically very slow under ambient conditions. The Cu(I)-catalyzed azide-alkyne

*Corresponding author,
crb@berkeley.edu

Received for review July 28, 2006
and accepted September 25, 2006.

Published online November 10, 2006
10.1021/cb6003228 CCC: \$33.50

© 2006 by American Chemical Society



Scheme 1. Bioorthogonal reactions with the azide: biomolecules containing the azide react *via* the Staudinger ligation (i), click chemistry (ii), or a strain-promoted cycloaddition (iii) to give ligated products.

cycloaddition (Scheme 1, reaction ii), also known as “click chemistry”, accelerates the reaction by use of a copper catalyst (20–22). The strain-promoted [3 + 2] cycloaddition (Scheme 1, reaction iii) removes the requirement for cytotoxic copper by employing cyclooctynes that are activated by ring strain (11, 23).

Among these options, the ideal choice of a reaction for a given biological application is not always obvious. In many cases, sensitivity is the most important parameter. Defined as the number of azides that are reacted in a given time period, sensitivity is governed by the intrinsic kinetics of the reaction and the reagent concentrations. In practice, the concentration of azides in the biological system is limited by the abundance of the target molecule and the efficiency of azide labeling. The concentration of the secondary tagging reagent (*i.e.*, the phosphine or alkyne) is typically limited by solubility and, in live cell or animal experiments, toxicity. Given these constraints, the ability to improve intrinsic reaction kinetics can be critical for optimizing a labeling strategy with respect to sensitivity. In cases where live cells or organisms are under study, biocompatibility may trump sensitivity as the most important parameter.

In order to provide a framework for choosing the optimal reaction for a given purpose, we compared the three ligations in three situations: labeling of isolated proteins, labeling of low abundance proteins from mixtures, and labeling of live cell surfaces. For high-sensitivity protein labeling, click chemistry was found to be ideal because of its superior kinetics. By contrast, live cell labeling required either the Staudinger ligation

or strain-promoted [3 + 2] cycloaddition because of their superior biocompatibility. This comparison identified a conflict between sensitivity and biocompatibility that should ultimately be resolved for an ideal labeling reaction. As a first step toward this goal, we improved the kinetics of the strain-promoted [3 + 2] cycloaddition using physical organic chemistry principles.

Initially we sought to identify the optimal reagents for each reaction with respect to intrinsic kinetics. In previous work, our attempts to improve the kinetics of the Staudinger ligation focused on increasing the electron density of the phosphine substituents (19). Although rate enhancements were observed, these were accompanied by increased rates of phosphine oxidation by air, an unwanted side reaction. Thus, the parent phosphine (Scheme 1, reaction i) remains the best reagent for biological Staudinger ligations. Other groups have directed considerable effort toward optimizing click chemistry for biological labeling reactions (7). We employed these previously reported reagents and conditions in our study.

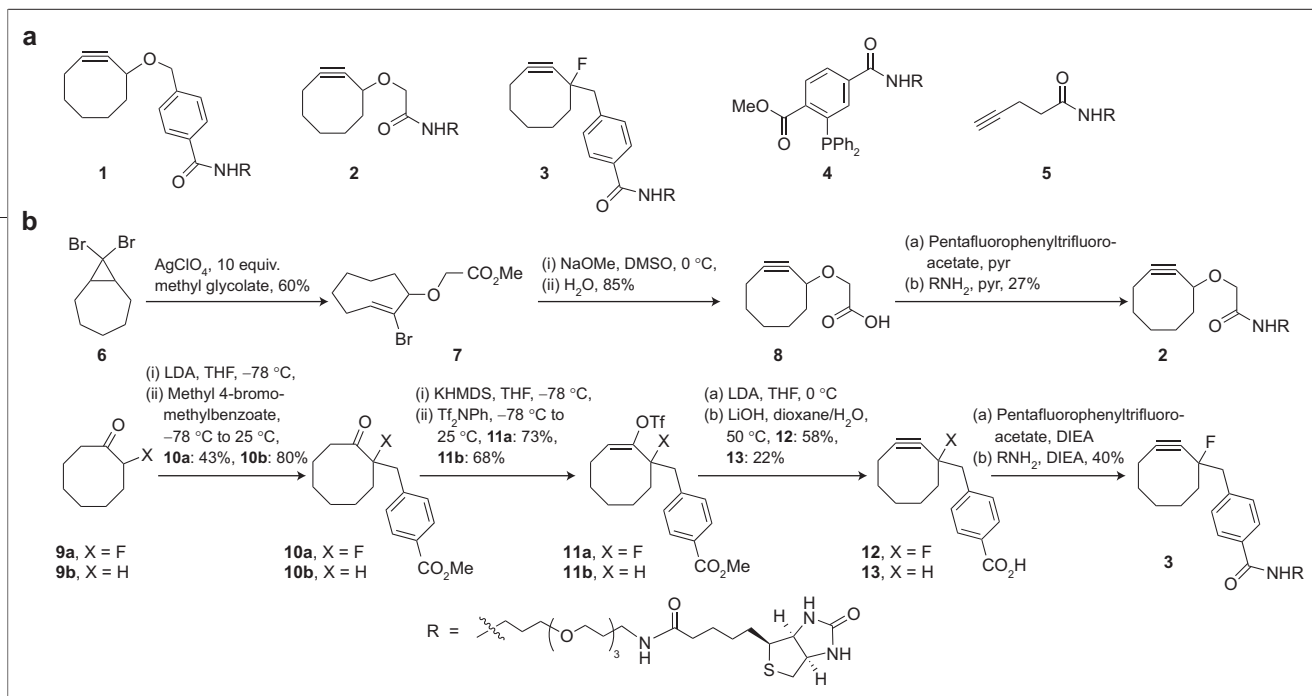
The only reaction that had not yet been explored with respect to kinetic enhancement is the strain-promoted [3 + 2] cycloaddition. Thus, we began our study by varying the substituents on the cyclooctyne scaffold. Two approaches were taken to improve upon previously reported cyclooctyne **1** (Scheme 2, panel a). First, the phenyl ring was excised (**2**) to improve the solubility of the reagent and possibly increase the rate of the reaction by decreasing steric bulk near the reactive center. In the second approach, an electron-withdrawing group (fluorine) was introduced adjacent to the alkyne (**3**) in order to lower the energy of its LUMO and promote reaction with the azide (24). As an additional benefit, this modification replaces the oxidatively labile ether link-

ages in **1** and **2** (Supplementary Scheme 1) with a stable carbon–carbon bond.

Synthesis of New Cyclooctyne Probes.

The synthesis of **2** proceeded similarly to that of compound **1** (23) (Scheme 2, panel b); silver perchlorate-mediated electrocyclic ring opening of **6** yielded a transient allylic cation that was trapped by methyl glycolate. One-pot elimination and hydrolysis of *trans*-bromocyclooctene **7** yielded the desired cyclooctyne **8**. The synthesis of cyclooctyne **3** was accomplished *via* alkylation of 2-fluorocyclooctanone (**9a**) (25) to give substituted cyclooctanone **10a**. Vinyl triflate formation and elimination, followed by saponification of the methyl ester, yielded cyclooctyne **12**. Non-fluorinated analogue **13** was synthesized *via* a similar route in order to directly assess the effect of the fluorine substituent. Biotinylation of the panel of cyclooctynes was achieved *via* formation of the pentafluorophenyl ester followed by condensation with an amine-modified biotin (26).

Kinetic Evaluation of Cyclooctynes. The relative reactivities of cyclooctynes, **8**, **12**, and **13** were determined in model reactions with benzyl azide in CD₃CN. The electron-withdrawing fluorine atom on **12** provided enhanced reactivity ($k = 4.3 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$) compared to the free acid of **1** ($k = 2.4 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$) (23) and compounds **8** ($k = 1.3 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$) and **13** ($k = 1.2 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$). In all cases, the only products observed were the triazole regioisomers as a ~1:1 mixture. Interestingly, the reaction of **12** with 2-azidoethanol in aqueous acetonitrile afforded the expected 1,5-substituted triazole, but the 1,4-isomer had undergone hydrolysis of the C–F bond to form a hydroxy-substituted product (Supplementary Scheme 2). Compound **12** in isolation was not subject to aqueous decomposition. The Staudinger ligation with benzyl azide proceeds at a similar rate in CD₃CN ($k = 2.0 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$) (19), while click chemistry is generally faster but subject to more complex kinetic behavior



Scheme 2. Biotinylated detection reagents used to probe for the presence of azides (a) and synthesis of new cyclooctynes (b).

because of its multiple reaction components (27). As an additional control, we investigated the chemical stabilities of cyclooctynes **8**, **12**, and **13** by incubating them in aqueous acid, aqueous base, β -mercaptoethanol, or phosphate-buffered saline (PBS) (see Supporting Information for details). In all cases, no decomposition was observed by ^1H NMR and ^{19}F NMR, where applicable.

Protein Labeling. We next sought to compare the reactions in the context of biological labeling experiments. The least demanding situation involves modification of a purified azide-labeled biomolecule. Thus, we expressed dihydrofolate reductase (DHFR) in which methionine residues were replaced with the unnatural amino acid azidohomoalanine (DHFR- N_3) (4, 28). First, the time dependencies of the reactions of compounds **1–5** with DHFR- N_3 were compared. DHFR- N_3 was incubated with **1–5** (100 μM) for 4, 12, or 24 h, and reaction progress was monitored by Western blot using anti-biotin antibody-HRP conjugate (Figure 1, panel a). For click chemistry, the relative concentrations of **5**, triazolyl ligand, tris-carboxyethylphosphine (TCEP), and CuSO_4 were held at a ratio of 1:1:10:10, as optimized by Cravatt *et al.* (7). Preliminary investigations found that urea and ionic detergents used to solubilize the protein inhibited click chemistry and the Staudinger

ligation, respectively (data not shown). Consequently, Staudinger ligations were performed in 8 M urea while the remainder of the ligations were run in 1% SDS. All of the protein labeling reactions were highly specific at 37 $^\circ\text{C}$, as native DHFR showed no detectable biotinylation. However, nonspecific biotinylation was observed for both the strain-promoted cycloadditions and click chemistry when the reactions were boiled (data not shown). All five reagents tested showed time-dependent protein labeling that was consistent with their kinetics in model reactions (Figure 1, panel a). Click chemistry afforded the highest sensitivity of labeling, with maximal labeling observed by 4 h.

Next, the concentration dependence of the reactions was compared (Figure 1, panel a). The reactions were performed with 50, 100, or 200 μM reagent for a period of 8 h. All three ligations exhibited labeling proportional to reagent concentration. Interestingly, the large shift in apparent molecular weight of DHFR- N_3 in the click chemistry lanes underscores the higher-order concentration dependence that this ligation obeys because of its multicomponent nature (27). This observation is particularly relevant for applications that require low concentrations of labeling reagent to avoid toxicity. In addition, in living animals, where the site of injection can be far removed from the tar-

geted biomolecule, trafficking can limit the concentration of reagents near the azide.

A more stringent test of the reactions' bioorthogonality is their ability to specifically label azides in the presence of complex mixtures, a feature essential for enrichment from lysates in preparation for proteomic analysis. To simulate a labeled lysate, we combined DHFR- N_3 with 200 μg of *Escherichia coli* lysate (total protein concentration of 10 mg/mL). These mixtures were then labeled with reagents **1–5** (200 μM , 8 h). Specific protein labeling was observed with all five reagents, with click chemistry affording the highest sensitivity, consistent with the trend observed using purified protein.

Live Cell Labeling. Having demonstrated selective labeling of proteins within complex lysates, we sought to extend the comparison to labeling of living cells. Jurkat cells were grown in the presence of 25 μM peracetylated *N*-azidoacetylmannosamine (Ac_4ManNAz) for 2–3 d, leading to the metabolic incorporation of the corresponding *N*-azidoacetyl sialic acid (SiaNAz) into their cell surface glycoproteins (5, 29). The cells were then treated with compounds **1–5** (50 or 100 μM for 1 h), incubated with FITC-avidin, and analyzed by flow cytometry (Figure 1, panel c). We observed significant cell death in the presence of the click reagents (Supplementary Figure 1), which

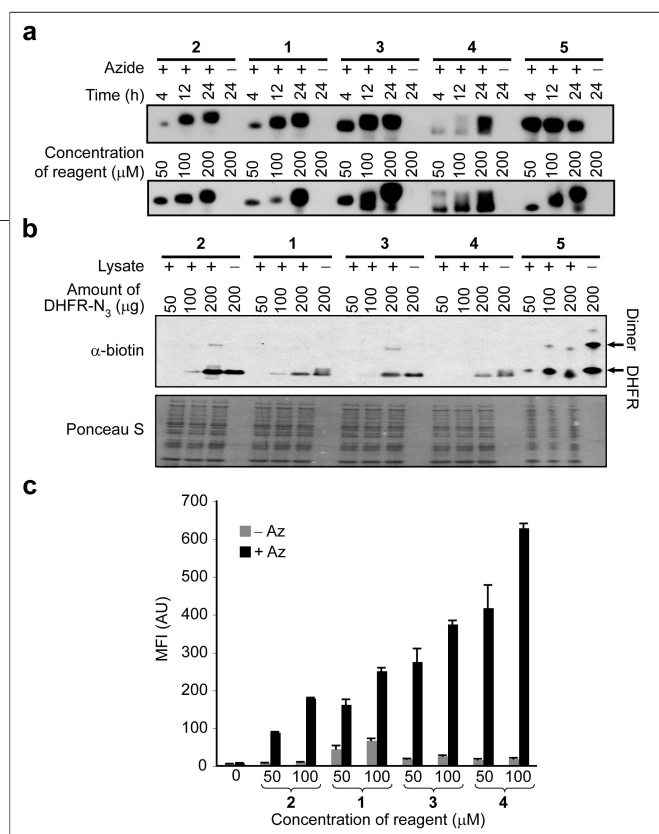


Figure 1. Relative labeling efficiencies of the reactions on isolated proteins, cell lysate, and live cells. **a)** DHFR (200 μg) grown in the presence (+) or absence (-) of azidohomoalanine was labeled with biotinylated reagents 1–5 under given conditions followed by Western blot. Top: Time course analysis at 100 μM reagent. Bottom: Concentration dependence over 8 h. The protein bands broaden and increase in molecular weight over time and concentration, reflecting multiple sites of modification. **b)** DHFR-N₃ in the presence (+) or absence (-) of 10 mg mL⁻¹ *E. coli* cell lysate is labeled with 200 μM reagent for 8 h. Top: Labeled proteins are detected by Western blot analysis. Labeling of DHFR-N₃ and its dimer was found to be largely lysate-independent for each of these reactions. Bottom: Total protein content was determined using Ponceau S. **c)** Jurkat cells grown for 3 d in the presence (+ Az) or absence (- Az) of peracetylated *N*-azidoacetylmannosamine (25 μM) were labeled for 1 h with 0, 50, or 100 μM reagent followed by secondary labeling with FITC-avidin. The resulting mean fluorescence intensity (MFI) of the cell populations was determined by flow cytometry. Error bars represent the standard deviation from three separate experiments.

we attribute to the high concentration of CuSO₄, as previously suggested in the context of an *E. coli* labeling study (11, 20). Cyclooctyne reagents 1–3 labeled cells in proportion to their rate constants in the model reaction, while phosphine 4 showed a higher degree of labeling than expected based on its relative rate of reaction with benzyl azide. To determine the basis of this discrepancy, we performed additional model reactions of phosphine 4 and cyclooctyne 1 with α -azido acetamides that better mimic the reactivity of SiaNAz. In these reactions, the phosphine outperformed the cyclooctyne by ~2-fold (Supplementary Table 1). These results suggest that the Staudinger ligation is more efficient for labeling azides bearing electron-withdrawing or resonance stabilizing groups. The strain-promoted cycloaddition is relatively insensitive to the electronics of the azide and is more efficient than the Staudinger ligation with unactivated alkyl azides.

Reaction Guide for Specific Applications.

Collectively, these experiments provide a guide for choosing the optimal ligation chemistry for specific applications (Table 1). Each of the reactions is competent to label isolated biomolecules. Click chemistry is the

most efficient application, although some groups have reported difficulty in separating intact modified biomolecules from catalytic copper (22). The residual heavy metal was found to interfere with analysis by mass spectrometry and could alter the activity of the modified protein. For situations where copper is a concern and purification by chromatography is prohibitive, either phosphine 4 or cyclooctyne 3 is an appropriate choice.

For proteomic applications, click chemistry is the clear choice. Its superior sensitivity should provide the most efficient detection of low abundance species. Furthermore, the difficulties encountered when intact proteins bind copper can be eliminated by trypsinization and LC (6). The only caveat to this approach is that the chemistry is not compatible with all detergents. For rare cases where labeling in the presence of specific detergents is necessary, cyclooctyne 3

is the best alternative, as phosphine 4 has also shown detergent sensitivity.

For labeling of live cells, the Staudinger ligation and strain-promoted [3 + 2] cycloadditions are both suitable choices, with the best reaction being dictated by the nature of the azide. At this point, the toxicity of the copper catalyst necessary for click chemistry limits its utility for labeling live cells. The Staudinger ligation has already been shown to perform in living mice without discernible toxicity (13, 15). An interesting future direction is to explore whether the cyclooctynes are amenable to such applications.

An important conclusion from this study is that a need remains for azide ligation reactions that are both highly sensitive and biocompatible. Within the current repertoire of reactions, these attributes are in conflict. Optimization of the cycloaddition reagents might be possible *via* synthesis of tight-binding copper ligands to mitigate click

is the best alternative, as phosphine 4 has also shown detergent sensitivity.

TABLE 1. Selection of reaction depends on application^a

Reaction	Optimal for labeling
Staudinger ligation	Surfaces of live cells; live organisms
Click chemistry	Proteomic samples
Strain-promoted [3 + 2] cycloaddition	Surfaces of live cells

^aReagents used in each reaction are shown in Scheme 1.

chemistry's cytotoxicity or further substitutions on the cyclooctyne scaffold to increase its rate. Although attempts to increase the rate of the Staudinger ligation have led to increases in non-specific oxidation, its increased selectivity at elevated temperatures might enable rate enhancement by localized heating (*i.e.*, via focused ultrasound) (30).

METHODS

New Compounds. Synthetic methods and characterization of new compounds is provided in Supplementary Methods.

Kinetic Evaluation of Cyclooctynes. The rates of reaction between cyclooctynes and various azides in CD₃CN were monitored by the disappearance of starting materials and appearance of the two regioisomeric products in the ¹H NMR spectrum. Second-order rate constants for the reaction were determined by plotting the 1/[reagent] vs time, followed by subsequent analysis by linear regression. The rate constants correspond to the determined slope. For further details, please see the Supporting Information.

Western Blot Analysis of DHFR-N₃. For time-dependent Western blot analysis of DHFR and DHFR-N₃, 200 ng portions of protein samples were incubated in 20 μL of **1–4** (100 μM final concentration) or **5** (100 μM final concentration) with 1 mM TCEP, 100 μM tris-triazolyl ligand (TBTA) and 1 mM CuSO₄ for the indicated periods of time. Prior to electrophoresis, samples were incubated with an equal volume of 100 mM 2-azidoethanol in 2X SDS-PAGE loading buffer for 8 h at RT (to quench unreacted **1–5**). Samples were subject to SDS-PAGE analyzed by Sypro Ruby Red or transferred to nitrocellulose membranes and detected by Western blot analysis with anti-biotin-HRP antibody. Concentration-dependent labeling was determined in the same manner, but with the indicated concentration of **1–4** or **5** with triazolyl ligand TBTA, TCEP and CuSO₄ in a 1:1:10:10 ratio for 8 h.

For labeling in the presence of complex mixtures, the indicated amount of protein was incubated with 200 μg (10 mg mL⁻¹ of soluble protein from *E. coli* lysate and buffers as above (final concentration of **1–5** was 200 μM). The extent and specificity of labeling were determined by Western blot analysis as described above. Detailed experimental procedures for Western blot experiments are included in the Supporting Information.

Cell Surface Azide Labeling and Detection. Jurkat cells bearing azides were treated with 0–100 μM of the biotinylated probes **1–4** in labeling buffer (PBS, pH 7.4 containing 1% (v/v) fetal calf serum) and analyzed by flow cytometry as previously described (30). For all flow cytometry experiments, data points were collected in triplicate and are representative of three separate experiments. For further details, see the Supporting Information.

Acknowledgment: This research was supported by a grant from the National Institutes of Health (GM58867). The authors thank Isaac Carrico and Jason Rush for materials and the Bertozzi lab for helpful discussions. J.A.P. was supported by a Howard Hughes Medical Institute predoctoral fellowship and J.M.B. was supported by a National Defense Science and Engineering Graduate fellowship.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Zacharias, D. A., and Tsien, R. Y. (2006) Molecular biology and mutation of green fluorescent protein, *Methods Biochem. Anal.* **47**, 83–120.
- Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat. Chem. Biol.* **1**, 13–21.
- Kho, Y., Kim, S. C., Jiang, C., Barma, D., Kwon, S. W., Cheng, J., Jaunbergs, J., Weinbaum, C., Tamanoi, F., Falck, J., and Zhao, Y. (2004) A tagging-via-substrate technology for detection and proteomics of famesylated proteins, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12479–12484.
- Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 19–24.
- Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction, *Science* **287**, 2007–2010.
- Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003) Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition, *J. Am. Chem. Soc.* **125**, 4686–4687.
- Speers, A. E., and Cravatt, B. F. (2004) Profiling enzyme activities in vivo using click chemistry methods, *Chem. Biol.* **11**, 535–546.
- Hang, H. C., and Bertozzi, C. R. (2001) Chemoselective approaches to glycoprotein assembly, *Acc. Chem. Res.* **34**, 727–736.
- Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering, *Curr. Opin. Biotechnol.* **14**, 603–609.
- Mahal, L. K., Yarema, K. J., and Bertozzi, C. R. (1997) Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis, *Science* **276**, 1125–1128.
- Link, A. J., Vink, M. K. S., Agard, N. J., Prescher, J. A., Bertozzi, C. R., and Tirrell, D. A. (2006) Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10180–10185.
- Vocadlo, D. J., and Bertozzi, C. R. (2004) A strategy for functional proteomic analysis of glycosidase activity from cell lysates, *Angew. Chem., Int. Ed.* **43**, 5338–5342.
- Dube, D. H., Prescher, J. A., Quang, C. N., and Bertozzi, C. R. (2006) Probing mucin-type O-linked glycosylation in living animals, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4819–4824.
- Sprung, R., Nandi, A., Chen, Y., Kim, S. C., Barma, D., Falck, J. R., and Zhao, Y. (2005) Tagging-via-substrate strategy for probing O-GlcNAc modified proteins, *J. Proteome Res.* **4**, 950–957.
- Prescher, J. A., Dube, D. H., and Bertozzi, C. R. (2004) Chemical remodelling of cell surfaces in living animals, *Nature* **430**, 873–877.
- Dube, D. H., and Bertozzi, C. R. (2003) Metabolic oligosaccharide engineering as a tool for glycobiology, *Curr. Opin. Chem. Biol.* **7**, 616–625.
- Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code, *Science* **301**, 964–967.
- Weller, R. L., and Rajsiki, S. R. (2005) DNA methyltransferase-mediated click chemistry, *Org. Lett.* **7**, 2141–2144.
- Lin, F. L., Hoyt, H. M., Van Halbeek, H., Bergman, R. G., and Bertozzi, C. R. (2005) Mechanistic investigation of the Staudinger ligation, *J. Am. Chem. Soc.* **127**, 2686–2695.
- Link, A. J., and Tirrell, D. A. (2003) Cell surface labeling of *Escherichia coli* via copper(I)-catalyzed [3 + 2] cycloaddition, *J. Am. Chem. Soc.* **125**, 11164–11165.
- Link, A. J., Vink, M. K. S., and Tirrell, D. A. (2004) Presentation and detection of azide functionality in bacterial cell surface proteins, *J. Am. Chem. Soc.* **126**, 10598–10602.
- Wang, Q., Chan, T. R., Hilgraf, R., Fokin, V. V., Sharpless, K. B., and Finn, M. G. (2003) Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition, *J. Am. Chem. Soc.* **125**, 3192–3193.
- Agard, N. J., Prescher, J. A., and Bertozzi, C. R. (2004) A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems, *J. Am. Chem. Soc.* **126**, 15046–15047.
- Nguyen, L. T., De Proft, F., Dao, V. L., Nguyen, M. T., and Geertlings, P. (2003) A theoretical approach to the regioselectivity in 1,3-dipolar cycloadditions of diazoalkanes, hydrazoic acid and nitrous oxide to acetylenes, phosphoalkynes and cyanides, *J. Phys. Org. Chem.* **16**, 615–625.
- Nyffeler, P. T., Duron, S. G., Burkhart, M. D., Vincent, S. P., and Wong, C.-H. (2005) Selectfluor: mechanistic insight and applications, *Angew. Chem., Int. Ed.* **44**, 192–212.
- Wilbur, D. S., Hamlin, D. K., Vessella, R. L., Stray, J. E., Buhler, K. R., Stayton, P. S., Klumb, L. A., Pathare, P. M., and Weerawarna, S. A. (1996) Antibody fragments in tumor pretargeting. Evaluation of biotinylated Fab' colocalization with recombinant streptavidin and avidin, *Bioconjugate Chem.* **7**, 689–702.
- Rodionov, V. O., Fokin, V. V., and Finn, M. G. (2005) Mechanism of the ligand-free CuI-catalyzed azide-alkyne cycloaddition reaction, *Angew. Chem., Int. Ed.* **44**, 2210–2215.
- Mangold, J. B., Mischke, M. R., and LaVelle, J. M. (1989) Azidoalanine mutagenicity in Salmonella: Effect of homologation and alpha-methyl substitution, *Mutat. Res.* **216**, 27–33.
- Saxon, E., Luchansky, S. J., Hang, H. C., Yu, C., Lee, S. C., and Bertozzi, C. R. (2002) Investigating cellular metabolism of synthetic azidosugars with the Staudinger ligation, *J. Am. Chem. Soc.* **124**, 14893–14902.
- Wu, F. (2006) Extracorporeal high intensity focused ultrasound in the treatment of patients with solid malignancy, *Minim. Invasive Ther. Allied Technol.* **15**, 26–35.

Mechanisms of Cellular Avidity Regulation in CD2–CD58-Mediated T Cell Adhesion

De-Min Zhu^{†,*}, Michael L. Dustin[§], Christopher W. Cairo^{†,¶}, Hemant S. Thatte^{†,||}, and David E. Golan^{†,***}

[†]Departments of Biological Chemistry and Molecular Pharmacology, Surgery, and Medicine, Harvard Medical School, Boston, Massachusetts 02115, [§]Skirball Institute of Biomolecular Medicine and the Department of Pathology, NYU School of Medicine, New York, New York 10016, ^{||}Department of Surgery, Veterans Affairs Boston Healthcare System, West Roxbury, Massachusetts 02132, and ^{**}Hematology Division, Brigham and Women's Hospital, Boston, Massachusetts 02115, ^{*}Current address: WP78-302, Merck Research Laboratories, Merck & Co. Inc., West Point, Pennsylvania 19486, [¶]Current address: Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

The interaction of T lymphocytes with antigen presenting cells (APCs) and target cells is critically dependent on CD2–CD58-mediated adhesion (1, 2). Antibody blocking of CD2 prevents T cell receptor (TCR)-mediated cell activation, and CD2 itself contributes to cell activation (3, 4). CD2 localizes to the T cell uropod, where the fast on and off rates of the receptor facilitate scanning of APCs (5, 6). The low molecular profile of the CD2–CD58 complexes allows these molecules to integrate with the TCR–peptide-bound major histocompatibility complex (pMHC) complexes (7, 8), providing a permissive environment in which the low-affinity TCR–pMHC interactions (9–13) can achieve the single receptor sensitivity required for antigen recognition (14). CD2-mediated T cell adhesion is modulated by the activation state of the cell. T cell activation, induced by TCR engagement or by phorbol-12-myristate-13-acetate (PMA) stimulation, has been reported to cause a 2.5-fold enhancement of CD2-mediated T cell adhesion (15). Several molecular mechanisms have been proposed to regulate CD2-mediated adhesion, including changes in the conformation and cytoskeletal interaction of CD2 (6), although the biophysical manifestations of CD2 adhesion modulation are not well defined.

Based on the results of experiments performed in solution, the CD2–CD58 interaction has a low equilibrium binding affinity and a fast kinetic dissociation rate (16–18). In the cellular context, however, adhesion receptors interact in a 2D rather than a 3D presentation; in the 2D environment, receptor mobility is restricted by association with the plasma membrane and by attachment to cytoskeletal proteins. At sites of cell–cell

ABSTRACT The CD2 receptor on T lymphocytes is essential for T cell adhesion and stimulation by antigen presenting cells (APCs). Blockade of CD2 function is immunosuppressive in both model systems and humans, indicating the importance of CD2 for the cellular immune response. Although the affinity of the molecular interaction between CD2 and its counter-receptor, CD58, is relatively low when measured in solution, this interaction mediates tight adhesion within the 2D cell–cell interface. To understand the mechanisms responsible for regulating the avidity of the CD2–CD58 interaction, we measured the number, affinity, and lateral mobility of CD2 molecules on resting and activated T cells. Cell activation caused a 1.5-fold increase in the number of CD2 sites on the cell surface, and the 2D affinity of CD2 for CD58 increased by 2.5-fold. The combination of T cell activation and CD2 ligation to CD58 decreased the laterally mobile fraction of the ligated CD2. Together, these changes would substantially enhance CD2 avidity and strengthen T cell–APC adhesion. The change in CD2 mobile fraction suggests that the cell uses cytoskeletal regulators to immobilize the receptor selectively at the site of contact with surfaces expressing CD58. Our observations are consistent with a model in which T cell activation initially induces increased CD2 2D affinity, cell surface receptor expression, and lateral mobility, allowing the CD2 molecules to diffuse to sites of contact with CD58-bearing APCs. Subsequently, T cell activation causes the CD58-bound CD2 to be recognized and immobilized at sites of cell–cell contact, thereby strengthening T cell–APC adhesion.

*Corresponding author,
dgolan@hms.harvard.edu.

Received for review June 14, 2006
and accepted October 23, 2006.

Published online November 17, 2006

10.1021/cb6002515 CCC: \$33.50

© 2006 by American Chemical Society

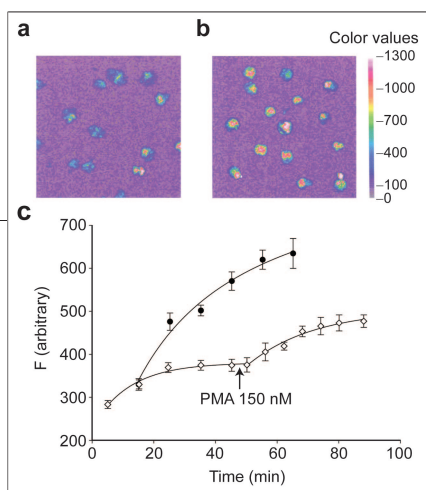


Figure 1. Kinetics of CD58 accumulation in cell bilayer contact area. Jurkat cells were incubated for 30 min with bilayers containing FITC-CD58 ($244 \text{ molecules } \mu\text{m}^{-2}$). Control cells (a) showed less accumulation of CD58 at the contact area than did cells treated with PMA (150 nM) (b). Color values represent fluorescence intensity. c) Jurkat cells were incubated with bilayers containing FITC-CD58, and fluorescence intensity (f) in the contact area was measured as a function of incubation time. Data points represent the mean fluorescence intensity for 13–34 cells at each point. Error bars represent standard error of the mean (SEM). (\diamond) Resting Jurkat cells were allowed to adhere to bilayers containing $50 \text{ molecules } \mu\text{m}^{-2}$ CD58, and 150 nM PMA was added at the indicated time. (\bullet) Jurkat cells were pre-incubated with 150 nM PMA for 20 min, and then the cells were allowed to adhere to bilayers containing $106 \text{ molecules } \mu\text{m}^{-2}$ CD58.

contact, dynamic polyvalent interactions between receptors and counter-receptors result from the mobile nature of receptors in biological membranes and the presentation of multiple binding sites on each cell, particularly as the systems approach equilibrium. Therefore, the physiological CD2-CD58 interaction can be appreciated only by determining the 2D affinity of the interacting partners. The 2D dissociation constant is defined by the law of mass action as $2D K_d = [\text{CD2}] \times [\text{CD58}] / [\text{CD2-CD58}]$, where $[\text{CD2}]$, $[\text{CD58}]$, and $[\text{CD2-CD58}]$ represent surface densities of free CD2, free CD58, and bound CD2-CD58 complexes in molecules μm^{-2} , with the 2D K_d in the same units (19, 20).

It has been difficult to obtain experimental data of suitable quality for 2D K_d measurements, both because of the technical difficulties of observing receptors in an adhesion between two cells and because of the problem of deconvolving adhesion receptor-counter-receptor interactions in the interface from cytoskeletal and membrane domain interactions with the receptors. A breakthrough was achieved in both respects by the application of supported planar bilayers (21). Here, glycosylphosphatidylinositol (GPI)-linked CD58 is fluorescently labeled and reconstituted into planar phospholipid bilayers to simulate the membrane of an APC, and T cell adhesion to CD58 is observed by accumulation of CD58 fluorescence at the cell-bilayer interface (22). This experimental method allows measurement of free bilayer CD58 density (molecules μm^{-2}), cell-bilayer contact area (μm^2), and the density of CD2-CD58 interactions in the contact area (molecules μm^{-2}) at equilibrium. We have also incorporated a linearized plot for calculation of the 2D K_d of the CD2-CD58 interaction and the total number of laterally mobile CD2 molecules on the cell surface (23). We refer to *affinity* as the inverse of the 2D K_d measured by the Zhu-Golan analysis (22, 24), and we refer to *avidity* as the total adhesive strength of the interaction. Two previous studies utilized this experimental and analytical approach and found that the CD2-CD58 interaction should result in strong adhesion, with a 2D K_d of 1.1–7.6 molecules μm^{-2}

for resting Jurkat T cells adhering to model bilayers reconstituted with laterally mobile CD58 at 24 °C (23, 25). A comparison of this 2D K_d with the physiological densities of CD2 and CD58 on T cell and APC surfaces, respectively, suggests that >85% of CD2 molecules are bound to CD58 in the contact area.

Here we apply our T cell-planar bilayer system to identify and characterize the mechanisms that regulate the avidity of T cell CD2 for its counter-receptor CD58. We quantify the number of receptors, receptor affinity, and receptor lateral mobility on both resting and activated T cells. By using quantitative fluorescence imaging, fluorescence photobleaching recovery (FPR), and single particle tracking (SPT), we observe that cell activation causes an increased rate of CD58 accumulation and an increased density of CD58 at the contact area. This increased accumulation is due to increases in cell surface CD2 expression and receptor affinity. We also provide evidence that CD2 is selectively immobilized at the contact site, further enhancing the number of potential CD2-CD58 bonds in the adhesion zone.

RESULTS AND DISCUSSION

Accumulation of CD58 at the Contact Area. Jurkat T cells were used as a model, because CD2 expressed on these cells has a 2D affinity for CD58 similar to that of human peripheral blood lymphocytes (23). Jurkat cells adhered to bilayers containing FITC-CD58 and caused the local density of CD58 to increase at the contact areas. This interaction was specifically inhibited by the adhesion-blocking monoclonal antibody (mAb) TS2/18 (24), and it was consistent with the mass-action equation for CD2 binding to CD58. We tested the effect of cell activation on CD58 accumulation by comparing the accumulation of fluorescence in the contact area for control cells with that for cells stimulated with PMA. After equivalent incubation times, the accumulated fluorescence in the contact area was greater for PMA-stimulated cells than for control cells (Figure 1, panels a and b). Control cells showed a plateau of accumulated CD58 after 30–40 min of incubation (Figure 1, panel c),

consistent with previous results (22, 25). Introduction of PMA after this time period caused an additional increase in CD58 accumulation. Similarly, cells stimulated with PMA for 20 min before incubation with CD58-containing bilayers showed a higher density of accumulated CD58 at equilibrium and required longer for the density to reach a plateau. These experiments indicated that the cell surface expression, binding affinity, or lateral mobility of CD2 molecules or a combination of these was altered on PMA-stimulated cells. The rate of CD58 accumulation was also 2-fold faster in PMA-treated cells than in control cells, probably due to the higher CD2 surface density on the activated cells (*vide infra*).

Consistent with the mass-action equation, the initial density of CD58 in the bilayer was an important determinant of the level of CD58 accumulation at equilibrium. We determined the plateau level of CD58 accumulation as a function of initial CD58 density for control and PMA-treated cells (Figure 2, panel a). PMA stimulation caused the equilibrium density of CD58 in the contact area to increase. The maximum density of CD58 was ~ 650 molecules μm^{-2} for control cells and ~ 950 molecules μm^{-2} for PMA-activated cells. Cell stimulation did not produce a significant change in the contact area of the cells: the maximum contact area size was $65 \mu\text{m}^2$ per cell for resting cells and $68 \mu\text{m}^2$ per cell for PMA-activated cells. From these data, we confirmed that CD58 accumulation (signifying the accumulation of CD2–CD58 complexes) depended directly on the initial CD58 density and determined that CD58 accumulation also depended on the activation state of the cell.

Lateral Mobility of CD58 in the Bilayer. Accumulation of CD58 in the contact area could have been related to a change in the lateral mobility of CD58 in the bilayer as a result of binding to CD2. We used FPR to measure the mobility of CD58 within and outside the cell–bilayer contact area (Table 1). FPR was performed under conditions that did not deplete FITC–CD58 in the contact area, such that diffusion of CD58 from the bilayer into the contact area was not a limiting factor. The diffusion coefficient (D) of FITC–CD58 in bilayer regions outside the contact area was high, as expected for diffusion of a GPI-linked protein in a model bilayer membrane. Within the area of contact with resting cells, the diffusion coefficient of CD58 was reduced. Treatment of cells with PMA further slowed CD58 diffusion in the contact area. The fractional mobility (f) of FITC–CD58 showed a similar trend: virtually no immobile CD58 was observed outside

the contact area ($f = 94\%$), while progressive (albeit modest) CD58 immobilization was found within the area of contact with resting and PMA-activated cells. Therefore, dynamic interaction with CD2 on the cell surface caused reduced mobility of CD58 in the contact area, as previously observed (22).

Lateral Mobility of CD2 on Cells. CD2 was labeled with the FITC-conjugated non-adhesion-blocking mAb CD2.1 to determine the lateral mobility of CD2 on cells adherent to bilayers reconstituted with unlabeled CD58. A low initial CD58 density was used to minimize migration of CD2 to the contact area while allowing cells to adhere to the bilayer. Measurements of CD2 lateral mobility were performed at the surface of the cell in contact with the bilayer and at the opposite surface of the cell not in contact with the bilayer (Table 1). The diffusion coefficient of CD2 was $(6-7) \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$ in both resting and PMA-treated cells. This observation was consistent with previous results (26) and with the diffusion coefficient of other adhesion receptors on Jurkat cells, *e.g.*, the lateral diffusion coefficient of LFA-1 is $(5-7) \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$, depending on the mAb used to label the receptor (27, 28).

The fractional mobility of CD2 was high outside the contact area, in resting and PMA-treated cells. In resting cells, CD2 fractional mobility was also high within the contact area, within experimental error of that outside the contact area (Table 1). However, in PMA-treated cells, the fractional mobility of CD2 within the contact area showed a dramatic and time-dependent decrease (Table 1). To characterize the time dependence of this immobilization, we measured CD2 fractional mobility in the area of contact with activated cells as a function of time (Figure 2, panel b). The fractional mobility of CD2 was low for the first 40 min after cell activation. The f value then increased gradually to 50% 120 min after PMA treatment. These results suggested that i) the cell recognized the ligation of CD2 by CD58, ii) the cell selectively immobilized the bound CD2 only when the cell was activated, and iii) the cell gradually released the immobilized CD2 over a time period of several hours. Furthermore, the combination of profound CD2 immobilization and modest CD58 immobilization was consistent with the rapid off rate for CD2–CD58 binding (16–18), and the slowing of CD58 diffusion in the context of cell activation was consistent with a reduced off rate (increased affinity) for CD2–CD58 binding. The selective immobilization of CD58-bound CD2 was likely due to

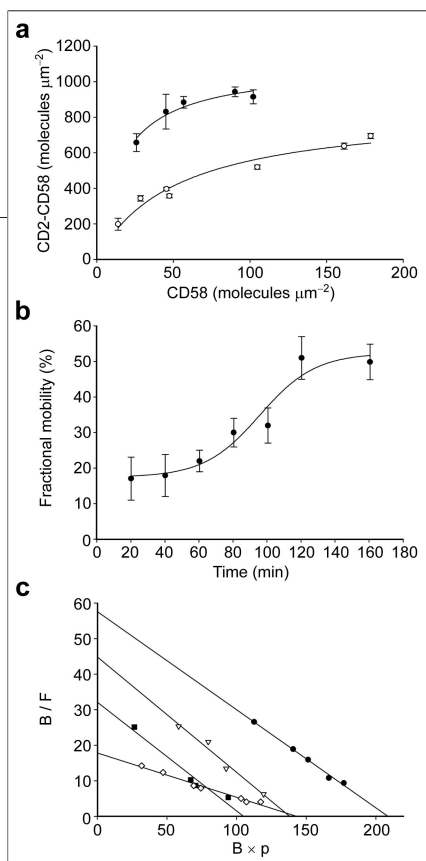


Figure 2. Effect of PMA activation on CD58 accumulation, CD2 lateral mobility, and CD2-CD58 affinity. a) Accumulation of bound CD58 molecules in the contact area as a function of CD58 density in the bilayer. Cells were incubated for 50 min with bilayers containing varying initial densities of CD58, and the density of bound CD58 (reflecting the density of CD2-CD58 complexes) was measured as described in the text: (○) control cells; (●) cells pre-incubated with 150 nM PMA for 20 min. Error is shown as SEM; $n = 20-130$ cells per data point. b) Time dependence of recovery from CD2 immobilization induced by the combination of CD2-CD58 binding and cell activation. Jurkat cells were incubated with bilayers containing 100-270 molecules μm^{-2} CD58, and 150 nM PMA was added immediately thereafter (time zero). CD2 was labeled with mAb FITC-CD2.1, and FPR was used to determine the fractional mobility (f) of CD2 in the contact area. Data points represent the mean values of f for 9-28 cells from 2-3 experiments at each point. Error is shown as SEM. c) Determination of 2D affinity ($2D K_d$) of CD2-CD58 interaction. The Zhu-Golan plot (B/F vs $B \times p$) was used to determine the 2D affinity of adhesion in control and activated Jurkat cells (23) (summarized in Table 2). Data are shown for the following: (◇) control cells; (●) PMA-activated cells; (▽) OKT3-GaM-treated cells; (■) CD2.1-treated cells.

increased association of CD2 with the T cell cytoskeleton, and the slow release of immobilized CD2 was likely caused by time-dependent changes in the strength of the CD2-cytoskeleton association (*vide infra*). The time course of immobilized CD2 release could be consistent with that of cellular deadhesion required for rapid scanning of APCs by migrating T cells (29).

We also tested whether different modes of cell stimulation could affect CD2 lateral mobility in the absence of ligation by CD58 (Table 2). PMA activates the cell through

result in a significant reduction in fractional mobility.

Number and 2D Affinity of CD2 Molecules on Cells.

FPR and quantitative fluorescence imaging experiments showed that cell activation alters the fractional mobility and density of CD2 in the contact area. We reasoned that the number of cell surface CD2 molecules and the 2D affinity of the CD2-CD58 interaction could also have a role in avidity regulation. We used the Zhu-Golan analysis (described in Methods and elsewhere (23, 24)) to measure the number and affinity of CD2 receptors in the

contact area. For a range of initial CD58 densities in the glass-supported planar bilayer, plots of B/F vs $B \times p$ (see eq 2) were constructed from experimental measurements of bound CD2-CD58 complex density (B), free CD58 density (F), CD2 fractional mobility (f), and ratio of contact area to cell surface area (p). Cell surface area was calculated as

TABLE 1. Lateral mobility of CD58 and CD2

FITC-labeled molecule	Treatment	Region	f (%) ^a	D [$\times 10^{-10}$ $\text{cm}^2 \text{s}^{-1}$] ^b	N
FITC-CD58 ^c	Control	Outside contact area	94 \pm 2	26.8 \pm 3.7	20
FITC-CD58 ^c	Control	Contact area	87 \pm 2	9.8 \pm 0.8	32
FITC-CD58 ^c	PMA-activated	Contact area	81 \pm 2	5.1 \pm 0.5	53
FITC-CD2.1 Ab ^d	Control	Top of cell ^e	77 \pm 6	5.9 \pm 0.7	18
FITC-CD2.1 Ab ^d	Control	Contact area	72 \pm 4	6.9 \pm 0.7	35
FITC-CD2.1 Ab ^d	PMA-activated	Top of cell ^e	74 \pm 4	7.3 \pm 0.9	25
FITC-CD2.1 Ab ^d	PMA-activated	Contact area	15-50	6.9 \pm 0.6	12

^aFractional mobility of the FITC-labeled molecule in the indicated region. Values represent mean \pm SEM. ^bDiffusion coefficient of the FITC-labeled molecule in the indicated region. For comparison, the diffusion coefficient of NBD-labeled phosphatidylethanolamine in egg PC planar bilayers was $80 \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$. Values represent mean \pm SEM. ^cThe initial density of FITC-CD58 in the glass-supported planar lipid bilayer was 200 molecules μm^{-2} . ^dThe initial density of CD58 in the bilayer was 30 molecules μm^{-2} for measurements at top of cell and 100-270 molecules μm^{-2} for measurements at contact area. Cell surface CD2 was labeled with FITC-CD2.1. ^eFocal plane was 10 μm above the plane of cell-bilayer contact.

$S_{\text{cell}} = 4\pi r^2 \times 1.8$, where r was the measured radius of the cell and 1.8 was a correction factor for surface roughness (30). We confirmed the validity of the S_{cell} equation by osmotic swelling experiments (25). Jurkat cells had a mean cell surface area of $700 \mu\text{m}^2 \text{ cell}^{-1}$.

In the Zhu–Golan plots (Figure 2, panel c), the X-intercept was used to determine the number of CD2 receptors (N_t) according to eq 3. The number of CD2 receptors per cell was also determined using radiolabeled mAb binding (Table 2). Because the mAb-binding method averages over all cells in a sample, whereas the Zhu–Golan method examines only cells that bind to CD58-containing bilayers, the latter method is expected to be biased toward cells expressing high levels of CD2. Nonetheless, the measurements obtained using the two methods were within 1.5–2.5-fold of one another; as expected, estimates of N_t obtained using the mAb-binding method were systematically lower than those determined by the Zhu–Golan method. N_t values measured using the Zhu–Golan method showed that resting and CD2.1-labeled cells expressed similar numbers of CD2 molecules on the cell surface, whereas cells activated with PMA or OKT3–GaM showed a 1.6-fold increase in CD2 surface expression.

In a Zhu–Golan plot, the negative reciprocal of the slope provides the 2D affinity ($2D K_d$) of the CD2–CD58 interaction (23, 25). All three cell activation treatments resulted in a 2.5-fold decrease in the $2D K_d$ (Table 2). Taken together, we found that cell activation by PMA or OKT3–GaM induced an increase in CD2 expression and that cell activation by PMA, OKT3–GaM, or CD2.1–CD58 caused an increase in receptor affinity. At equilibrium, the change in affinity would result in an increase from 87% to 96% of bound CD2 in the T cell–APC contact area. We concluded that increases in both 2D affinity and receptor expression were used by the cell to increase the avidity of the CD2–CD58 interaction. The magnitude and direction of this shift in CD2 avidity upon cell activation have been observed previously (15).

Our results are consistent with the involvement of a conformational change in CD2 as the basis for the increase in CD2 affinity, because treatment of T cells with the activating anti-CD2 mAb pair TS2/18 and CD2.1 produces an increase in affinity identical to that associated with cell activation through the PKC or TCR pathways (15, 31). Conformational changes in CD2 have

TABLE 2. Fractional mobility, surface expression, and 2D affinity of CD2

Treatment	$f(\%)^a$	$N_t[\times 10^4 \text{ molecules cell}^{-1}]$			r^2
		mAb assay	Zhu–Golan analysis	2D K_d^c	
Control	73 ± 3	4.3–8.6	13.2	8.1	0.97
PMA	70 ± 4	4.0–8.0	20.9	3.6	0.99
OKT3–GaM	35 ± 2	^b	21.0	3.3	0.96
CD2.1	74 ± 4	^b	13.1	3.1	0.97

^aMean \pm SEM from >40 cells determined in >3 experiments. CD2 was labeled with FITC–TS2/18. ^bNot determined. ^cMeasured in molecules μm^{-2} .

been observed by antibody labeling, immunofluorescence, and NMR studies (32–36). Other potential mechanisms for altered affinity, such as changes in the topology or intermembrane separation of the T cell–bilayer interface, are unlikely to be important here because the interface is flat, uniform, and stable when the bilayer is reconstituted with CD58 (36).

Changes in CD2–CD58 affinity are likely to have a fundamental role in adhesion strengthening. A low $2D K_d$ results in a small confinement region of the adhesion receptors (19, 25). Our findings indicate that the resting T cell holds CD2 in a lower affinity conformation that is capable of engaging in an initial “loose” interaction with APCs and that cell stimulation (*e.g.*, by TCR engagement) converts CD2 from the lower affinity conformation to a higher affinity form that mediates tighter adhesion through confinement of both receptors. We note that our initial report of CD2–CD58 affinity used a different clone of Jurkat cells than that used here (25); our current data (23, and the present study) suggest that the earlier clone could have been activated. Cell-based functional studies have found that PMA activation of a T cell hybridoma enhances by 2.5-fold the avidity of adhesion mediated by the CD2–CD58 interaction (15).

The activation-induced increase in the number of CD2 molecules at the cell surface occurs within 1 h of cell activation, and it is therefore unlikely that synthesis of new CD2 molecules is responsible for the increase (37). Thatté *et al.* (38) (and others) have shown that transmembrane receptors are stored in cytoplasmic vesicles and that disruption of the microtubule network with colchicine can induce complete surface expression of the stored protein within 60 min. Treatment of Jurkat cells with $10 \mu\text{M}$ colchicine for 30 min causes a 60%

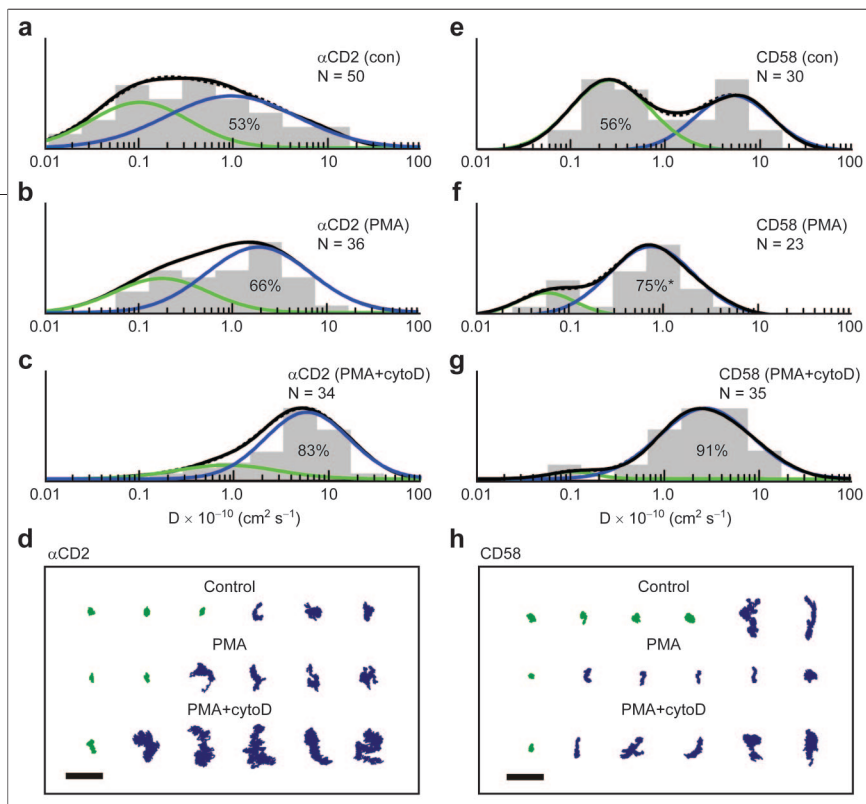


Figure 3. SPT of CD2. The diffusion of CD2 on Jurkat cells was observed by high-speed SPT. Beads were labeled with either mAb TS2/18 F(ab)' (a–d) or purified GPI–CD58 (e–h). Cells were treated for 30 min with buffer containing the following: a, e) DMSO (0.1%); b, f) PMA (150 nM, 0.1% DMSO); or c, g) cytoD and PMA (5 μM and 150 nM, 0.1% DMSO). Tracking data were analyzed to determine the lateral diffusion coefficients corresponding to the trajectories of individual particles; these data are shown as a histogram (gray) and as a smoothed population density function (black line). The best fit of the population density is a dotted line, and the best-fitted subpopulations are green and blue lines. The relative fraction of the dominant subpopulation is indicated. Representative trajectories (2 s, 1000 FPS) are shown for control, PMA-treated, and cytoD+PMA-treated cells (d, h). * denotes a subpopulation showing directed motion. Scale bar = 1 μm .

increase in the surface expression of CD2 as measured by FITC–TS2/18 labeling (data not shown). Longer incubations with colchicine do not increase the intensity of FITC–TS2/18 fluorescence at the cell surface, indicating that the CD2 reservoir is depleted with 30 min of incubation. These results suggest that T cells retain a cytoplasmic pool of CD2 that can be rapidly translocated to the cell surface in order to increase CD2-mediated adhesion.

SPT of CD2 on Cells. Our FPR experiments suggested that the average lateral mobility of CD2 was regulated by cell activation and by ligand engagement. We employed SPT to provide a non-ensemble measurement of CD2 mobility on Jurkat cells (28). Measurements of lateral mobility using FPR are proportional to the average mobility of the population over relatively long times and distances and are therefore biased toward faster diffusion. In contrast, SPT measurements sample the population of molecules over shorter times and distances and are therefore more sensitive than FPR to slow diffusion. In addition, SPT is capable of resolving the presence of multiple subpopulations of molecules (27). Because SPT resolves the immobile population observed in FPR experiments, SPT and FPR often measure different mean values for lateral diffusion (27). Nonetheless, the two methods are complementary because they provide information about the properties of molecules on different time and distance scales (28). All SPT experiments labeled CD2 by using 1 μm beads conjugated to either mAb TS2/18 or purified GPI–CD58,

and all observations were made at the top surface of Jurkat cells mounted on glass coverslips. Complete SPT results are given in the supporting information (See Supplementary Tables 1–3).

The average lateral diffusion of CD2 observed by SPT was ~ 3 -fold slower than that measured by FPR (Table 3), consistent with measurements on other cell surface proteins (28). Beads labeled with TS2/18 showed a small increase in average diffusion upon cell activation. Subpopulation analysis showed that there were two populations of trajectories within these data sets (Figure 3). Consistent with the FPR results, there was a significant slowly diffusing subpopulation in both resting and activated cells. The slow population was reduced upon cell activation; this change could have been responsible for the small increase in diffusion coefficient observed by FPR upon PMA stimulation of cells (Table 1). Because FPR measurements of lateral diffusion are biased toward faster diffusing receptors, we interpret the population of SPT trajectories centered at diffusion coefficients greater than $\sim (0.5\text{--}1.0) \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$ to approximate the mobile fraction measured by FPR. The fractions of faster-diffusing molecules in control and PMA-treated cells by SPT were reasonably consistent with FPR measurements of CD2 fractional mobility under the same conditions. (See ref 39 for a more extended discussion of the differences between measurements of lymphocyte receptor lateral mobility by FPR and SPT.)

The bead used to label CD2 in SPT experiments was also employed to mimic a small area of CD2–CD58

contact. Beads were conjugated to purified CD58 with a stoichiometry of ~ 5000 sites μm^{-2} and then used in SPT measurements on Jurkat cells. On resting cells, CD58-ligated CD2 showed a modest increase in average lateral diffusion relative to antibody-labeled CD2. Upon treatment with PMA, however, CD58-ligated CD2 showed significantly reduced lateral diffusion.

Subpopulation analysis demonstrated that CD58-ligated CD2 was found in two major populations on resting cells. Cell activation changed the diffusion profile of CD58-ligated CD2: a small immobile population remained, but the major population showed an intermediate diffusion coefficient and evidence of directed motion (Figure 3). This analysis by SPT confirmed the FPR observation that CD2 mobility was synergistically affected by the combination of CD2–CD58 ligation and cell activation.

We also tested whether the effect of cell activation on CD2 lateral mobility was mediated by cytoskeletal interactions. Treatment of Jurkat cells with PMA and the actin cytoskeleton disrupting cytochalasin D (cytoD) revealed that the immobile populations of both antibody-labeled and CD58-ligated CD2 were constrained by interactions with the cytoskeleton (Figure 3). In both cases, cytoD treatment shifted the CD2 molecules into a single rapidly diffusing population. CytoD also increased the diffusion coefficients of CD2 in both cases (Table 3). The sensitivity of CD2 mobility to treatment with cytoD suggested that these molecules are attached directly or indirectly to the actin cytoskeleton. It is likely that this interaction is mediated through binding of the cytoplasmic tail of CD2 to the CD2 associated protein (CD2AP). CD2AP interacts dynamically with the proline-rich CD2 cytoplasmic tail (6), and it also contains a high-affinity binding site for the actin capping protein capZ (40).

TABLE 3. Single particle tracking of CD2

Label	Treatment	N	Mean D^a	Major subpopulations ^b			
				D_{immobile}	Percent	D_{mobile}	Percent
TS2/18	Control	50	1.3 ± 0.3	0.1	47	1.0	53
TS2/18	PMA	36	2.0 ± 0.4	0.2	34	2.0	66
TS2/18	PMA+cytoD	34	6.7 ± 1.0	0.9	17	6.2	83
CD58	Control	30	2.8 ± 0.7	0.3	56	5.4	44
CD58	PMA	23	0.9 ± 0.2	0.1	25	0.8	75
CD58	PMA+cytoD	35	3.9 ± 0.7	0.1	9	2.8	91

^aDiffusion coefficients are in units of $10^{-10} \text{ cm}^2 \text{ s}^{-1}$, and error is given as SEM. ^bThe two largest populations are shown for each condition, and diffusion coefficients are in units of $10^{-10} \text{ cm}^2 \text{ s}^{-1}$.

Model for Cellular Regulation of CD2 Avidity. The avidity of cell adhesion depends on a number of different variables, and new methods are needed to dissect quantitatively the mechanisms that contribute to adhesion strength (avidity). We have employed a model adhesion system and a new receptor–ligand binding analysis to determine the 2D affinity of the CD2–CD58 interaction in a physiological environment. Our analysis takes account of the dynamic nature of the adhesion by incorporating experimental measurements of receptor lateral mobility (22, 24). We observe the effects of cell activation on receptor mobility using FPR and SPT. The two methods find similar results: CD2 mobility is only slightly affected by activation of the cell or by ligation to CD58, but there is a synergistic effect on lateral mobility when CD2 on activated cells is bound to CD58. We also find that cell activation increases the affinity of CD2 for CD58 by 2.5-fold and the number of CD2 receptors at the cell surface by 1.5-fold. Increased affinity, increased receptor density, and decreased lateral mobility at the contact area combine to make CD2–CD58-mediated adhesion stronger and more selective (Figure 4). Our results provide the most complete analysis of CD2-mediated adhesion to date and demonstrate that CD2–CD58 avidity is regulated by the conjunction of cell activation and contact area; receptor conformational change and affinity; and receptor and ligand lateral mobility, expression level, and surface density.

METHODS

Cell Culture and Reagents. The Jurkat E6.1 T leukemia cell line (American Type Culture Collection) was maintained in RPMI

1640 medium (Sigma) supplemented with 10% heat-inactivated fetal bovine serum (Sigma) and 2 mM L-glutamine/100 units mL^{-1} penicillin/0.1 mg mL^{-1} streptomycin (Sigma) at 37 °C in a

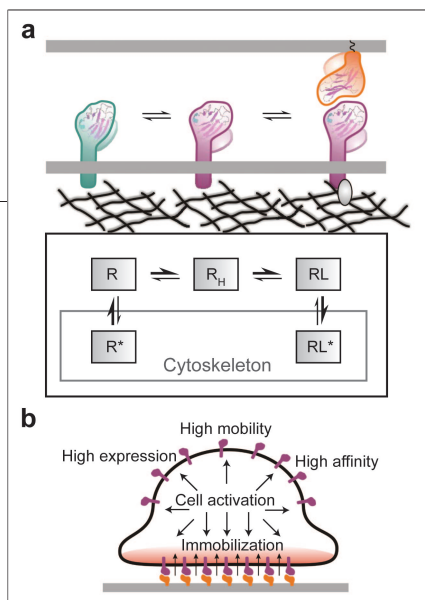


Figure 4. Model of CD2–CD58 avidity regulation. We propose a comprehensive model of CD2–CD58 avidity regulation that incorporates receptor affinity ($2D K_d$), number (N_d), and lateral mobility (f). **a** CD2 (R) is primarily mobile ($f \approx 50\text{--}70\%$) on resting cells, but there is a significant population of molecules that interacts with the cytoskeleton (R^*). Activation of the cell by PMA or by TCR/CD3 cross-linking induces a conformational change that increases the affinity of CD2 for CD58 by ~ 2.5 -fold (R_H). The mobile, high-affinity CD2 molecules can diffuse laterally on the cell surface and engage CD58 within a contact area (RL). The bound receptor is preferentially immobilized on activated cells to promote adhesion (RL^*), likely through association with CD2AP–capZ–actin complexes. **b** Cellular activation results in increased expression, affinity, and lateral mobility of unligated CD2. At the contact area, CD2–CD58 ligation generates a second signal, likely through a second conformational change in CD2, that immobilizes the receptor selectively in the contact area. High lateral mobility allows receptors outside the contact area to diffuse into the contact area and engage CD58, thereby increasing the number of bound CD2–CD58 complexes and promoting adhesion. This model couples passive diffusion with active recognition of conformational change to strengthen adhesion.

humidified atmosphere of 5% CO_2 . CytoD and PMA were purchased from EMD Biosciences. Other reagents were purchased from Sigma-Aldrich; cell media and buffers were purchased from Invitrogen. For SPT experiments, mAb TS2/18 was obtained from Endogen and digested to F(ab)' fragments by standard procedures. GPI-linked CD58 (GPI–CD58) was purified from human red cells as previously described (25).

Antibody Binding Assays. Jurkat cells were washed twice with N -(2-hydroxyethyl)piperazine- N' -ethanesulfonic acid (HEPES)-buffered saline containing 2% bovine serum albumin (BSA) and resuspended to 5×10^6 cells mL^{-1} . Cells were incubated with 2.3–230 nM iodinated mAb, in the presence and absence of 1.33 μM cold antibody, for 1 h at 4 $^\circ\text{C}$. TS2/18 and 35.1 mAbs displayed saturation binding to cells. Because this binding could have represented bivalent or monovalent binding to cell surface CD2 molecules, mAb molecular weights of 75,000 and 150,000 were used to calculate the limits of a range of specific activities. Bound and free counts were separated by centrifugation at $4000 \times g$ for 3 min through an oil cushion (1.5 parts dibutylphthalate to 1 part dioctylphthalate (47)) in prechilled microsediment tubes. The cell pellet and supernatant were separated and counted. Scatchard analysis was used to calculate the average number of CD2 sites per cell, as described (24).

Preparation of Planar Bilayers Reconstituted with FITC–CD58. GPI–CD58 was conjugated to FITC as previously described (25). To remove additional free FITC, the FITC–CD58 conjugate was subjected to ultrafiltration using a 30,000 MW cut off membrane (Centricon, Amicon) immediately after elution from the affinity column.

Unilamellar liposomes were prepared from egg phosphatidylcholine (PC) (Avanti Polar Lipids). Liposomes with or without reconstituted CD58 were prepared by OG dialysis (22, 42–44), stored under an argon atmosphere at 4 $^\circ\text{C}$, and used to prepare glass-supported planar bilayers (22, 42–44). Briefly, 10 μL of liposome suspension was deposited at the bottom of a plastic Petri dish (50 mm \times 9 mm), and a cleaned glass coverslip was placed on the droplet to fuse the liposomes at the glass surface. The coverslip was incubated for 30 min at RT and then gently washed with binding buffer containing 25 mM HEPES (pH 7.4), 147 mM NaCl, 5 mM KCl, 0.8 mM MgCl_2 , 1.8 mM CaCl_2 , 5 mM D -glucose, and 1% BSA (Calbiochem). The coverslip was rinsed 3–5 times with binding buffer and then immersed in a tank containing 1 L of binding buffer. A stainless steel slide with a 15-mm-diam hole at the center was coated on both sides with vacuum grease and a clean blank coverslip was sealed to one side of the slide. The slide was immersed in the tank, and the bilayer-coated coverslip was carefully mounted to the other side of the slide, with the bilayer facing the inside of the hole. The slide assembly was removed from the tank and rinsed with distilled water. The blank coverslip was carefully removed to open the chamber, and cells or reagents were added. The coverslip was then replaced to seal the chamber. Planar bilayers were not

exposed to open air during the procedure. The presence of GPI–CD58 molecules in the bilayer was confirmed by the specific binding of fluorescent antibodies. We observed that all of the reconstituted FITC–CD58 molecules were oriented away from the glass surface (23).

Imaging and Analysis. Two-dimensional 108 $\mu\text{m} \times$ 108 μm fluorescence images of FITC–CD58 redistribution in egg PC bilayers were acquired using a laser-scanning confocal fluorescence microscope (Meridian ACAS 570) with a pinhole diameter of 225 μm and a step size of 0.6 μm . The laser excitation wavelength was 514 nm. The focus was adjusted to the plane of the bilayer as determined by the maximum fluorescence intensity. A quantitative measure of cellular autofluorescence and background fluorescence was obtained from cells in contact with bilayers containing unlabeled CD58. Cells did not bind to CD58-containing bilayers in the presence of the adhesion-blocking CD2 mAb TS2/18 (data not shown). The contact area of a cell with a bilayer (S_c) was defined as the area of the accumulated fluorescence under the cell; the fluorescence intensity threshold for each image was set equal to the sum of cellular autofluorescence/background fluorescence and the free FITC–CD58 fluorescence intensity (F), where F was measured from a cell-free area of the bilayer (*vide infra*). Contact area determinations using this method were in generally good agreement with those determined using interference reflection microscopy (25). The fluorescence intensity was converted to CD58 surface density using a standard curve of fluorescence intensity vs FITC–CD58 density, as determined by the binding of ^{125}I -labeled anti-CD58 mAb TS2/9. The free (or unbound) CD58 density (f) in the bilayer was determined by the fluorescence intensity in a cell-free area after subtraction of the background fluorescence of a bare glass coverslip. Cellular autofluorescence/background fluorescence and free FITC–CD58 fluorescence (F) were then subtracted from the fluorescence intensity in the cell–bilayer contact area to yield the surface density of bound CD58 (B).

FPR. Fluorescence photobleaching recovery (FPR) was used to measure the fractional mobility and lateral diffusion coefficient of FITC–CD58 in the planar bilayer membrane and of FITC–mAb-labeled CD2 in the T cell membrane. Experiments were performed using a confocal laser-scanning microscope (Meridian ACAS 570) as described (22, 25). FITC-labeled anti-CD2 mAbs TS2/18 (adhesion-blocking) and CD2.1 (non-adhesion-blocking) were used to label CD2 molecules on Jurkat cells. FPR measurements of CD2 mobility were performed at either the cell bilayer contact area or the upper surface of the cell. For the latter experiments, the focal plane of the microscope was first oriented at the cell–bilayer contact plane and then moved up by 10 μm from the plane of the bilayer to a region at the upper surface of the cell. (*N.B.*, the average diameter of Jurkat cells was 11.1 μm .) All measurements were carried out at RT.

2D Affinity Analysis. Scatchard analysis cannot be used to measure the affinity of receptors at a contact area in which the

receptors are accumulating over time. Instead, we have derived a new analysis that takes account of this dynamic condition, and we have successfully used the analysis to measure the 2D K_d of the CD2–CD58 interaction in resting Jurkat T cells (23, 24). Briefly, the analysis uses quantitative fluorescence imaging to determine both the number of receptors in the contact area and the size of the contact area. FPR is then used to determine the fraction of laterally mobile receptors. The formation of CD2–CD58 complexes in the contact area is described by the general Zhu–Golan equation:

$$\frac{B}{F} = \frac{N_t \times f}{K_d \times S_{\text{cell}}} - \frac{B \times p}{K_d} \quad (1)$$

where B is the density of bound CD58 in the contact area, F is the density of free CD58 in the bilayer, N_t is the total number of CD2 binding sites on the cell, f is the percentage of laterally mobile CD2 on the cell, K_d is the 2D dissociation constant, S_{cell} is the total surface area of the cell, and p is the ratio of the contact area (S_b) to S_{cell} (i.e., $p = S_b/S_{\text{cell}}$). ($N_t \times f$) represents the number of available binding sites, since immobile receptors are unable to migrate into the contact area. According to eq 1, we can plot the data derived from fluorescence imaging and FPR experiments according to the Zhu–Golan form:

$$\frac{B}{F} \text{ vs } B \times p \quad (2)$$

This plot generates a line with a negative reciprocal slope equal to the 2D K_d , and an x-intercept (X) that can be used to determine N_t from the following relationship:

$$N_t = \frac{(X \times S_{\text{cell}})}{f} \quad (3)$$

SPT. Jurkat cells were used for all SPT experiments. Cells were cultured as described above and labeled with beads as described elsewhere (27, 28). Briefly, cells were harvested from cultures in exponential growth and washed three times with Hank's balanced salt solution (HBSS) supplemented with 1% (m/v) BSA (HBSSB). The cells were then resuspended in HBSSB containing DMSO (0.1% v/v), DMSO with PMA (150 nM), or DMSO with cytochalasin D (cytoD) (5 $\mu\text{g mL}^{-1}$). Aliquots of the cell suspension (0.25 mL) were incubated for 30 min at 37 °C and then labeled with polystyrene microspheres (*vide infra*) for another 15 min at 37 °C. The samples were then diluted to 1 mL with HBSSB and transferred to a 24-well plate containing 12 mm circular coverslips treated with Cell-Tak (BD Pharmingen). The plate was centrifuged at \sim 500 rpm for 7 min, and then the wells were carefully washed seven times with fresh HBSSB (1 mL). One of the coverslips was then transferred to a microslide, seated using a thin circle of vacuum grease, and sealed with Cytoseal 60 (Richard-Allan Scientific). Samples were observed at 37 °C within 90 min of sealing.

One-micrometer polystyrene microspheres were obtained from Polysciences (2.6% m/v). The beads were diluted to a stock solution of 1.3% (m/v) in deionized water with 0.1% NaN_3 and sonicated for 15 min before each use. Beads (10 μL of stock solution) were labeled by incubation with monoclonal F(ab)' fragments or purified GPI–CD58 (0.1–10 μg) for 1 h in 0.2 mL of borate buffer (100 mM borate, 1 mM EDTA, 0.1% NaN_3 , pH = 8.5) at a final bead concentration of 0.05% (m/v). In all cases, we used the minimum amount of protein required to achieve selective binding of beads to cells. Control beads were labeled with polyclonal F(ab)' or BSA under identical conditions. After adsorption of the protein to the beads, the samples were diluted to 1 mL with blocking buffer (10 mM HEPES, 140 mM NaCl, 1 mM

EDTA, 2% dextran, 1% BSA, 0.1% NaN_3 , 0.1 $\mu\text{g mL}^{-1}$ poly(ethylene glycol) compound, pH 7.4) and incubated for 1 h. The samples were then sonicated for 15 min and centrifuged at 5000 rpm for 7 min. The supernatant was aspirated to 0.1 mL final volume, resuspended, and sonicated for 15 min immediately before use. All labeled bead samples were used within 48 h of preparation. Selectivity of binding was confirmed at the beginning of each experiment by manually counting the number of positively labeled cells in \sim 20 random fields. Samples were used for tracking experiments if the selectivity of bead binding was $>$ 4-fold greater than the negative control. Bead binding was typically 0.4% of cells for control beads, 2.8% (7-fold selectivity) for TS2/18-labeled beads, and 2% (5-fold selectivity) for CD58-labeled beads.

Cells were observed on a Nikon TE2000-E microscope equipped with DIC optics using a 60 \times oil objective with an oil condenser (NA = 1.4) (45). Images were captured at 1000 FPS using a Fastcam Super 10K camera (Photron). Video data were processed using Metamorph (Universal Imaging) and converted to trajectories. Trajectory data were analyzed using mean square displacement analysis (28) implemented in a custom program written in Matlab (Mathworks) as described (27).

Acknowledgments: C.W.C. was funded by NIH NRSA Postdoctoral Fellowship F32 GM067292 and NIH Training Grant T32 HL07623. D.E.G. acknowledges funding from NIH Grants HL32854 and HL70819, and M.L.D. acknowledges funding from NIH Grant AL43542.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Bachmann, M. F., Barner, M., and Kopf, M. (1999) CD2 sets quantitative thresholds in T cell activation, *J. Exp. Med.* **190**, 1383–1391.
- Shaw, S., Luce, G. E. G., Quinones, R., Gress, R. E., Springer, T. A., and Sanders, M. E. (1986) Two antigen-independent adhesion pathways used by human cytotoxic T-cell clones, *Nature* **323**, 262–264.
- Dustin, M. L., Sanders, M. E., Shaw, S., and Springer, T. A. (1987) Purified lymphocyte function-associated antigen-3 binds to CD2 and mediates lymphocyte-T adhesion, *J. Exp. Med.* **165**, 677–692.
- Moingeon, P., Chang, H. C., Wallner, B. P., Stebbins, C., Frey, A. Z., and Reinherz, E. L. (1989) CD2-mediated adhesion facilitates lymphocyte-T antigen recognition function, *Nature* **339**, 312–314.
- Tibaldi, E. V., Salgia, R., and Reinherz, E. L. (2002) CD2 molecules redistribute to the uropod during T cell scanning: Implications for cellular activation and immune surveillance, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7582–7587.
- Dustin, M. L., Olszowy, M. W., Holdorf, A. D., Li, J., Bromley, S., Desai, N., Widder, P., Rosenberger, F., van der Merwe, P. A., Allen, P. M., and Shaw, A. S. (1998) A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts, *Cell* **94**, 667–677.
- Wild, M. K., Cambiaggi, A., Brown, M. H., Davies, E. A., Ohno, H., Saito, T., and van der Merwe, P. A. (1999) Dependence of T cell antigen recognition on the dimensions of an accessory receptor–ligand complex, *J. Exp. Med.* **190**, 31–41.
- Bromley, S. K., Burack, W. R., Johnson, K. G., Somersalo, K., Sims, T. N., Sumen, C., Davis, M. M., Shaw, A. S., Allen, P. M., and Dustin, M. L. (2001) The immunological synapse, *Annu. Rev. Immunol.* **19**, 375–396.
- Doyle, C., and Strominger, J. L. (1987) Interaction between CD4 and class-II MHC molecules mediates cell-adhesion, *Nature* **330**, 256–259.
- Luescher, I. F., Vivier, E., Layer, A., Mahiou, J., Godeau, F., Malissen, B., and Romero, P. (1995) CD8 modulation of T-cell antigen receptor–ligand interactions on living cytotoxic T-lymphocytes, *Nature* **373**, 353–356.

11. Matsui, K., Boniface, J. J., Reay, P. A., Schild, H., Destroth, B. F., and Davis, M. M. (1991) Low affinity interaction of peptide-MHC complexes with T-cell receptors, *Science* **254**, 1788–1791.
12. O'Rourke, A. M., Appgar, J. R., Kane, K. P., Martz, E., and Mescher, M. F. (1991) Cytoskeletal function in CD8-cell and T-cell receptor-mediated interaction of cytotoxic lymphocytes-T with class-I protein, *J. Exp. Med.* **173**, 241–249.
13. Weber, S., Trauneker, A., Oliveri, F., Gerhard, W., and Karjalainen, K. (1992) Specific low-affinity recognition of major histocompatibility complex plus peptide by soluble T-cell receptor, *Nature* **356**, 793–796.
14. Irvine, D. J., Purbhoo, M. A., Krosggaard, M., and Davis, M. M. (2002) Direct observation of ligand recognition by T cells, *Nature* **419**, 845–849.
15. Hahn, W. C., Burakoff, S. J., and Bierer, B. E. (1993) Signal transduction pathways involved in T-cell receptor-induced regulation of CD2 avidity for CD58, *J. Immunol.* **150**, 2607–2619.
16. Davis, S. J., Ikemizu, S., Wild, M. K., and van der Merwe, P. A. (1998) CD2 and the nature of protein interactions mediating cell–cell recognition, *Immunol. Rev.* **163**, 217–236.
17. Dustin, M. L., Olive, D., and Springer, T. A. (1989) Correlation of CD2 binding and functional-properties of multimeric and monomeric lymphocyte function-associated antigen-3, *J. Exp. Med.* **169**, 503–517.
18. van der Merwe, P. A., Barclay, A. N., Mason, D. W., Davies, E. A., Morgan, B. P., Tone, M., Krishnam, A. K. C., Ianelli, C., and Davis, S. J. (1994) Human cell-adhesion molecule CD2 binds CD58 (LFA-3) with a very-low affinity and an extremely fast dissociation rate but does not bind CD48 or CD59, *Biochemistry* **33**, 10149–10160.
19. Bell, G. I., Dembo, M., and Bongrand, P. (1984) Cell-adhesion - competition between nonspecific repulsion and specific bonding, *Biophys. J.* **45**, 1051–1064.
20. Bell, G. I. (1978) Models for the specific adhesion of cells to cells, *Science* **200**, 618–627.
21. McConnell, H. M., Watts, T. H., Weis, R. M., and Brian, A. A. (1986) Supported planar membranes in studies of cell–cell recognition in the immune system, *Biochim. Biophys. Acta* **864**, 95–106.
22. Dustin, M. L., Ferguson, L. M., Chan, P. Y., Springer, T. A., and Golan, D. E. (1996) Visualization of CD2 interaction with LFA-3 and determination of the two-dimensional dissociation constant for adhesion receptors in a contact area, *J. Cell Biol.* **132**, 465–474.
23. Zhu, D. M., Dustin, M. L., Cairo, C. W., and Golan, D. E. (2006) Analysis of two-dimensional dissociation constant of laterally mobile cell adhesion molecules, *Biophys. J.*, in press.
24. Dustin, M. L. (1997) Adhesive bond dynamics in contacts between T lymphocytes and glass-supported planar bilayers reconstituted with the immunoglobulin-related adhesion molecule CD58, *J. Biol. Chem.* **272**, 15782–15788.
25. Dustin, M. L., Golan, D. E., Zhu, D. M., Miller, J. M., Meier, W., Davies, E. A., and van der Merwe, P. A. (1997) Low affinity interaction of human or rat T cell adhesion molecule CD2 with its ligand aligns adhering membranes to achieve high physiological affinity, *J. Biol. Chem.* **272**, 30889–30898.
26. Liu, S. J., Hahn, W. C., Bierer, B. E., and Golan, D. E. (1995) Intracellular mediators regulate CD2 lateral diffusion and cytoplasmic Ca²⁺ mobilization upon CD2-mediated T-cell activation, *Biophys. J.* **68**, 459–470.
27. Cairo, C. W., Mirchev, R., and Golan, D. E. (2006) Cytoskeletal regulation couples LFA-1 conformational changes to receptor lateral mobility and clustering, *Immunity* **25**, 297–308.
28. Saxton, M. J., and Jacobson, K. (1997) Single-particle tracking: Applications to membrane dynamics, *Annu. Rev. Biophys. Biomol. Struct.* **26**, 373–399.
29. Gunzer, M., Schafer, A., Borgmann, S., Grabbe, S., Zanker, K. S., Brocker, E. B., Kampgen, E., and Friedl, P. (2000) Antigen presentation in extracellular matrix: Interactions of T cells with dendritic cells are dynamic, short lived, and sequential, *Immunity* **13**, 323–332.
30. Mege, J. L., Capo, C., Benoliel, A. M., Foa, C., Galindo, R., and Bongrand, P. (1986) Quantification of cell-surface roughness - A method for studying cell mechanical and adhesive properties, *J. Theor. Biol.* **119**, 147–160.
31. Bockenstedt, L. K., Goldsmith, M. A., Dustin, M., Olive, D., Springer, T. A., and Weiss, A. (1988) The CD2 ligand LFA-3 activates T-cells but depends on the expression and function of the antigen receptor, *J. Immunol.* **141**, 1904–1911.
32. Li, J., Smolyar, A., SunderPlassmann, R., and Reinherz, E. L. (1996) Ligand-induced conformational change within the CD2 ectodomain accompanies receptor clustering: Implication for molecular lattice formation, *J. Mol. Biol.* **263**, 209–226.
33. Kitao, A., and Wagner, C. (2000) A space-time structure determination of human CD2 reveals the CD58-binding mode, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2064–2068.
34. Wang, J., Smolyar, A., Tan, K. M., Liu, J., Kim, M. Y., Sun, Z. J., Wagner, G., and Reinherz, E. L. (1999) Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors, *Cell* **97**, 791–803.
35. Wyss, D. F., Dayie, K. T., and Wagner, G. (1997) The counterreceptor binding site of human CD2 exhibits an extended surface patch with multiple conformations fluctuating with millisecond to microsecond motions, *Protein Sci.* **6**, 534–542.
36. Shaw, A. S., and Dustin, M. L. (1997) Making the T cell receptor go the distance: A topological view of T cell activation, *Immunity* **6**, 361–369.
37. Alberola-Ila, J., Places, L., Delacalle, O., Romero, M., Yague, J., Gallart, T., Vives, J., and Lozano, F. (1991) Stimulation through the TCR/CD3 complex up-regulates the CD2 surface expression on human lymphocytes-T, *J. Immunol.* **146**, 1085–1092.
38. Thatte, H. S., Bridges, K. R., and Golan, D. E. (1994) Microtubule inhibitors differentially affect translational movement, cell-surface expression, and endocytosis of transferrin receptors in K562 cells, *J. Cell. Physiol.* **160**, 345–357.
39. Dustin, M. L., Bivona, T. G., and Philips, M. R. (2004) Membranes as messengers in T cell adhesion signaling, *Nat. Immunol.* **5**, 363–372.
40. Hutchings, N. J., Clarkson, N., Chalkley, R., Barclay, A. N., and Brown, M. H. (2003) Linking the T cell surface protein CD2 to the actin-capping protein CAPZ via CMS and CIN85, *J. Biol. Chem.* **278**, 22396–22403.
41. Sayre, P. H., Hussey, R. E., Chang, H. C., Ciardelli, T. L., and Reinherz, E. L. (1989) Structural and binding analysis of a two-domain extracellular CD2 molecule, *J. Exp. Med.* **169**, 995–1009.
42. Brian, A. A., and McConnell, H. M. (1984) Allogeneic stimulation of cyto-toxic T-cells by supported planar membranes, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6159–6163.
43. Chan, P. Y., Lawrence, M. B., Dustin, M. L., Ferguson, L. M., Golan, D. E., and Springer, T. A. (1991) Influence of receptor lateral mobility on adhesion strengthening between membranes containing LFA-3 and CD2, *J. Cell Biol.* **115**, 245–255.
44. Mimms, L. T., Zampighi, G., Nozaki, Y., Tanford, C., and Reynolds, J. A. (1981) Phospholipid vesicle formation and transmembrane protein incorporation using octyl glucoside, *Biochemistry* **20**, 833–840.
45. Mirchev, R., and Golan, D. E. (2001) Single-particle tracking and laser optical tweezers studies of the dynamics of individual protein molecules in membranes of intact human and mouse red cells, *Blood Cells Mol. Dis.* **27**, 143–147.

Design and Characterization of an Active Site Selective Caspase-3 Transnitrosating Agent

Douglas A. Mitchell[†], Sarah U. Morton^{*}, and Michael A. Marletta^{†,§,¶,*}

[†]Department of Chemistry, [§]Department of Molecular and Cell Biology, and [¶]Division of Physical Biosciences, Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720-1460, and ^{*}Department of Biochemistry, University of California, San Francisco, California 94158

In recent years, it has become clear that nitric oxide (NO) targets proteins other than soluble guanylate cyclase (sGC) (1). S-Nitrosation, an example of sGC/cyclic guanosine monophosphate (cGMP)-independent NO signaling, is a post-translational modification in which a cysteine thiol is converted to a nitrosothiol by the oxidative addition of NO. Modification of a cysteine residue that directly participates in an enzymatic reaction leads to inhibition of activity. The apoptotic protease caspase-3 is one such protein, and a number of reports have recently begun to characterize the molecular mechanism of how NO inhibits the caspases and, thus, apoptosis (2–7). Procaspase-3 is S-nitrosated under basal conditions in human B- and T-cell lines on the catalytic cysteine, Cys163 (3). Similarly, the active site cysteine of procaspase-9 is S-nitrosated in human colon adenocarcinoma cells (8).

The limited supply of NO (nanomolar under signaling concentrations), coupled with its promiscuous reactivity towards both thiol and non-thiol functionality, argues for the existence of molecular machinery that confers reaction specificity (5). For example, NO reacts indiscriminately and rapidly with heme proteins, reactive oxygen, and nitrogen species (9). Also, when a NO donor or transnitrosating agent is allowed to react with purified proteins *in vitro*, many cysteines are modified to nitrosothiols, even when a modest concentration of nitrosating reagent is used. In the case of caspase-3, multiple cysteines are modified to nitrosothiols (5, 7); however, it is known that only one cysteine is modified in cells (3). The same can be said for the ryanodine receptor (10, 11). To date, poly-S-nitrosation has only been described in cells when an exogenous source of NO was applied. A reasonable hypothesis is that specificity is achieved by S-nitrosation being a protein-assisted process. Without

ABSTRACT The oxidative addition of nitric oxide (NO) to a thiol, S-nitrosation, is a focus of studies on cyclic guanosine monophosphate (cGMP)-independent NO signaling. S-Nitrosation of the catalytic cysteine of the caspase proteases has important effects on apoptosis and consequently has received attention. Here we report on a small molecule that can directly probe the effects of S-nitrosation on the caspase cascade. This chemical tool is capable of permeating the mammalian cell membrane, selectively transnitrosating the caspase-3 active site cysteine, and halting apoptosis in cultured human T-cells. The efficacy of this reagent was compared with the commonly used reagent S-nitrosoglutathione and an esterified derivative.

*Corresponding author,
marletta@berkeley.edu

Received for review September 12, 2006
and accepted October 11, 2006.

Published online November 3, 2006

10.1021/cb600393x CCC: \$33.50

© 2006 by American Chemical Society

assistance and regulation, accounting for previous observations proves difficult.

To further support a role for protein-mediated S-nitrosation and denitrosation reactions *in vivo*, the active site of caspase-3 is effective at excluding glutathione (GSH). Cellular concentrations of GSH (low millimolar) do not inhibit enzymatic activity and large excesses of GSH do not efficiently restore the activity of caspase-3-Cys163-SNO (5, 7, 12). Since the caspase active site nitrosothiol is not freely exchangeable with GSH, S-nitrosoglutathione (GSNO) reductase cannot be directly responsible for recovery of caspase-3 activity upon apoptotic stimulation (13–16). The mechanism of how cells that produce NO are capable of restoring caspase activity remains to be determined.

To aid the study of how S-nitrosation regulates the caspase cascade and apoptosis, we have developed and characterized a peptide-based reagent (PepSNO) designed to carry out a specific transnitrosation reaction with the catalytic cysteine of caspase-3 in apoptotic human T-cells. Other small molecule reagents, for example, GSNO and S-nitroso-N-acetyl penicillamine (SNAP), do not exhibit specificity with the multiple reactive cysteines within caspase-3 (5). Peptide nitrosothiol reagents designed from known protease specificities have been previously used in attempts to engineer faster-acting nitrosating agents for papain (17) and human rhinovirus 3C protease (HRV 3C) (18). The strategy worked in the case of HRV 3C where a nitrosating agent was developed that exhibited a second-order rate of deactivation 8-fold higher than GSNO. Specificity of the reagent for the catalytic cysteine over the other two cysteines was inferred but not explicitly demonstrated. The nitrosation rate enhancement in the case of papain was <2-fold. In both studies, the reagents were tested only on purified enzyme.

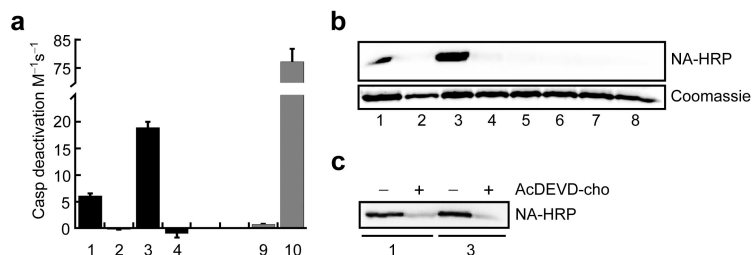
In this work, we implemented a strategy similar to that of Xian *et al.* (17, 18) and designed a peptide based on the optimized substrate recognition sequence of human caspase-3, DEVD (12). In the design employed here, the more tolerant P2 position was changed from valine to a thiol-containing amino acid (cysteine, C; homocysteine, hC; mercaptonorvaline, mnV; and penicillamine, penA) for subsequent conversion to the corresponding nitrosothiol. As anticipated, the distance from the electrophilic nitrogen of the nitrosothiol moiety to the nucleophilic sulfur of caspase-3 was a crucial design element. The crystal structure of caspase-3 bound to

both a peptide aldehyde (19) and a chloromethyl ketone inhibitor (20) indicated that the distance from the C_β of valine (P2 of the peptide) to the sulfur of caspase-3-Cys163 is 4.3 Å. The side chain of the nitrosothiol-containing amino acid must be capable of spanning this distance in order to react with the catalytic cysteine. The distances from C_β to the electrophilic nitrogen of the nitrosothiol in the extended conformation of S-nitroso-C, S-nitroso-hC, and S-nitroso-mnV are ~2.7, 3.9, and 5.1 Å, respectively.

The orthogonally protected L-amino acids C, hC, and penA were commercially available (Bachem, N-Fmoc, and S-Trityl) for use in solid-phase peptide synthesis, but the similarly protected mnV reagent was not. We therefore developed a synthetic route to L-5-[S-trityl]-[N-9-fluorenylmethyloxycarbonyl]-mnV starting from L-glutamic acid (Supplementary Figure 1). Fmoc chemistry was then used to synthesize a panel of peptides that fit the caspase-3 recognition motif, AcDEXD, where X is the thiol-containing amino acid (Supplementary Methods). The thiol-containing peptides were converted to nitrosothiols by reaction with acidified nitrite. The purified peptides were studied for the ability to deactivate recombinant human caspase-3 using a fluorimetric activity assay (5). As illustrated (Figure 1, panel a), both the S-nitroso-C and S-nitroso-hC containing peptides were capable of deactivating caspase-3, while the corresponding thiol peptides were not. Both the S-nitroso-C and S-nitroso-hC peptides displayed a rapid rate of caspase-3 deactivation when compared with GSNO (6- and 18-fold, respectively) but were not as efficient as thioredoxin-Cys73-SNO or an irreversible inhibitor (AcDEVd-chloromethylketone, $(1.4 \pm 0.1) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) under identical conditions. This suggested that the rate-determining step was the transnitrosation reaction, not peptide binding.

To further address the mechanism of deactivation, the biotin switch method was utilized to confirm S-nitrosation of caspase-3. When treated with the S-nitroso-C and S-nitroso-hC peptides, caspase-3 gave a strong signal on a horseradish peroxidase conjugated NeutrAvidin Western (NA-HRP), indicating S-nitrosation of the large subunit of caspase-3. Surprisingly, caspase-3 treated with the S-nitroso-mnV and S-nitroso-penA peptides did not yield a signal above controls (Figure 1, panel b). Additional evidence implicating transnitrosation as the mechanism for loss of activity, as opposed to active site occupation, was provided by a competitive inhibition

Figure 1. S-Nitrosopeptides designed from the caspase-3 recognition sequence selectively S-nitrosate the caspase-3 catalytic cysteine *in vitro*. a) Recombinant caspase-3 was treated with the peptide nitrosating reagents, and the activity was measured. All peptides were based on the caspase-3 AcDEXD recognition motif. Assignments for the X axis are as follows: 1, S-nitrosocysteine;



2, cysteine; 3, S-nitrosomocysteine; 4, homocysteine. Two transnitrosating agents not based on the AcDEXD motif were also studied under equivalent conditions: 9, S-nitrosoglutathione; 10, thioredoxin-C73-SNO. b) Caspase-3 was treated with the peptidic nitrosating agents and subjected to the biotin switch method. After SDS-PAGE and transfer to a nitrocellulose membrane, the blot was probed with NA-HRP. A loading control duplicate gel was stained with Coomassie. Assignments for AcDEXD are as follows: 1–4, same as in panel a; 5, S-nitrosomercaptonorvaline; 6, mnV; 7, S-nitrosopenicillamine; 8, penA. c) Caspase-3 was treated with a mixture of the indicated nitrosating agent and the reversible and competitive peptide-aldehyde inhibitor AcDEVD-cho.

experiment, in which the K_i for the hC-containing peptide was determined to be 75 μM (Supplementary Figure 2). This inhibition constant is too weak to account for the observed loss of activity. Furthermore, competitive inhibition cannot explain the observed time-dependent loss of activity. Finally, all of the AcDEXD reagents were tested for the ability to carry out an S-thiolation reaction (formation of a mixed disulfide with loss of HNO) with caspase-3-Cys163, but in no case was this product detected after trypsin digestion and matrix-assisted laser desorption/ionization (MALDI) analysis (data not shown). Previous small molecule studies provide a possible explanation where transnitrosation reactions were shown to be ~ 100 -fold faster than the related S-thiolation reaction (21, 22). Even so, our result was unexpected since similar conditions with GSNO led to detectable S-glutathiolation of caspase-3-Cys163 (5). Negative-mode MALDI and other sequencing-grade protease screens were also unsuccessful in detecting the AcDECD or AcDEhCD S-thiolation product. It is unlikely that the S-thiolation product was formed to a significant extent because MALDI readily detects the AcDECD and AcDEhCD peptides and the corresponding homodisulfides (data not shown).

Two methods were employed to demonstrate that the transnitrosation reaction was specific for the caspase-3 active site cysteine, Cys163. First, a potent yet reversible caspase-3 peptide aldehyde inhibitor, AcDEVD-cho, was added with either the S-nitroso-C or S-nitroso-hC peptides to caspase-3, and the biotin switch method was carried out as usual. As shown (Figure 1, panel c), addi-

tion of AcDEVD-cho significantly reduced the Western signal, consistent with S-nitrosation at the catalytic cysteine. Additionally, MALDI mass spectrometry (MS) analysis on trypsin-digested caspase-3 that had undergone the biotin switch verified specificity for Cys163 (Supplementary Figure 3). It should be noted that selectivity for the active site cysteine is lost if caspase-3 is treated with >3 equiv of the S-nitrosating agent. MALDI data confirmed that caspase-3 treated with 4 molar equiv of the S-nitroso-hC peptide formed a nitrosothiol on Cys264, a surface-accessible cysteine (data not shown) (20).

The S-nitroso-hC peptide was additionally screened against papain, a much less specific cysteine protease, and caspase-7 to probe reagent specificity with other cysteine nucleophile enzymes. Reaction of papain with the S-nitroso-hC peptide (3 molar equiv) did not inhibit proteolysis of bovine serum albumin (data not shown). Because of the high sequence and structural homology of caspase-7 to caspase-3 (23) and the fact that both proteases optimally cleave after DEVD (12), it was anticipated that the S-nitroso-hC peptide would deactivate caspase-7. This peptide did react with caspase-7, albeit 30-fold slower ($0.62 \pm 0.03 \text{ M}^{-1} \text{ s}^{-1}$) compared with caspase-3. The slower reaction of the S-nitroso-hC peptide with caspase-7 may stem from this enzyme being less efficient (k_{cat}/K_M (caspase-3) for AcDEVD-para-nitroanilide = $2.2 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$; caspase-7 = $3.7 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) (24).

Because of a faster transnitrosation reaction with caspase-3-Cys163, the S-nitroso-hC-containing peptide

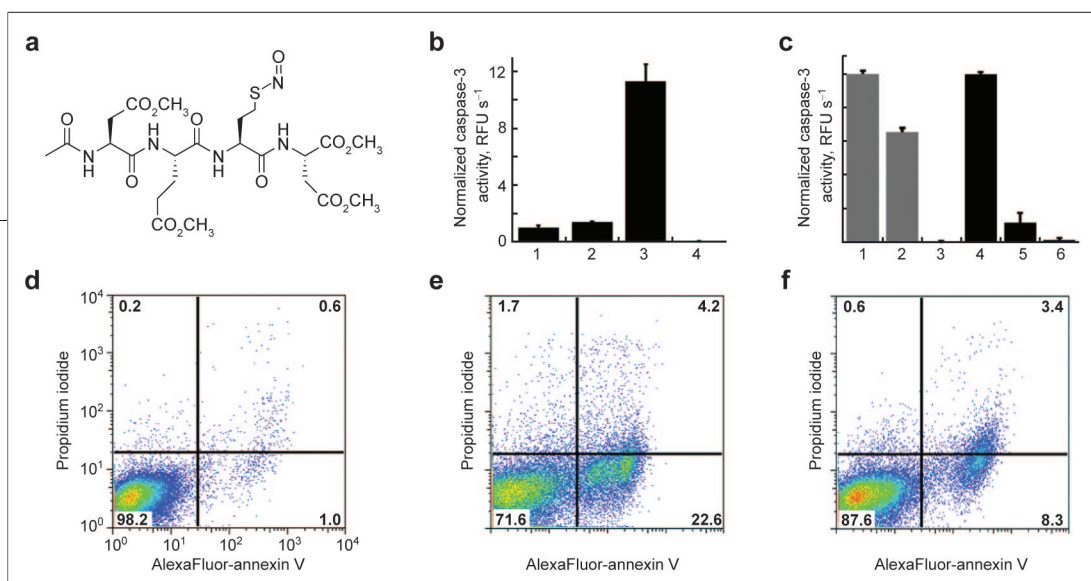


Figure 2. PepSNO inhibits apoptosis in Jurkat cells more effectively than GSNO. **a)** Structure of PepSNO. **b)** Etoposide ($2.5 \mu\text{g mL}^{-1}$) was added to Jurkat cells under reduced light conditions for 5 h to induce apoptosis before the addition of peptide reagent (1 mM). The peptide reagents were allowed to react with the cells for an additional 5 h at 37°C before harvesting, lysing, and measurement of the AcDEVD-ase activity by fluorimetry, reported as relative fluorescent units (RFUs): column 1, PepSH; column 2, PepSNO; column 3, PepSH and etoposide; column 4, PepSNO and etoposide. Error is standard deviation (SD) ($n = 3$). **c)** Sample preparation was the same as that for panel b. The activity of each sample was baseline subtracted from a sample that did not receive etoposide. The relative activities shown are normalized to the peptide thiol activity, which were equivalent within error (SD, $n = 3$): column 1, 2.5 mM GSH; column 2, 0.5 mM GSNO; column 3, 2.5 mM GSNO; column 4, 2.5 mM PepSH; column 5, 0.5 mM PepSNO; column 6, 2.5 mM PepSNO. **d–f)** Apoptosis was also quantified by the Vybrant flow cytometry assay. Apoptotic cells appear in the lower right quadrant: **d)** 0.5 mM PepSNO; **e)** 0.5 mM PepSH and etoposide; **f)** 0.5 mM PepSNO and etoposide.

was used in subsequent cellular experiments. Anticipating difficulties with transport into cells, we converted the four carboxylate moieties of the peptide to methyl esters (Figure 2, panel a). The efficacy of this esterified compound, hereafter “PepSNO”, was examined first by adding the reagent to intact Jurkat cells (human T-cell lymphoma) that were treated with either etoposide or a buffer control for 5 h before addition of PepSNO and the thiol control, PepSH. After an additional 5 h treatment, cells were lysed and tested for caspase-3 activity (Figure 2, panel b). The caspase-3/7 (AcDEVD-ase) activity remained virtually unchanged in the non-etoposide-treated cells, regardless of treatment with PepSH or PepSNO. In the cells that were treated with etoposide, however, treatment with PepSNO reduced the AcDEVD-ase activity to barely detectable levels, while the PepSH-treated cells exhibited a high level of activity. The effectiveness of PepSNO was then compared with GSNO at two different concentrations (Figure 2, panel c). PepSNO (0.5 mM) was capable of reducing AcDEVD-ase activity by almost 90%, while an equal concentration of GSNO reduced the activity by 35%. Each reagent at 2.5 mM completely inhibited AcDEVD-ase activity. The two most probable explanations for the increased potency of the PepSNO reagent are (i) a greater rate of association with caspase-3 and transnitrosation with Cys163 and (ii) increased cellular permeability. To address the contribution of permeability, the two carboxylates in GSH and

GSNO were converted to their methyl esters (GSHe and GSNOe, respectively, Supplementary Methods) and examined under identical conditions. Treatment of intact Jurkat cells with GSNOe (0.5 mM) inhibited AcDEVD-ase activity (65%) to a greater extent than treatment with GSNO but still not as efficiently as treatment with PepSNO, $P < 0.005$ (Student's *t*-test, Supplementary Figure 4, panel a). AcDEVD-ase activity is a relatively early stage marker for apoptosis. A later marker in Jurkat cells is the translocation of phosphatidylserine (PS) to the cell surface (25). Using the Vybrant apoptosis assay kit (Invitrogen), we quantified the extent of apoptosis (annexin V, binds specifically to PS) and necrosis (propidium iodide) in cells treated with PepSH and PepSNO (Figure 2, panels d–f). Without etoposide treatment, the percentage of apoptotic Jurkat cells was 1% (Figure 2, panel d). Cells that were stimulated to undergo apoptosis with etoposide and were treated with PepSH (0.5 mM) contained a high percentage of apoptotic cells (Figure 2, panel e), but the same treatment with PepSNO reduced the extent of apoptosis by almost two-thirds (Figure 2, panel f). The same analysis was carried out for GSH, GSNO, GSHe, and GSNOe (Supplementary Table 1). Treatment of cells with GSNO and GSNOe led to an increase in the percentage of apoptotic cells in the absence of etoposide, while treatment of cells with GSH and GSHe inhibited etoposide-induced apoptosis, which was in accord with earlier observations (26).

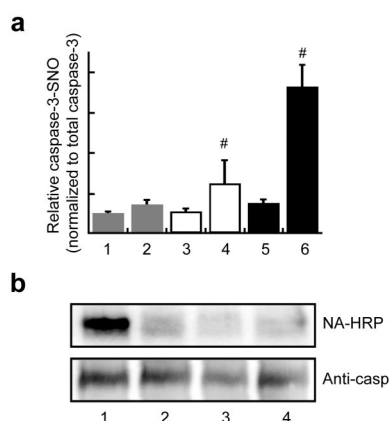


Figure 3. PepSNO transnitrosates the caspase-3 active site cysteine in apoptotic Jurkat cells. a) Relative S-nitrosation of caspase-3 in etoposide treated Jurkat cells. After a 5 h treatment with etoposide, the indicated peptide reagent (0.5 mM) was added for an additional 5 h. Anti-caspase-3 was used to purify caspase-3 by immunoprecipitation, which was then subjected to the biotin switch. Blots were probed with NA-HRP and anti-caspase-3 to ascertain the extent of caspase-3 S-nitrosation and to normalize for total caspase-3: column 1, GSH; column 2, GSNO; column 3, GSHe; column 4, GSNOe; column 5, PepSH; column 6, PepSNO. *P*-values were calculated by the Student's *t*-test: (gray) *P* = 0.02; (white) *P* = 0.10; (black) *P* < 0.001; (#) *P* = 0.006. b) A cell-permeable, caspase-3 reversible inhibitor (CPCI) competes with PepSNO for the active site of caspase-3. Cells were treated with etoposide and PepSH or PepSNO as usual, but 30 min prior to addition of PepSH or PepSNO, CPCI or DMSO was added to the cells: lane 1, PepSNO without CPCI; lane 2, PepSNO with CPCI; lane 3, PepSH without CPCI; lane 4, PepSH with CPCI.

These results underscore the necessity for more selective reagents to study the effects of S-nitrosation in apoptosis and other apoptotic processes (27). While we cannot rule out the possibility that PepSNO is participating in side reactions at this concentration and time course, the data confirm that this reagent is superior to GSNO and GSNOe in the inhibition of caspase-3-dependent apoptosis *via* S-nitrosation of the catalytic cysteine.

A further demonstration that PepSNO is a more selective reagent than GSNO and GSNOe is observed when the biotin switch method is performed on the entire cellular lysate using 0.5 mM of each peptide reagent. This method detects only abundant S-nitrosated proteins; therefore, caspase-3 is not detected (28). After treatment with GSNOe, the lysate contained numerous bands reactive toward NA-HRP, while the bands from GSH-, GSNO-, and GSHe-treated lysates were significantly more faint (Supplementary Figure 4, panel b). An identical assay using PepSH and PepSNO did not have any visible bands (data not shown), indicating that even though PepSNO was used at a higher concentration in the cellular experiments than in the *in vitro* experiments (0.5 mM vs 20 μ M), specificity for the caspase active site is retained within our detection. When one considers that PepSNO must penetrate the cell membrane and be processed by intracellular esterases before binding to the caspase-3 active site, it is understandable that a higher concentration is required in cells to attain inhibition of enzyme activity.

Despite the nonselective nature of GSNO and GSNOe, we were able to detect caspase-3-SNO in apoptotic Jurkat cells treated with these reagents using anti-

caspase-3 immunoprecipitation and the biotin switch method. The relative amount of caspase-3-SNO, which was normalized to total caspase-3, was determined for cells treated with 0.5 mM of the following: GSH, GSNO, GSHe, GSNOe, PepSH, and PepSNO (Figure 3, panel a). Cells treated with the thiol version of each reagent contained a basal amount of caspase-3-SNO, while the nitrosothiol derivative increased the percentage of caspase-3-SNO. The trend observed, PepSNO > GSNOe > GSNO, complied with the prediction based on the AcDEVD-ase activity of similarly treated cells (Figure 2, panel c and Supplementary Figure 4, panel a). Selectivity for the catalytic cysteine of caspase-3 in a cellular context was demonstrated for PepSNO by pretreating Jurkat cells with a cell-permeable, C-terminal aldehyde (reversible), caspase-3-selective active site inhibitor (Figure 3, panel b). The level of caspase-3 S-nitrosation was highest in cells treated with PepSNO lacking the inhibitor, while cells treated with PepSNO and the inhibitor had levels of caspase-3-SNO comparable to controls (PepSH with or without inhibitor). A low but detectable amount of procaspase-3 was found to be S-nitrosated. However, S-nitrosation on the processed caspase was much greater.

In summary, we have developed a cell-permeable peptide nitrosothiol (PepSNO) that inhibits etoposide-stimulated apoptosis in Jurkat cells *via* S-nitrosation of the catalytic cysteine of caspase-3. We characterized this process using a fluorimetric caspase-3/7 activity assay, flow cytometry, and the biotin switch method on purified enzymes, Jurkat cell lysate, and immunoprecipitated caspase-3. Specificity for the active site cysteine of caspase-3 was shown by active site competitive inhi-

bition experiments and MS. In each of these studies, the efficacy of PepSNO was superior to the commonly used biological transnitrosating agent GSNO and a more cell-permeable version, GSNOe. Derivatives of PepSNO may be tailored in future studies to be selective for individual caspase isoforms. If the optimal recognition sequences

are used as a blueprint (12), reagents could be designed with the expectation to be as caspase-selective as the related proteolytic substrate or aldehyde inhibitor (24). Reagents such as these will help elucidate the mechanism of NO- and nitrosothiol-dependent regulation of the caspase cascade.

METHODS

General. Purification of recombinant human caspase-3 was carried out as previously described (5). Jurkat cells were maintained in a 5% CO₂, water-saturated atmosphere at 37 °C in RPMI-1640 with glutamine (Gibco) supplemented with fetal bovine serum (FBS, 10%), penicillin (100 U mL⁻¹), and streptomycin (100 µg mL⁻¹).

Rate of *in Vitro* Caspase Deactivation. Recombinant caspase-3 and caspase-7 (0.1–5 µM) were reacted with varying concentrations of peptide nitrosating reagents (1–300 µM, dark, 25 °C, pH 7.4) under nonreducing conditions. At timed intervals, caspase activity was measured by diluting the above solution to a final caspase concentration (20–30 nM) in assay buffer supplemented with the profluorescent aminomethylcoumarin substrate AcDEVD-amc (75 µM, Calbiochem). Rate of deactivation was quantified by plotting the remaining AcDEVD-ase activity versus time of reaction. Rate constants were calculated using the bimolecular rate equation, $V = k[\text{caspase}][\text{inhibitor}]$. Due to the rapid deactivation kinetics exhibited by the irreversible chloromethylketone inhibitor, AcDEVD-cmk (Calbiochem), a lower concentration of enzyme (10–50 nM) and inhibitor (5–50 nM) were used.

Biotin Switch Method. Recombinant caspase-3 (8 µM) was allowed to react under reduced light with various peptide nitrosating agents (20 µM) for 20 min at RT and then subjected to the biotin switch method (5, 29). To demonstrate specificity for the active site cysteine of caspase-3 (Cys163), a separate experiment included a reversible peptide-aldehyde inhibitor, AcDEVD-cho (100 µM, Calbiochem). The biotin switch method was also used to show S-nitrosation of caspase-3 in human T-cells. Whole-cell lysates and caspase-3 immunoprecipitates (see below), using GSNO, GSNOe, and PepSNO as transnitrosating agents, were analyzed as previously described (30–32). The selectivity of PepSNO for the active site cysteine of caspase-3 was demonstrated in Jurkat cells using cells pretreated with the cell-permeable, reversible caspase-3 inhibitor CPCI, AcaAVALLPAVLLALLAP-DEVD-cho (50 µM, Alexis Biochemicals), 30 min prior to the addition of PepSH/SNO. After the NA-HRP Westerns were imaged, blots were stripped (Restore, Pierce), blocked, and probed with anti-caspase-3 (Cayman Chemical) to ensure equal loading.

Induction of Apoptosis and Peptide Treatment. Jurkat cells were harvested and resuspended in media containing reduced serum (1% FBS) and etoposide (2.5 µg mL⁻¹, Sigma). After 5 h at 37 °C, PepSNO, PepSH, GSNO, GSH, GSNOe, and GSHe (0.5 and 2.5 mM) were added to the cultures (10 mL) under reduced light conditions. Cells were harvested by centrifugation after a further 5 h incubation and washed twice (PBS) prior to lysis.

Flow Cytometry. Flow cytometry was performed with a FACS-Calibur flow cytometer (BD Biosciences) using CellQuest software. Apoptosis and necrosis were quantified using 50,000 intact cells per sample that were labeled with propidium iodide and Alexa Fluor 488 conjugated annexin V as suggested by the manufacturer's instructions (Vybrant Kit no. 2, Invitrogen).

AcDEVD-ase Activity of T-Cells. Jurkat cells were lysed under reduced light in RIPA/HEN buffer (100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 150 mM NaCl, 1 mM EDTA, 1 mM NaF, 100 µM neocuproine, 5% glycerol, 1% Triton X-100, 1% sodium deoxycholate, 0.05% SDS, pH 7.4), containing Complete (Roche) protease inhibitor cocktail, by mild vortexing at 4 °C. The insoluble fraction was removed by centrifugation (14,000 rpm, 15 min, 4 °C). Total soluble protein was quantified by the Bradford assay and normalized to 1.0–2.0 mg mL⁻¹ by dilution into RIPA/HEN, and the solution was supplemented with AcDEVD-amc (75 µM). AcDEVD-ase (caspase-3/7) activity was measured on a Molecular Devices Spectra Max Gemini XS ($\lambda_{\text{ex}} = 380 \text{ nm}$, $\lambda_{\text{em}} = 460 \text{ nm}$) using 96-well black fluorimetry plates (Corning) in triplicate.

Immunoprecipitation. Under reduced light, Jurkat cells were lysed, and the insoluble fraction was removed as above. The supernatants were then precleared by addition protein G–Sepharose 4B (Invitrogen, 30 µL, equilibrated in RIPA/HEN) for 30 min at 4 °C with rocking. The beads were pelleted by centrifugation, and the supernatant was transferred to a clean tube containing the anti-caspase-3 polyclonal antibody (Cayman Chemical, 30 µL). The antibody was allowed to bind for 90 min at 4 °C with rocking, after which the protein G beads were again added (40 µL, equilibrated in RIPA/HEN) for an additional 90 min binding at the same temperature (reduced light). The immunoprecipitate was pelleted by centrifugation (2000 rpm, 2 min), and the supernatant was decanted. The beads were washed three times with RIPA/HEN (1 mL) before elution of the immunocomplex from the beads using the biotin switch blocking buffer (contains 2.5% w/w SDS and 30 mM *N*-ethylmaleimide). The elution was carried out under reduced light at 55 °C for 25 min with frequent vortexing. After removal of the beads by centrifugation, the supernatant was subjected to the biotin switch method as described above.

Acknowledgments: We thank Phung Gip and Carolyn Bertozzi (UC–Berkeley) for the use of equipment, Debajyoti Datta and James Wells (UC–San Francisco) for caspase-7, Nathaniel Martin, Joshua Woodward, and Jacquin Niles (UC–Berkeley) for helpful discussions, and finally, members of the Marletta laboratory for critical review of the manuscript. This work was supported in part by grants from the National Institutes of Health and DeBenedictis Fund of UC–Berkeley to M.A.M. D.A.M. is supported by an American Heart Association predoctoral fellowship.

Supporting Information Available: This material is free of charge via the Internet.

Competing interests statement: The authors declare that they have no competing financial interests.

REFERENCES

1. Hess, D. T., Matsumoto, A., Kim, S. O., Marshall, H. E., and Stamler, J. S. (2005) Protein S-nitrosylation: purview and parameters, *Nat. Rev. Mol. Cell. Biol.* 6, 150–166.

- Mannick, J. B., Asano, K., Izumi, K., Kieff, E., and Stamler, J. S. (1994) Nitric oxide produced by human B lymphocytes inhibits apoptosis and Epstein-Barr virus reactivation, *Cell* **79**, 1137–1146.
- Mannick, J. B., Hausladen, A., Liu, L., Hess, D. T., Zeng, M., Miao, Q. X., Kane, L. S., Gow, A. J., and Stamler, J. S. (1999) Fas-induced caspase denitrosylation, *Science* **284**, 651–654.
- Mannick, J. B., Miao, X. Q., and Stamler, J. S. (1997) Nitric oxide inhibits Fas-induced apoptosis, *J. Biol. Chem.* **272**, 24125–24128.
- Mitchell, D. A., and Marletta, M. A. (2005) Thioredoxin catalyzes the S-nitrosation of the caspase-3 active site cysteine, *Nat. Chem. Biol.* **1**, 154–158.
- Mohr, S., Zech, B., Lapetina, E. G., and Brune, B. (1997) Inhibition of caspase-3 by S-nitrosation and oxidation caused by nitric oxide, *Biochem. Biophys. Res. Commun.* **238**, 387–391.
- Zech, B., Wilm, M., van Eldik, R., and Brune, B. (1999) Mass spectrometric analysis of nitric oxide-modified caspase-3, *J. Biol. Chem.* **274**, 20931–20936.
- Kim, J. E., and Tannenbaum, S. R. (2004) S-Nitrosation regulates the activation of endogenous procaspase-9 in HT-29 human colon carcinoma cells, *J. Biol. Chem.* **279**, 9758–9764.
- McCleverty, J. A. (2004) Chemistry of nitric oxide relevant to biology, *Chem. Rev.* **104**, 403–418.
- Sun, J., Xin, C., Eu, J. P., Stamler, J. S., and Meissner, G. (2001) Cysteine-3635 is responsible for skeletal muscle ryanodine receptor modulation by NO, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11158–11162.
- Xu, L., Eu, J. P., Meissner, G., and Stamler, J. S. (1998) Activation of the cardiac calcium release channel (ryanodine receptor) by poly-S-nitrosylation, *Science* **279**, 234–237.
- Thornberry, N. A., Rano, T. A., Peterson, E. P., Rasper, D. M., Timkey, T., Garcia-Calvo, M., Houtzager, V. M., Nordstrom, P. A., Roy, S., Vaillancourt, J. P., Chapman, K. T., and Nicholson, D. W. (1997) A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis, *J. Biol. Chem.* **272**, 17907–17911.
- Liu, L., Hausladen, A., Zeng, M., Que, L., Heitman, J., and Stamler, J. S. (2001) A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans, *Nature* **410**, 490–494.
- Sanghani, P. C., Bosron, W. F., and Hurley, T. D. (2002) Human glutathione-dependent formaldehyde dehydrogenase. Structural changes associated with ternary complex formation, *Biochemistry* **41**, 15189–15194.
- Sanghani, P. C., Robinson, H., Bosron, W. F., and Hurley, T. D. (2002) Human glutathione-dependent formaldehyde dehydrogenase. Structures of apo, binary, and inhibitory ternary complexes, *Biochemistry* **41**, 10778–10786.
- Sanghani, P. C., Stone, C. L., Ray, B. D., Pindel, E. V., Hurley, T. D., and Bosron, W. F. (2000) Kinetic mechanism of human glutathione-dependent formaldehyde dehydrogenase, *Biochemistry* **39**, 10720–10729.
- Xian, M., Chen, X. C., Liu, Z. Y., Wang, K., and Wang, P. G. (2000) Inhibition of papain by S-nitrosothiols—formation of mixed disulfides, *J. Biol. Chem.* **275**, 20467–20473.
- Xian, M., Wang, Q. M., Chen, X. C., Wang, K., and Wang, P. G. (2000) S-nitrosothiols as novel, reversible inhibitors of human rhinovirus 3C protease, *Bioorg. Med. Chem. Lett.* **10**, 2097–2100.
- Rotonda, J., Nicholson, D. W., Fazil, K. M., Gallant, M., Gareau, Y., Labelle, M., Peterson, E. P., Rasper, D. M., Ruel, R., Vaillancourt, J. P., Thornberry, N. A., and Becker, J. W. (1996) The three-dimensional structure of apopain/CPP32, a key mediator of apoptosis, *Nat. Struct. Biol.* **3**, 619–625.
- Ganesan, R., Mittl, P. R., Jelakovic, S., and Grutter, M. G. (2006) Extended substrate recognition in caspase-3 revealed by high resolution X-ray structure analysis, *J. Mol. Biol.* **359**, 1378–1388.
- Hogg, N., Singh, R. J., and Kalyanaraman, B. (1996) The role of glutathione in the transport and catabolism of nitric oxide, *FEBS Lett.* **382**, 223–228.
- Wang, K., Wen, Z., Zhang, W., Xian, M., Cheng, J. P., and Wang, P. G. (2001) Equilibrium and kinetics studies of transnitrosation between S-nitrosothiols and thiols, *Bioorg. Med. Chem. Lett.* **11**, 433–436.
- Wei, Y., Fox, T., Chambers, S. P., Sintchak, J., Coll, J. T., Golec, J. M., Swenson, L., Wilson, K. P., and Charifson, P. S. (2000) The structures of caspases-1, -3, -7 and -8 reveal the basis for substrate and inhibitor selectivity, *Chem. Biol.* **7**, 423–432.
- Talanian, R. V., Quinlan, C., Trautz, S., Hackett, M. C., Mankovich, J. A., Banach, D., Ghayur, T., Brady, K. D., and Wong, W. W. (1997) Substrate specificities of caspase family proteases, *J. Biol. Chem.* **272**, 9677–9682.
- Ferraro-Peyret, C., Quemeneur, L., Flacher, M., Revillard, J. P., and Genestier, L. (2002) Caspase-independent phosphatidylserine exposure during apoptosis of primary T lymphocytes, *J. Immunol.* **169**, 4805–4810.
- Franco, R., and Cidlowski, J. A. (2006) SLCO/OATP-like transport of glutathione in FasL-induced apoptosis: Glutathione efflux is coupled to an organic anion exchange and is necessary for the progression of the execution phase of apoptosis, *J. Biol. Chem.* **281**, 29542–29557.
- Berger, A. B., Witte, M. D., Denault, J. B., Sadaghiani, A. M., Sexton, K. M., Salvesen, G. S., and Bogoy, M. (2006) Identification of early intermediates of caspase activation using selective inhibitors and activity-based probes, *Mol. Cell* **23**, 509–521.
- Zhang, Y., Keszlér, A., Broniowska, K. A., and Hogg, N. (2005) Characterization and application of the biotin-switch assay for the identification of S-nitrosated proteins, *Free Radical Biol. Med.* **38**, 874–881.
- Jaffrey, S. R., and Snyder, S. H. (2001) The biotin switch method for the detection of S-nitrosylated proteins, *Sci. STKE* **2001**, pl1.
- Erwin, P. A., Lin, A. J., Golan, D. E., and Michel, T. (2005) Receptor-regulated dynamic S-nitrosylation of endothelial nitric-oxide synthase in vascular endothelial cells, *J. Biol. Chem.* **280**, 19888–19894.
- Erwin, P. A., Mitchell, D. A., Sartoretto, J., Marletta, M. A., and Michel, T. (2006) Subcellular targeting and differential S-nitrosylation of endothelial nitric-oxide synthase, *J. Biol. Chem.* **281**, 151–157.
- Mitchell, D. A., Erwin, P. A., Michel, T., and Marletta, M. A. (2005) S-Nitrosation and regulation of inducible nitric oxide synthase, *Biochemistry* **44**, 4636–4647.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

ACS Chemical Biology Abstracts in PubMed!

We're pleased to announce that *ACS Chemical Biology* (ACS CB) is approved for indexing in MEDLINE, the National Library of Medicine (NLM) database of indexed journal citations and abstracts. Your abstracts will soon appear in PubMed searches. Our content can also be found in the Chemical Abstracts Service, with the ISI Web of Science, and through Google searches. For those of you who published with us in 2006, rest assured that all of our content will be indexed in PubMed. For those future authors who were waiting for this milestone, it's time to submit your papers to our journal. Many of you have asked why it takes so long to see the abstracts from new journals in PubMed. Here's why.

A journal must publish four issues before it can apply for indexing in MEDLINE. The application process is straightforward—fill in a web form describing the content of the journal, some stats, and review and production criteria. Once the application is submitted, and if the subject matter is within MEDLINE's scope, the journal is scheduled for review. Here's what takes so long: the review process only happens a few times a year, so a journal launched at the start of the year will not usually be reviewed until October.

The NLM's Literature Selection Technical Review Committee (LSTRC) determines which journals will be indexed in MEDLINE. This 15-member advisory committee includes physicians, researchers, educators, editors, health science librarians, and even historians. The review process is similar to that for a National Institutes of Health grant: primary and secondary reviewers are chosen, and many criteria are used, including subject matter, quality of content, importance of the papers to those in the biomedical field, the rigor of the editorial practices, and the quality of the production. According to the NLM, only 25–30% of titles reviewed are selected for indexing.

Several weeks after the LSTRC meets, the publisher is notified whether the title is selected for indexing. If the news is positive, the publisher must send a test feed of the journal content to PubMed to establish the lines of communication. Once the technical bugs are worked out and PubMed performs its quality-control checks on the delivered material, the journal content (including back issues) is automatically fed to the NLM. Soon afterward, abstracts of the journal's content appear in PubMed searches. We are currently sending all ACS CB content to the NLM, and we expect our abstracts to appear in PubMed searches in 2007 (a few short weeks from now).

Thanks to all of our 2006 authors for publishing with us and for their patience as we completed the NLM indexing process. For those researchers who have not yet published with ACS CB, our manuscript submission system (<https://acs.manuscriptcentral.com/acs>) is standing by to receive your great papers!

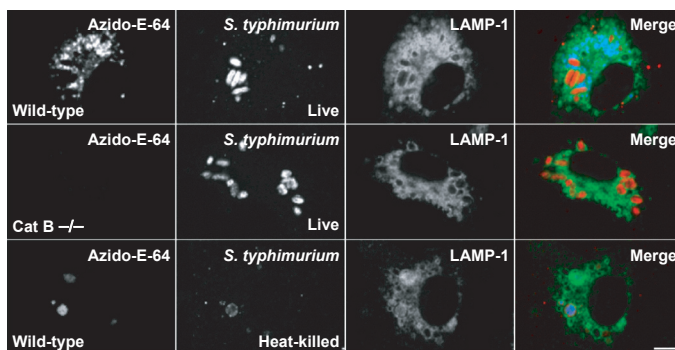


Evelyn Jabri
Executive Editor

Activity Probing the Proteome

Activity-based probes, a term coined for small-molecule enzyme inhibitors possessing chemical tags that enable visualization of affected proteins, are powerful tools for investigating enzyme activity. However, the presence of the tags can alter inhibitor activity or preclude entry into cells, frequently limiting the probes' ability to profile enzyme families *in vitro*. Now, Hang *et al.* (p 713) present a biorthogonal activity-based probe, termed azido-E-64, suitable for live cell labeling of cathepsin B, a cysteine protease involved in several important processes, including immune defense.

Azido-E-64 dodges the cell permeability issue thanks to an azide functionality that acts as a silent precursor to a visualization tag. After cell entry and enzyme inhibition, azido-E-64 can be selectively biotinylated at the azide



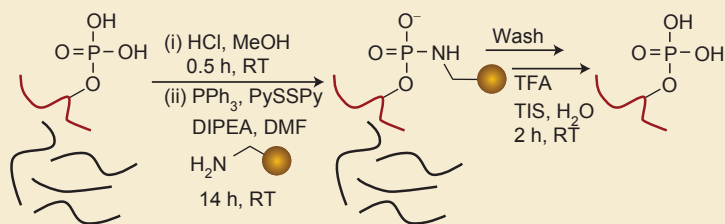
group by using Staudinger ligation chemistry. Using this approach along with others, the authors demonstrated that cathepsin B is selectively labeled by azido-E-64 in macrophages. Strikingly, they also showed that in *Salmonella*-infected macrophages, cathepsin

B is categorically excluded from *Salmonella*-containing vacuoles, which play a key role in *Salmonella*'s ability to elude destruction by host enzymes. These probes offer an exciting method for exploring enzyme activity *in vivo* and on a proteomic scale.

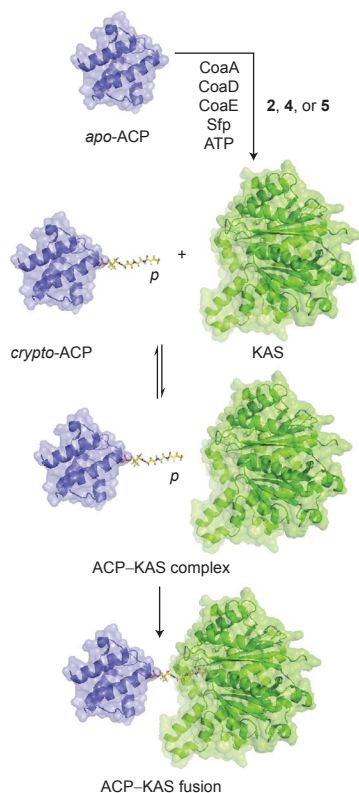
Finding Phosphorylation

The intricate signaling networks that govern cellular processes are orchestrated in part by protein phosphorylation events. However, detecting and characterizing phosphorylated proteins are complicated by the lack of efficient methods for enriching phosphorylated peptides and proteins from the concoction of components in cells and tissue samples. Now, Warthaka *et al.* (p 697) present oxidation–reduction condensation to selectively pull out phosphorylated peptides and proteins from complex mixtures.

This approach uses a phosphine and a disulfide to activate a phosphate for coupling with an alcohol or an amine. A mixture of synthetic peptides (a phosphorylated serine, threonine, or tyrosine and



several unphosphorylated peptides) was subjected to the reaction conditions in the presence of glycine-conjugated Wang resin. Mass spectrometry revealed that the method is selective for phosphate-containing peptides. The authors also showed that phosphorylated proteins could be captured and recovered from protein mixtures. This enrichment strategy offers greater selectivity and phosphopeptide compatibility than current techniques and will facilitate future investigations.



Synthetic Connections for Synthases

Like biological royalty, fatty acids, polyketides, and nonribosomal peptides are literally carried through their biosyntheses by “carrier” proteins (CPs). The biosynthetic precursors of these metabolites are covalently attached to the CPs, which present them to various enzymes (synthases) that sequentially build the final compounds in an assembly-line fashion. The structural complexity of the components, however, has hindered molecular characterization of this process. Now, Worthington *et al.* (p 687 and Point of View p 679) describe the synthesis and evaluation of small-molecule cross-linking agents selective for the *Escherichia coli* fatty acid ketosynthase elongation enzymes KASI and KASII.

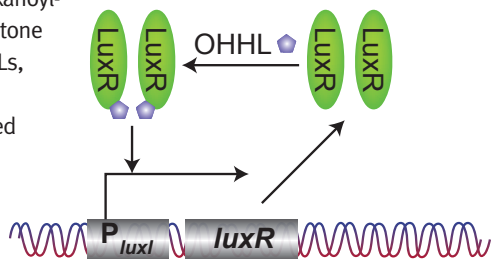
The agents were designed to undergo metabolic uptake that results in attachment to the acyl CP (ACP) and to react irreversibly with the ketosynthases. Indeed, incubation of the agents with the appropriate biosynthetic components caused covalent cross-linking of ACP and KASI or KASII. The agents were also selective for fatty acid and polyketide ACPs, because they did not react with CPs from nonribosomal peptide synthases. This innovative approach will facilitate various structural and proteomic investigations.

Positively Loopy!

Positive feedback loops (PFLs) are attractive tools for regulating gene expression and have many potential applications. Such artificial genetic circuits require a tightly regulated, inducible, and nontoxic system, and creative engineering strategies are often needed. Sayut *et al.* (p 692 and Point of View p 681) describe the generation of artificial PFLs through the use of components of a bacterial quorum-sensing system.

In the system, the transcriptional activator LuxR activates the gene expression promoter P_{luxI} in the presence of high concentrations of the small-molecule 3-oxo-hexanoyl-homoserine lactone (OHHL). Two PFLs, PFL1 and PFL2, were constructed such that LuxR expression is under the control of the

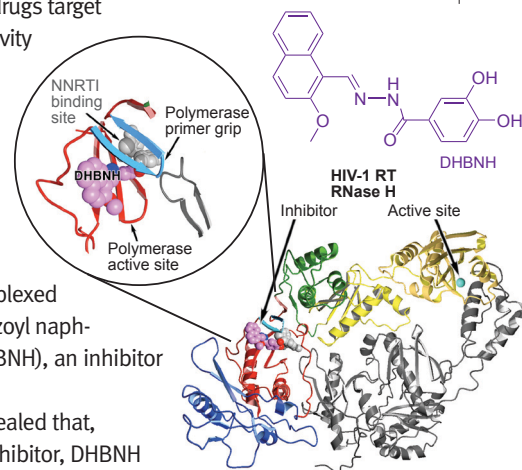
P_{luxI} promoter. Thus, addition of OHHL activates P_{luxI} and expression of *luxR*, triggering the loop. Indeed, PFL1 and PFL2 exhibited improved responses to OHHL over a control system and were completely activated at 10 nM OHHL, corresponding to six molecules per cell. To increase the sensitivity of the PFLs, the authors used error-prone PCR. With this directed evolution strategy, four LuxR mutants were discovered that exhibited an enhanced response to OHHL over the wild-type protein, likely from altered OHHL binding or LuxR function.



Inhibition from Afar

HIV is notoriously resistant to drugs, and novel compounds are desperately needed to inhibit it. HIV-1 reverse transcriptase (RT) is an essential enzyme for viral replication, possessing both a DNA polymerase and an RNase H (RNH) activity. All approved anti-HIV drugs target the polymerase activity of RT, so targeting RNH activity is an intriguing strategy for developing new inhibitors. Himmel *et al.* (p 702) now report the crystal structure of RT complexed with dihydroxy benzoyl naphthyl hydrazone (DHBNH), an inhibitor of RT RNH activity.

The structure revealed that, although an RNH inhibitor, DHBNH actually binds RT >50 Å from the RNH active site, ironically close to the polymerase active site. The authors hypothesize that binding the inhibitor perturbs the trajectory of the RNA–DNA duplex RNH substrate, preventing RNA cleavage. Additional analysis enabled the design of DHBNH analogues that also inhibited the polymerase activity, which could yield highly specific RT inhibitors. Because DHBNH interacts primarily with main-chain and C β atoms, mutant HIV strains will be less likely to evade inhibition.



Introducing our AUTHORS



Howard C. Hang

Current position: The Rockefeller University, Laboratory of Chemical Biology and Microbial Pathogenesis, Assistant Professor/Head of Laboratory

Education: University of California, Santa Cruz, B.S. in chemistry, 1998; University of California, Berkeley, Ph.D. in chemistry with Prof. Carolyn R. Bertozzi, 2003

Postdoctoral work: Whitehead Institute for Biomedical Research, Damon Runyon Cancer Research Foundation Postdoctoral Fellow with Prof. Hidde Ploegh, 2004–2006

Nonscientific interests: Traveling, cooking, watching movies, listening to music

The development of mechanism-based probes for various protein families is providing new insight into biological function. Nonetheless, the application of these chemical probes to living cells has been challenging. To address this limitation, we generated a cell-permeable mechanism-based probe that covalently labels cysteine proteases in living cells. This probe carried an azido group that could be visualized or retrieved with bioorthogonal labeling methods such as the Staudinger ligation. Using this approach, we visualized active cysteine proteases during bacterial infection of host cells. We also demonstrated their absence from organelles where bacterial pathogens reside; this absence may contribute to survival and replication of pathogenic bacteria in host cells. The arsenal of virulence mechanisms used by microbes to infect host cells is complex and not well understood. In this regard, I think that chemical tools can complement genetics and biochemistry to dissect the interactions between microbes and host cells. (Read Hang's article on p 713.)

Reversible protein phosphorylation regulates diverse cellular processes, including signal transduction, transcription, and metabolic pathways. Given the significance of phosphorylation, identifying phosphoproteins is important in the understanding of biological functions. Because analysis of phosphopeptides by mass spectrometry is difficult, enrichment usually precedes analysis, but current phosphopeptide enrichment methods are challenging. We used an oxidation–reduction condensation to develop a direct phosphate modification of phosphopeptides and phosphoproteins and demonstrated it as a method for their enrichment. The oxidation–reduction condensation method displays general reactivity toward phosphopeptides, high chemical selectivity, and dependence on common commercially available reagents. We expect that it will be a useful tool for phosphopeptide and phosphoprotein analysis and for development of new phosphoproteomic tools. (Read Warthaka's article on p 697.)

Current position: Wayne State University, Department of Chemistry, Ph.D. candidate with Prof. Mary Kay H. Pflum

Education: University of Kelaniya, Sri Lanka, B.Sc. sp. in chemistry, 1996

Nonscientific interests: Listening to my 8-year-old son play piano



Mangalika Warthaka

My research interests include investigating the protein interactions that occur during fatty acid and polyketide biosynthesis. These pathways are important targets for the treatment of diseases, including cancer and heart disease, as well as for the development of antibiotics. These complex biosynthetic systems contain a variety of enzymes that all interact with a carrier protein (CP). During these interactions, the CP functions as a scaffold for the growing product as it is assembled. In my research, I have co-opted the CP's post-translational modification pathway to generate modified CPs useful for investigating the interactions between the CP and its partner proteins. These studies will further our understanding of these complex biosynthetic systems. (Read Worthington's article on p 687 and Point of View on p 679.)



Andrew Worthington

Current position: University of California, San Diego, Department of Chemistry, Ph.D. candidate with Prof. Michael D. Burkart

Education: Kenyon College, B.A. in chemistry, 2003

Nonscientific interests: Surfing, knitting, boxing



Daniel J. Sayut

Current position: University of Massachusetts Amherst, Department of Chemical Engineering, Ph.D. candidate with Prof. Lianhong Sun

Education: Syracuse University, B.S. in chemical engineering, 2004

Nonscientific interests: Juggling, mountain biking, snowboarding, literature, hiking

Coming from a background in chemical engineering, I am amazed by the complexity and robustness of cellular systems. Currently, my favorite cellular systems are the quorum-sensing mechanisms used by bacteria for communication. The positive feedback loops constructed in this paper use components isolated from a quorum-sensing system and represent a novel application of this inherent cellular mechanism. The ease with which the behavior of the constructed feedback loops were modified by directed evolution demonstrates the power and elegance of this technique for engineering protein function. (Read Sayut's article on p 692 and Point of View on p 681.)

I use X-ray crystallography to determine biological macromolecular structures. I am currently focusing on HIV-1 reverse transcriptase (RT) in complex with RT RNase H (RNH) inhibitors and non-nucleoside RT inhibitors. These inhibitors, such as the RNH inhibitor dihydroxy benzoyl naphthyl hydrazone, may serve as lead molecules for the structure-based design of therapeutic agents against drug-resistant strains of HIV. I also collaborate with colleagues in computational chemistry to develop improved methods for X-ray structure determination and refinement. I am developing new phase-modification methods for the visualization of electron density maps used during structural refinement. My interests encompass employing structural biology to find new ways to combat disease, as well as understanding the organized mechanical motions that occur during the rearrangement of protein subdomains, such as during the transduction of chemical energy into mechanical work in molecular motors. (Read Himmel's article on p 702.)

Current position: Rutgers University, Department of Chemistry and Chemical Biology, Center for Advanced Biotechnology and Medicine, postdoctoral fellow with Prof. Eddy Arnold

Education: Rutgers University, Cook College, B.A. in biochemistry, 1983; University of Pittsburgh, M.S. in biochemistry and biophysics with Prof. Theresa Chay, 1986; Brandeis University, Ph.D. in biophysics and structural biology with Prof. Carolyn Cohen, 2000

Nonscientific interests: Hiking, camping, cycling, skiing, composing music



Daniel M. Himmel



Stefan G. Sarafianos

Current position: University of Missouri–Columbia, School of Medicine, Department of Molecular Microbiology and Immunology, Christopher Bond Life Sciences Center Assistant Professor

Education: University of Patras, Greece, B.S. in chemistry, 1986; Georgetown University, Ph.D. in chemistry and biochemistry with Prof. Soma Kumar, 1992

Postdoctoral work: University of Medicine and Dentistry of New Jersey, postdoctoral fellow with Dr. Mukund Modak, 1993–1995; Rutgers University, Center for Advanced Biotechnology and Medicine and Department of Chemistry and Chemical Biology, postdoctoral fellow and research faculty with Prof. Eddy Arnold, 2000–2006

Nonscientific interests: History, archaeology, music, spending time with my four children

My research has focused on the structural and biochemical mechanisms of the polymerase and RNase H (RNH) activities of HIV-1 reverse transcriptase (RT). For more than 10 years, I have been taking "snapshots" of RT in complexes with various substrates and inhibitors. This has been part of our effort to understand the molecular details of viral replication mechanisms, viral inhibition, and viral drug resistance. The present work reveals the first glimpses of molecular interactions of RT with an RNH inhibitor. The inhibitor binds at an unexpected site, closer to the polymerase rather than the RNH active site, and may offer opportunities for novel antiviral strategies. My other research interests include the replication mechanisms in positive-stranded RNA viruses, including the SARS coronavirus. (Read Sarafianos's article on p 702.)

Spotlight

Carbohydrate Channeling

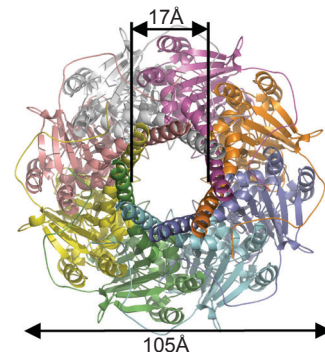
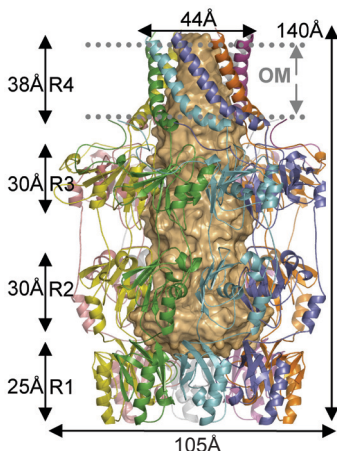
Many bacteria produce extracellular polysaccharides (EPSs) that play pivotal roles in their structural integrity and pathogenicity. Certain EPSs form a discrete structural layer called the capsule, and group 1 capsule assembly is characterized by the synthesis of individual lipid-linked polysaccharide repeats that are exported to the periplasm for assembly of the full-length polymer. Once assembled, the polymer is transported across the outer membrane by the lipoprotein Wza, a member of the outer membrane auxiliary protein family. Dong *et al.* (*Nature* 2006, 444, 226–229) now present the crystal structure of Wza, providing insight into the regulation and mechanism by which capsular polysaccharide

is exported through the bacterial membrane.

The structure revealed that Wza is an octamer containing a large internal cavity open to the extracellular environment but closed to the periplasm. Each mono-

mer of Wza has four domains, with domains 2 and 3 being duplicates. The eight copies of domain 1 form a ring structure with a concave surface at the base of the protein. Domain 4 is an amphipathic helix, and the eight monomers create an

α -helical barrel at the top of the structure. The central cavity and the exterior surface of the protein are polar, with the exception of the markedly hydrophobic α -helical barrel that is predicted to be the transmembrane region. Notably, all other known integral outer membrane proteins possess transmembrane β -barrels; Wza is the first example of a transmembrane α -helical barrel. On the basis of this structure, the authors propose that the carbohydrate moves from the periplasm to the central cavity of Wza. The polar environment enables water molecules to effectively lubricate the cavity for translocation of the polysaccharides to the membrane, where they exit through the α -helical barrel. They also hypothesize that the opening of the cavity is regulated *via* a conformational change at the concave base of the protein, possibly through interaction with the inner membrane tyrosine autokinase Wzc. **Eva Gordon**



Both images reprinted by permission from Macmillan Publishers Ltd: *Nature*, Dong, C., *et al.*, 444, 226–229, copyright 2006.

Pluripotency Aplenty

Embryonic stem (ES) cells are pluripotent, which means they have the extraordinary ability to differentiate into all cell types and thus harbor tremendous therapeutic promise. Under self-renewal conditions, however, ES cells can propagate in the pluripotent state indefinitely. Molecular tools that help define the mechanisms governing ES cell self-renewal will contribute to our

understanding of basic developmental biology and facilitate potential ES cell-based therapies. Chen *et al.* (*Proc. Natl. Acad. Sci. U.S.A.* 2006, 103, 17,266–17,271) now report the discovery of a synthetic small molecule, SC-1 (also named pluripotin), that on its own sustains the self-renewal of murine ES (mES) cells.

SC-1 was discovered from a high-throughput screen of 50,000 het-

erocytes against an mES cell line that, in the absence of self-renewal conditions such as feeder cells or exogenous factors, has limited self-propagation ability. Remarkably, SC-1-treated cells could self-renew for >10 passages, were phenotypically indistinguishable from ES cells grown under self-renewal conditions, and, upon removal of SC-1, could differentiate into various cell

(continued on page 673)

Pluripotency Aplenty,

continued from page 672

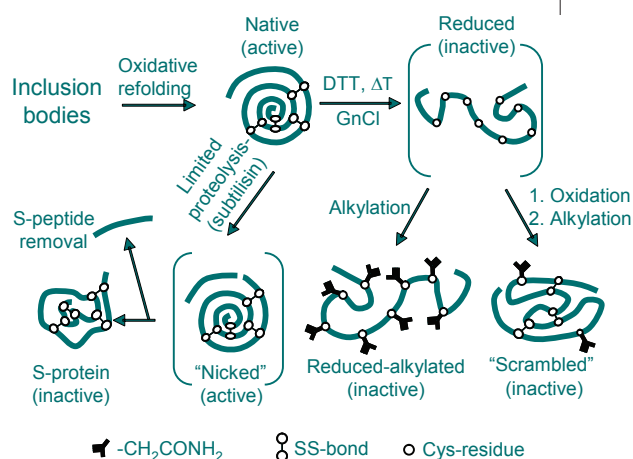
types in culture and contribute to the germ line in mice *in vivo*. After initial studies failed to link SC-1 to known self-renewal pathways, target identification experiments using an SC-1-linked affinity matrix led to the isolation of two proteins, extracellular signal-regulated protein kinase-1 (Erk1) and Ras GTPase-activating protein (RasGAP). In cells, SC-1 was shown to inhibit Erk1 phosphorylation and to activate the Ras pathway by inhibiting RasGAP. The authors subsequently demonstrated that overexpression of Erk1 or RasGAP promotes ES cell differentiation, either of which can be partially rescued by exposure to SC-1. Notably, the dual inhibition of Erk1 and RasGAP was necessary for maintaining the cells in an undifferentiated state; blocking either pathway alone by other methods was not sufficient to sustain self-renewal. The authors propose that SC-1 functions by simultaneously blocking Erk1 pathways, which are known to contribute to cell differentiation, and by inhibiting RasGAP pathways, which may activate self-renewal pathways through Ras signaling. **Eva Gordon**

Flexibly Folded

N-Linked glycosylation of proteins serves important structural and functional roles, including involvement in protein folding, stability, cell adhesion, and the immune response. In eukaryotes, it is known that N-linked glycosylation is coupled to protein translocation and that the enzyme responsible, oligosaccharyltransferase (OTase), recognizes its substrate in an unfolded state. However, the homologous glycosylation pathway in bacteria is less well characterized. Now, Kowarik *et al.* (*Science* 2006, 314, 1148–1150) report that bacterial N-glycosylation is not linked to the protein translocation machinery.

Using pulse chase experiments, the authors first observed that the bacterial protein AcrA was glycosylated by the bacterial oligosaccharyltransferase (PglB) post-translationally, a sign of a translocation-independent reaction. When they explored this further, they found that an AcrA fusion that was directed to the periplasm as a folded protein was also glycosylated, an indication that PglB can glycosylate folded, exported proteins. They next demonstrated that PglB is capable of glycosylating purified AcrA *in vitro*, in sharp contrast to eukaryotic OTases, which do not possess this ability. When a peptide

containing the glycosylation consensus sequence of AcrA was inserted into a loop in GFP and the GFP-derivative was exposed to PglB, the protein was glycosylated and retained full fluorescence, a further demonstration that PglB is capable of glycosylating folded proteins *in vitro*. Finally, four folding variants of bovine ribonuclease A (RNaseA), which included a folded, enzymatically active species, a partially folded species, and two random coils, were examined for their ability to act as PglB substrates. PglB glycosylated all four forms of RNaseA, but notably, the nonstructured species were better PglB substrates than the folded proteins. The authors propose that in bacteria, glycosylation sites likely adopt specific conformations in locally flexible areas of folded proteins, providing insight into the different mechanisms that have evolved for prokaryotic and eukaryotic glycosylation systems. **Eva Gordon**

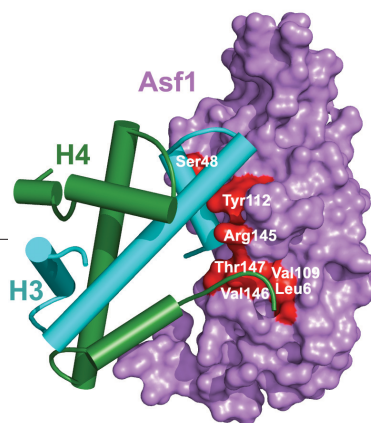


From Kowarik, M., *et al.*, *Science*, Nov 17, 2006, DOI: 10.1126/science.1134351. Reprinted with permission from AAAS.

Chaperoning Chromatin

Not unlike teenagers at the prom, the biomolecules in the nuclei of cells that make up chromatin need chaperones to ensure that they behave properly. Chromatin consists of DNA wound around proteins called histones; whereas the interaction between histones and DNA has been extensively characterized, the interactions of histones with each other and with chaperones in the absence of DNA are not well understood. Now, English *et al.* (*Cell* 2006, 127, 495–508) present the crystal structure of the histone-chaperone protein anti-silencing function 1 (Asf1) bound to the H3/H4 histone heterodimer, providing novel insights into the interactions between histones and their chaperones.

The Asf1–H3/H4 structure revealed that Asf1 interacts extensively with both H3 and H4, an unexpected finding because previous studies had suggested that Asf1 directly interacts only with H3. Notably, the last 10 residues of H4 add an additional β -strand to an antiparallel β -sheet within Asf1. Moreover, the carboxyl terminus of H4 undergoes a dramatic conformational change in order to form this β -sheet with Asf1. Interest-



Reprinted from *Cell*, 127, English, C. M., *et al.*, Structural basis for the histone chaperone activity of asf1, 495–508, Copyright 2006, with permission from Elsevier.

ingly, it is these same residues that form a β -sheet with histone 2A in the nucleosome, stabilizing nucleosome structure. To explore the biological relevance of the key Asf1–H3/H4 contacts revealed in the crystal structure, the authors mutated various residues within the binding interface between Asf1, H3, and H4. Yeast expressing these mutant proteins exhibited growth defects, reduced chromatin-disassembly activity, and diminished Asf1–H3/H4 interactions; this confirms the importance of the Asf1–H3/H4 interaction in chromatin regulation. On the basis of structural and biological data from these studies and the established role of the H4 carboxyl terminus in stabilizing histone structure, the authors propose that once this region of H4 is exposed, the interprotein β -sheet formed between Asf1 and H4 captures the H3/H4 complex and facilitates further chromatin disassembly, putting forth a potential mechanism by which histone chaperones mediate nucleosome assembly and disassembly. **Eva Gordon**

Channeling Autoimmunity

Certain autoimmune diseases such as multiple sclerosis (MS), rheumatoid arthritis (RA), or type-1 diabetes mellitus (T1DM) are caused in part by a specific type of autoreactive white blood cell called an effector memory T (T_{EM}) cell. TEM cells from patients with MS have distinct phenotypic characteristics, including elevated expression levels of Kv1.3 potassium channels. Beeton *et al.* (*Proc. Natl. Acad. Sci. U.S.A.* 2006, 103, 17,414–17,419) now demonstrate that T_{EM} cells from patients with RA or T1DM also exhibit high levels of Kv1.3 and that Kv1.3 blockers selectively target these T_{EM} cells and can diminish RA progression and T1DM incidence in rats.

Using patch-clamp technology, immunostaining, and flow cytometry, the authors first showed that disease-associated autoreactive T_{EM} cells from T1DM and RA patients express high levels of Kv1.3. Next, they evaluated immunological synapse formation, Ca^{2+} signaling, cytokine production, and [3H]thymidine incorporation and determined that selective Kv1.3 blockers preferentially block T_{EM} cell function without affecting the activity of other T cells, an important distinction to ensure that other aspects of immune function are not compromised. The therapeutic potential of Kv1.3 blockers was next assessed in RA and T1DM disease models in rats. Treatment with Kv1.3 blockers significantly alleviated RA symptoms and reduced the incidence of experimental autoimmune diabetes in the animals. Notably, toxicity studies using repeated doses of the blockers revealed that they did not appear to cause adverse effects. These exciting findings indicate that Kv1.3 channels may be a promising therapeutic target for any autoimmune disease in which levels of Kv1.3 on T_{EM} cells are elevated. Thus, discovery of new small-molecule Kv1.3 blockers could lead to effective drugs for a variety of autoimmune diseases. **Eva Gordon**

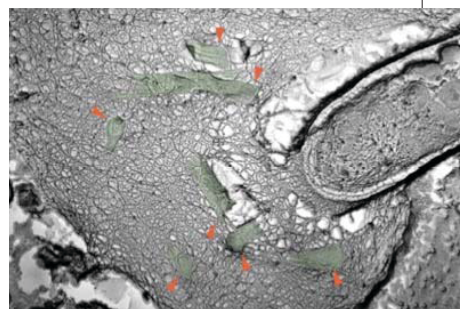
Shigella's Path of Destruction

Shigella, the pathogenic bacteria responsible for causing dysentery, motor through the cytoplasm of host cells by inducing actin polymerization and hitching a ride on the resulting propulsive force. This motility helps firmly establish infection and facilitates access to neighboring host cells, but the mechanisms that guide the movement are not clearly defined. Now, Yoshida *et al.* (*Science* 2006, 314, 985–989) report that the *Shigella* effector protein VirA degrades host cell microtubules (MTs) *via* a tubulin-specific cysteine-protease-like activity, effectively clearing a path for movement through the cytoplasm.

In a comparison of wild-type (WT) and *virA* mutant (*virA*⁻) bacterial movement through mammalian cells, immunofluorescence microscopy revealed that *virA*⁻ mutants did not move, an indication of the critical role that VirA plays in *Shigella* motility. The authors also demonstrated that MT structure acts as a barrier to movement; WT bacteria exhibited smoother movement with reduced variability in cells treated with the MT-destabilizing drug nocodazole compared with untreated cells. Further probing with immunofluorescence and freeze-fracture electron microscopy revealed that as the bacteria move, they create tunnel-like areas through MT networks in the precise area where the actin tail had formed. Moreover, the MT networks were frequently fragmented in the tunnel zone, a sign that the bacteria were break-

ing down MTs that were in their path. Suspectingly, VirA just happened to be secreted on the bacterial surface *via* the type III secretion system in the exact vicinity of the MT fragmentation. Perusal of the secondary structure of VirA revealed a strong resemblance to a family of cysteine proteases; thus, VirA was examined for cysteine-protease-like activity. Indeed, VirA was capable of degrading α -tubulin, and this activity was inhibited by cysteine protease inhibitors. Importantly, mice infected with various *Shigella* mutants with compromised VirA activity had increased survival rates compared with the WT bacteria. This study illuminates how VirA contributes to the remarkable strategy by which *Shigella* sabotages its intracellular obstacles.

Other cytoplasmic invading bacterial pathogens such as *Listeria monocytogenes*, *Mycobacterium marinum*, *Rickettsia prowazekii*, and *Burkholderia pseudomallei* also move through the host cell cytoplasm by inducing actin polymerization at one pole of the bacterium; thus, similar methods to deforest the MT jungle may be used by these pathogens as well. **Eva Gordon**



From Yoshida, S., *et al.*, *Science*, Nov 10, 2006, DOI: 10.1126/science.1133174. Reprinted with permission from AAAS.

Driving Cancer Cells to Suicide

Cancer cells have the notorious ability to sidestep the innate pathways that direct a cell to commit suicide, an important and natural part of a normal cell's life cycle. Moreover, many cancer cells develop resistance to the very drugs designed to kill them by inducing cell suicide. The Bcl-2 protein family consists of several proteins that play important roles in guiding cells toward survival, such as the prosurvival proteins Bcl-2, Bcl-x_L, or Mcl-1, or death, including the prodeath

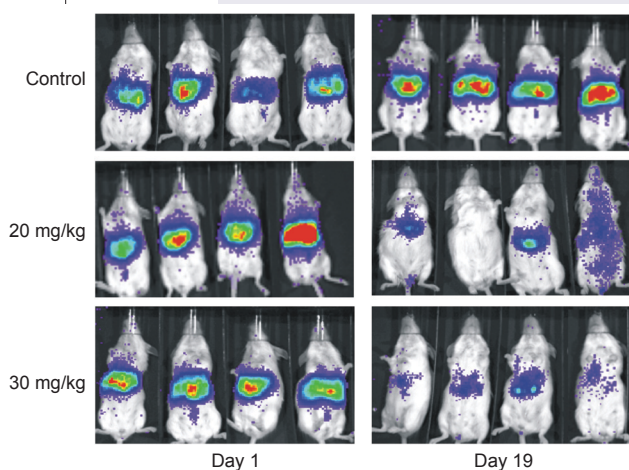
proteins Noxa, Bim, Bad, Bax, and Bak. Interaction of a conserved dimerization motif, the BH3 domain on certain prodeath members, with a hydrophobic groove on prosurvival members can

initiate cell suicide, otherwise known as apoptosis. This suggests that small molecules that mimic the BH3 domain could be effective anticancer agents. Two recent reports (Konopleva *et al.*, *Cancer Cell* 2006, 10, 375–388, and van Delft *et al.*, *Cancer Cell* 2006, 10, 389–399) dissect the molecular mechanism of ABT-737, a small-molecule BH3 mimetic, and provide enticing evidence for the consideration of ABT-737 as a therapeutic agent for a variety of cancers.

apoptosis if Bak and/or Bax are present. Investigation of several putative small-molecule BH3 mimics using Bax and Bak-deficient cells revealed that only ABT-737 actually induced apoptosis *via* the same mechanism that BH3-containing proteins do. This is an important consideration in the design of effective anticancer agents to ensure selectivity and to reduce the likelihood of adverse side effects. ABT-737 was shown to bind selectively to Bcl-2 and Bcl-x_L and to disrupt the association of Bcl-2 with Bax in the leukemia cell line HL-60. Notably, however, ABT-737 did not bind to the prosurvival protein Mcl-1.

The human leukemia cell line HL-60, which does not express high levels of Mcl-1, underwent apoptosis upon exposure to ABT-737, whereas other cell lines that express high levels of Mcl-1 were resistant to ABT-737-induced apoptosis. Both research groups hypothesized that treatment of resistant cell lines with ABT-737 in combination with agents that reduce Mcl-1 expression might lead to enhanced cytotoxicity. Several strategies known to diminish Mcl-1 levels, including treatment with short hairpin RNA against Mcl-1, exposure to extracellular signal-regulated protein kinase (Erk) phosphorylation inhibitors, overexpression of Noxa, or IL-3 deprivation, each resulted in greatly enhanced sensitivity of otherwise resistant cells to ABT-737. Additionally, treatment of resistant cell lines with other agents, such as the chemotherapeutic drugs doxorubicin, cytosine arabinoside, or etoposide; the cyclin-dependent kinase inhibitor R-roscovitine; or the protein synthesis inhibitor cycloheximide, also resulted in enhanced sensitivity to ABT-737, likely in part by down-regulating Mcl-1 expression.

(continued on page 677)



Reprinted from *Cancer Cell*, 10, Konopleva, M., *et al.*, Mechanisms of apoptosis sensitivity and resistance to the BH3 mimetic ABT-737 in acute myeloid leukemia, 375–388, Copyright 2006, with permission from Elsevier.

FRET No More Over O-GlcNAc

Post-translational modification of proteins with β -O-N-acetyl-D-glucosamine (O-GlcNAc) is a dynamic event that, like phosphorylation, contributes to cell signaling in nuclear and cytosolic proteins and has been linked to certain disease states. Unfortunately, little is known about the spatial and temporal regulation of O-GlcNAc during cell signaling because of a lack of methods to study this modification in live cells. Now, Carrillo *et al.* (*J. Am. Chem. Soc.* 2006, 128, 14,768–14,769) present a FRET-based sensor for detecting the dynamics of O-GlcNAc modifications in cells.

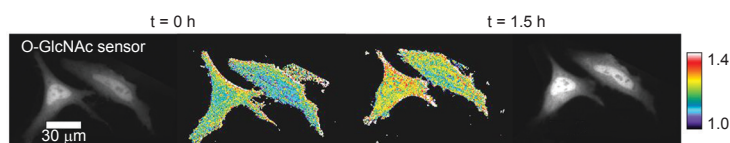
The sensor was designed so that the addition of an O-GlcNAc to an appropriate substrate results in a binding event, which in turn brings two fluorophores together to cause an increase in FRET. A fusion protein is composed of an enhanced cyan fluorescent protein, the bacterial lectin GafD (which binds terminal β -O-GlcNAc), the casein kinase II-derived peptide substrate for O-GlcNAc transferase, and a yellow fluorescent protein variant made up the components of the sensor. To establish that the sensor could detect the addition and the removal of O-GlcNAc from proteins, the authors used bacterially expressed sensors to conduct an *in vitro* assay. As expected, a significant increase in fluorescence emission at 528 nm was observed when the sensor was incubated with recombinant O-GlcNAc transferase and

Driving Cancer Cells to Suicide,

continued from page 676

Both studies also revealed the efficacy of ABT-737 as an anticancer agent in animal models. Konopleva *et al.* showed that ABT-737 significantly reduced leukemia burden and extended survival of mice in leukemia models, with no adverse effects on normal tissue. In addition, van Delft *et al.* demonstrated that ABT-737 prolonged survival of mice transplanted with either control or Bcl-2 expressing lymphoma cells. Notably, mice transplanted with Mcl-1 expressing lymphoma cells were resistant to ABT-737 treatment.

Exploration of ABT-737's mechanism of action enabled both research groups to determine why certain cancer cells become suicidal upon exposure to ABT-737 while others resist being pushed down that path. On the basis of these studies, ABT-737 has exciting prospects as a single agent in tumors with low Mcl-1 expression or in combination with other anticancer agents in high Mcl-1-expressing tumors. Moreover, these discoveries have revealed that Mcl-1 expression levels could be a valuable prognostic marker for response to ABT-737. **Eva Gordon**



Reprinted with permission from Carrillo, L. D., *et al.*, *J. Am. Chem. Soc.*, 128, 14,768–14,769. Copyright 2006 American Chemical Society.

the requisite sugar donor UDP-GlcNAc. Likewise, a significant loss of signal was observed when O-GlcNAcase, the enzyme that removes O-GlcNAc from proteins, was added to a sensor that contained O-GlcNAc. The sensor was next tested for its ability to detect O-GlcNAc dynamics in live cells. Indeed, treatment of a human cancer cell line transiently transfected with the sensor that had reagents that promoted a dynamic increase in O-GlcNAc resulted in a significant increase in FRET. This sensor represents the first method for exploring O-GlcNAc dynamics in a cellular setting, and its modular design will enable examination of the interplay between O-GlcNAc and phosphorylation networks. **Eva Gordon**

Trapping Transient Protein–Protein Interactions in Polyketide Biosynthesis

Nathan A. Schnarr and Chaitan Khosla*

Departments of Chemistry, Chemical Engineering, and Biochemistry, Stanford University, Stanford, California 94305

Despite the enormous impact of transient biomolecular interactions on chemistry and biology, reliable strategies for characterizing them remain largely elusive. Chemical cross-linking of proteins with variable binding affinities provides a means of assaying critical contact sites (1–6). For certain specialized cases, natural enzymatic activities may be exploited to serve as highly selective covalent stabilization. In this issue of *ACS Chemical Biology*, Burkart and co-workers describe an impressive application of this concept to fatty acid synthase components (7). This work not only will open doors to an increased understanding of thiotemplate assembly but will offer significant potential for improved biosynthetic engineering efforts.

Carrier proteins (CPs) serve as primary workhorses in fatty acid, polyketide, and nonribosomal peptide biosynthesis, supplying the appropriate extender unit for each (8–10). In fatty acid and polyketide synthases, CPs collaborate with specific ketosynthases (KSs) to elongate the growing fatty acid or polyketide backbone by two carbon atoms. Ultimately, KS–CP recognition dictates the flow of intermediates as well as turnover numbers in these systems. Despite their importance, these protein–protein interactions are relatively weak. For example, the dissociation constant for the interaction between the actinorhodin KS and appropriate CPs is estimated to be in the 1–10 μM range (11). Consequently, insight into the structural determinants of

KS–CP recognition has been limited. Recent high-resolution structures from fatty acid and polyketide synthases have renewed scientific interest in modular architecture (12–14). However, because the CP position is unknown, the mechanistic basis for several steps in the overall catalytic cycle remains mysterious. Utilizing the native reactivity of both CP and KS domains, Burkart's group has discovered a clever means of trapping KS–CP complexes for further analysis.

Initially, a chemically synthesized alkylating group is enzymatically coupled to a CP via the phosphopantetheine prosthetic arm. When introduced to the partner KS domain, reversible binding promotes irreversible alkylation of the active-site cysteine residue, creating the aforementioned chemical cross-link (Figure 1, panel a). The authors clearly demonstrate that selectivity arises from protein–protein interactions and not protein–substrate specificity by examining biochemically unrelated CPs harboring the same alkylating functionality. Importantly, common and scalable synthetic methods can be used to prepare these reactive end groups, an asset too commonly overlooked. Attempts to generate a workable model for KS–CP recognition through crystallization of novel cross-links are ongoing and may finally reveal definitive roles for CP recognition elements.

Clearly, much of our ability to effectively engineer biosynthetic assemblies relies on suitable reprogramming of individual enzymes. Because protein–protein interac-

ABSTRACT Transient biomolecular interactions are essential for biological processes, but strategies for studying them have remained elusive. A paper in this issue shows how natural enzymatic activities can be exploited to examine protein–protein interactions in fatty acid synthase.

*Corresponding author,
khosla@stanford.edu.

Published online December 15, 2006

10.1021/cb600451d CCC: \$33.50

© 2006 by American Chemical Society

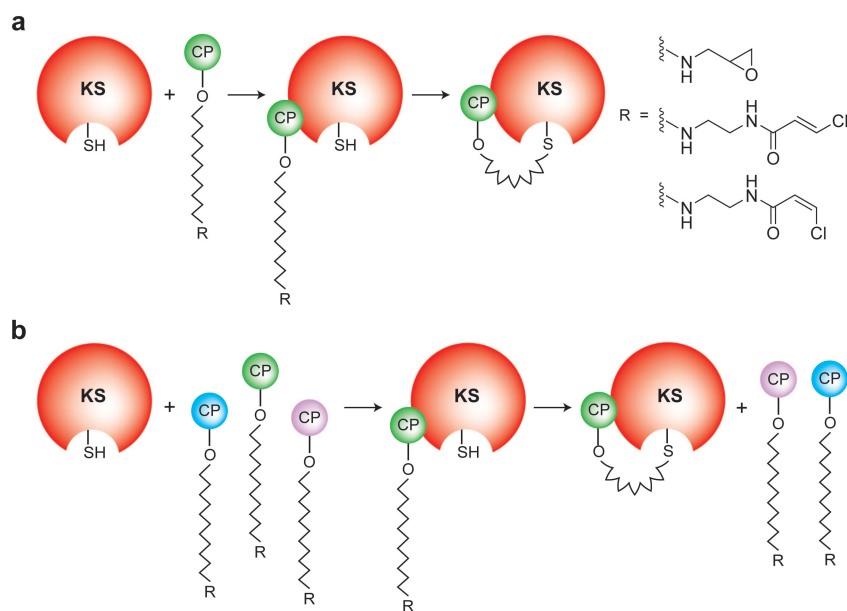


Figure 1. Cross-linking strategy to study fatty acid synthase. a) Schematic representation of KS–CP cross-linking via phosphopantetheine-tethered thiol traps. Transient binding of CP to KS leads to irreversible active-site alkylation. b) Theoretical application of cross-linking strategy to competitive CP binding. A panel of CPs treated with the KS of interest. A correlation between primary sequence and binding is readily extractable by comparison of bound and unbound material.

tions play a substantial role in all thiotemplate-based systems, significant attention has been focused on determining the factors governing these events. This is where a reliable cross-linking strategy may benefit researchers most. For a given KS domain, an array of similar CPs can be readily assayed for competitive binding advantage (Figure 1, panel b). The diminutive size of CPs should facilitate identification of bound material if a method of release from the partner KS is developed.

The work presented by Burkart and co-workers will provide important mechanistic insights into assembly-line biosynthesis. Whether toward improving or disrupting specific biomolecular interfaces, the impact of a viable cross-linking strategy for mapping recognition elements is far-reaching. This paper, together with recent structural information, may soon bring us a near-complete picture for both fatty acid and

polyketide biogenesis. In all, this is another fitting example of integrative chemistry and biology.

REFERENCES

- Sinz, A. (2005) Chemical cross-linking and FTICR mass spectrometry for protein structure characterization, *Anal. Bioanal. Chem.* **381**, 44–47.
- Schulz, D. M., Ihling, C., Clore, G. M., and Sinz, A. (2004) Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTICRMS: direct demonstration of multiple binding modes, *Biochemistry* **43**, 4703–4715.
- Dihazi, G. H., and Sinz, A. (2003) Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry, *Rapid Commun. Mass Spectrom.* **17**, 2005–2014.
- Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5802–5806.
- Bennett, K. L., Kussmann, M., Björk, P., Godzwon, M., Mikkelsen, M., Sørensen, P., and Roepstorff, P. (2000) Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—a novel approach to assess intermolecular protein contacts, *Protein Sci.* **9**, 1503–1518.
- Rappsilber, J., Siniosoglou, S., Hurt, E. C., and Mann, M. (2000) A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry, *Anal. Chem.* **72**, 267–275.
- Worthington, A., Rivera, H., Jr, Torpey, J., Alexander, M., and Burkart, M. (2006) Mechanism-based protein crosslinking probes to investigate carrier protein mediated biosynthesis, *ACS Chem. Biol.* **1**, 687–691.
- Rawlings, B. J. (1998) Biosynthesis of fatty acids and related metabolites, *Nat. Prod. Rep.* **15**, 275–308.
- Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: a millennium review, *Nat. Prod. Rep.* **18**, 380–416.
- Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology polyketide and non-ribosomal peptide antibiotics: login, machinery, and mechanisms, *Chem. Rev.* **106**, 3568–3496.
- Dreier, J., Shah, A. N., and Khosla, C. (1999) Kinetic analysis of the actinorhodin aromatic polyketide synthase, *J. Biol. Chem.* **274**, 25108–25112.
- Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006) Architecture of a fungal fatty acid synthase at 5 Å resolution, *Science* **311**, 1263–1267.
- Maier, T., Jenni, S., and Ban, N. (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution, *Science* **311**, 1258–1262.
- Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7 angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11124–11129.

Evolving Sensitivity

Hao Song and Lingchong You*

Department of Biomedical Engineering and Institute for Genome Sciences and Policy, Duke University, Durham, North Carolina 27708

De novo engineering of gene circuits with well-defined functions is at the heart of the nascent field of synthetic biology (1–7). Such engineered systems may offer insights into underlying design principles of biological control and lead to innovative applications in biotechnology, computing, and medicine.

However, engineering gene circuits with desired, nontrivial properties is challenging. When assembled, a circuit may not function as designed because of improper or unbalanced interactions among its components (DNAs, RNAs, and proteins). Multiple rounds of circuit revision and characterization are common practice. To this end, modeling is often used to explore how qualitative and quantitative behaviors of the circuit depend on its parameters. Guided by modeling, one may choose circuit components with specific kinetic properties. If existing components do not satisfy the design criteria for optimal circuit performance, one may choose to modulate their properties by structure-based rational design. However, rational design is often limited by the lack of detailed knowledge of the structure–function relationships of individual circuit components.

Directed evolution serves as a powerful alternative that complements the rational design approach. It entails the generation of a large pool of mutant components *via* random mutagenesis and subsequent screening and selection of mutants with desired function. It relies less on detailed knowledge about the structure–function relationships of the components to be opti-

mized. Traditionally, directed evolution has proved highly efficient in optimizing the function of a wide variety of biomolecules, without the need to resort to detailed understanding of their structures (8). In recent years, however, applications of directed evolution for the optimization of a circuit or its individual components have increased (9–14). An important element of this development is that components evolved in one context may expand the design scope of gene circuits in a different context (15). In this issue, Sun and co-workers (16) present an elegant study on optimizing positive feedback loops (PFLs) by directly evolving an individual component. The evolved circuits exhibit significantly reduced activation thresholds in response to *N*-(3-oxo-hexanoyl)-homoserine lactone (OHHL), an inducer. The improved response sensitivity of these circuits may be helpful for potential applications in metabolic engineering and gene therapy.

In the first PFL (PFL1) (16), upon binding the OHHL inducer molecule and P_{luxR} promoter, the LuxR transcription factor taken from *Vibrio fischeri* is autoactivated by its own expression in *Escherichia coli*, forming the PFL. A *gfpuv* gene upstream of *luxR* acts as a reporter for the circuit dynamics. The second circuit (PFL2) is the same as PFL1, except for the incorporation of a constitutively expressed LuxR. Both wild-type PFL1 and PFL2 exhibit ultrasensitivity in their response to OHHL. OHHL stimulation above a threshold value (OHHL₅₀, where the output GFP_{uv} level is at half of its maximum) can elicit a significant amplification of

ABSTRACT Engineering gene circuits with novel functions holds promise for broad applications in biology, engineering, and medicine. Directed evolution complements rational design as an important strategy for optimizing gene circuits and circuit elements.

*Corresponding author,
you@duke.edu.

Published online December 15, 2006

10.1021/cb6004596 CCC: \$33.50

© 2006 by American Chemical Society

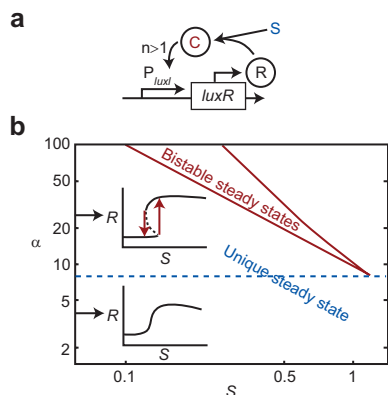


Figure 1. Mathematical modeling suggests the possibility of generating bistable behavior with the starting and evolved positive feedback circuits. a) Schematic of the PFL under the control of a signal OHHL (S). The transcription factor LuxR (R), upon binding with S, forms an activated complex C that forms a homodimer and binds to promoter P_{luxI} to activate LuxR expression, effecting positive feedback. This system can be modeled by two dimensionless equations: $dR/d\tau = \alpha(C^n/(1 + C^n)) - \beta R - kSR + \alpha_0$; $dC/d\tau = kSR - C$. Here, S, R, and C in the equations represent concentrations of OHHL, LuxR, and complex C, respectively. Synthesis of R is modeled by a single Hill function, lumping transcription and translation together. The Hill coefficient is set to be 2 to reflect fast dimerization of C and binding of the dimer to P_{luxI} . α is the synthesis rate constant of R due to feedback. α_0 is the basal or constitutive synthesis rate constant of R. β is the degradation rate constant of R. k is the binding constant of R with S. To make the model dimensionless, time is scaled with respect to the decay rate constant of C, and concentrations are scaled with respect to the half activation threshold of C. Biologically feasible parameter values are used for bifurcation analysis, with base values of $\beta = 5$, $\alpha_0 = 1$, and $k = 1$. b) Dynamic behaviors of the PFL are determined by α . When α is larger than the critical value (blue dashed line), the system exhibits bistability (upper inset). For each α , the two red lines define the boundaries of the bistable region in terms of S.

GFPuv expression. To further improve the sensitivity of these circuits, the authors created a LuxR mutant library by directed

evolution. Favorable LuxR mutants with improved circuit behavior were subsequently screened and identified. They hypothesized that the identified LuxR mutants are most likely to increase the binding interaction with OHHL or the promoter P_{luxI} , thereby increasing the strength of the PFL.

PFLs play essential roles in diverse cellular functions, including tissue development, cell fate decision, and long-term memory (17, 18). They can result in either graded or “all-or-none” bistable responses to an external cue. For the work by Sun and colleagues (16), one may wonder how the evolved parts may impact the overall dynamics. To gain insight, we developed a simple kinetic model to analyze the switching behavior of the systems (Figure 1). When the positive feedback is weak (*i.e.*, the α value is smaller than the value of the blue dashed line, Figure 1, panel b), the circuit can only demonstrate monostable behavior, exemplified by a monotonic dependent “R vs S” curve (lower inset, Figure 1, panel b). For sufficiently strong feedback regulation, however, the circuit may demonstrate bistability (see the bistable “R vs S” curve, upper inset, Figure 1, panel b).

Bistable gene switches controlled by positive feedback were experimentally examined in a few cellular contexts (19, 20). The current data in ref 16, which were measured at the population level, cannot indicate whether the system is bistable. Additional analyses by examining single-cell behavior and detecting the presence of hysteresis are needed to elucidate these interesting dynamics. If this becomes the authors’ design goal, further rounds of evolution on LuxR, other circuit components, or the circuit might help.

Acknowledgments: Current research in the You lab is supported by the National Academies Keck Futures Initiative, the National Institutes of Health, the National Science Foundation, and the David and Lucile Packard Foundation.

REFERENCES

- Hasty, J., McMillen, D., and Collins, J. J. (2002) Engineering gene circuits, *Nature* 420, 224–230.
- Endy, D. (2005) Foundations for engineering biology, *Nature* 438, 449–453.
- Sprinzak, D., and Elowitz, M. B. (2005) Reconstruction of genetic circuits, *Nature* 438, 443–448.
- Benner, S. A., and Sismour, A. M. (2005) Synthetic biology, *Nat. Rev. Genet.* 6, 533–543.
- Andrianantonandro, E., Basu, S., Karig, D. K., and Weiss, R. (2006) Synthetic biology: new engineering rules for an emerging discipline, *Mol. Syst. Biol.* 2, 1–14.
- Voigt, C. A. (2006) Genetic parts to program bacteria, *Curr. Opin. Biotechnol.* 17, 548–557.
- Chin, J. W. (2006) Modular approaches to expanding the functions of living matter, *Nat. Chem. Biol.* 2, 304–311.
- Arnold, F. H. (2001) Combinatorial and computational challenges for biocatalyst design, *Nature* 409, 253–257.
- Yokobayashi, Y., Weiss, R., and Arnold, F. H. (2002) Directed evolution of a genetic circuit, *Proc. Natl. Acad. Sci. U.S.A.* 99, 16587–16591.
- Collins, C. H., Arnold, F. H., and Leadbetter, J. R. (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones, *Mol. Microbiol.* 55, 712–723.
- Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR, *Nat. Biotechnol.* 24, 708–712.
- Chockalingam, K., Chen, Z., Katzenellenbogen, J. A., and Zhao, H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches, *Proc. Natl. Acad. Sci. U.S.A.* 102, 5691–5696.
- Alper, H., Fisher, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering, *Proc. Natl. Acad. Sci. U.S.A.* 102, 12678–12683.
- Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S. (2002) Combinatorial synthesis of genetic networks, *Science* 296, 1466–1470.
- Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R. (2005) A synthetic multicellular system for programmed pattern formation, *Nature* 434, 1130–1134.
- Sayut, D. J., Niu, Y., and Sun, L. (2006) Construction and engineering of positive feedback loops, *ACS Chem. Biol.* 1, 692–696.
- Laurent, M., and Kellersohn, N. (1999) Multistability: a major means of differentiation and evolution in biological systems, *Trends Biochem. Sci.* 24, 418–422.
- Bhalla, U. S., and Iyengar, R. (1999) Emergent properties of networks of biological signaling pathways, *Science* 283, 381–387.
- Becskei, A., Seraphin, B., and Serrano, L. (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion, *EMBO J.* 20, 2528–2535.
- Isaacs, F. J., Hasty, J., Cantor, C. R., and Collins, J. J. (2003) Prediction and measurement of an autoregulatory genetic module, *Proc. Natl. Acad. Sci. U.S.A.* 100, 7714–7719.

Small Molecules Keep Mitotic Kinases in Check

Klaus Strebhardt*

Department of Obstetrics and Gynecology, School of Medicine, J. W. Goethe-University, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany

The Human Genome Project has revealed the existence of 518 kinases that are likely to play essential roles at virtually all stages of the life cycle of mammalian cells. A detailed (yet still incomplete) mosaic has arisen from this knowledge that depicts cell-cycle control as a series of different kinases that promote the cellular journey through each stage of the cell cycle. Significant interest has been generated by the family of polo-like kinases (Plk's), which has been highly conserved throughout evolution and strongly contributes to the fine-tuning of the mitotic network. Detailed functional studies now reveal that the Plk's have multiple functions during the entry into mitosis, centrosome maturation, bipolar spindle formation, the segregation of chromosomes and cytokinesis, and crucially, the fidelity-monitoring of checkpoint control (1–3). In humans, Plk1 is the best-characterized member of this family. Because Plk1 is associated with tumorigenesis and belongs to the family of serine/threonine kinases, an attractive target for novel chemotherapeutics, it is thought also to be a promising target for anticancer drug development (4). Small molecules that are cell-permeable and can inhibit their target kinases with fast kinetics are powerful tools for probing cell-cycle mechanisms. Indeed, two novel compounds that belong to the first generation of small-molecule inhibitors of Plk1 have allowed the investigation of mitotic phenotypes in mammalian cells at a desired level of temporal control (5, 6).

The targeted inactivation of protein kinases is often accomplished by using ATP-

competitive small-molecule inhibitors that hamper enzymatic activity by blocking ATP binding sites. However, current progress in developing such specific inhibitors for Plk1 is lacking, so the development of potent and selective compounds that target this kinase is of great interest. Although an overall structure for Plk1 is not yet available, the crystal structure of its polo-box domain and a homology-modeled structure of its kinase domain have been established (7–9).

McInnes and colleagues (5) adopted a structure-guided design approach to the identification of Plk1-specific inhibitors by using a homology model of the Plk1 kinase domain. A hypothetical 3D structure of Plk1, based on the staurosporine-bound conformation of protein kinase A, was also validated by using several micromolar Plk1 inhibitors with known structure–activity relationships. Examples of this include purvalanol A and some thiazoloanilopyrimidine compounds that dock with the active site of Plk1. Given the energetic aspects of these interactions, a conformation was identified that constitutes the optimal molecular interactions that will inhibit Plk1 kinase activity.

After the homology model was validated with various Plk1 inhibitors, LIDAEUS (a tool for rapid docking of ligands into protein binding sites) (10) was used to dock a library of small molecules with this kinase to enable the structure-guided identification of novel inhibitors. A total of 350 compounds ranked *in silico* were subsequently tested in kinase assays. Various derivatives of a benzothiazole *N*-oxide core structure, the cyclapolins, were found to inhibit Plk1 with an

ABSTRACT The remarkable progress in the discovery of small molecules that target protein kinases continues, and two recent reports have described alternative approaches to this task: homology modeling and phenotype-based screening. Compounds targeting the nucleotide binding pockets of polo-like kinases in particular provide significant new insights into the molecular mechanisms controlling cell-division phenotypes.

*Corresponding author, strebhardt@em.uni-frankfurt.de.

Published online December 15, 2006

10.1021/cb600462z CCC: \$33.50

© 2006 by American Chemical Society

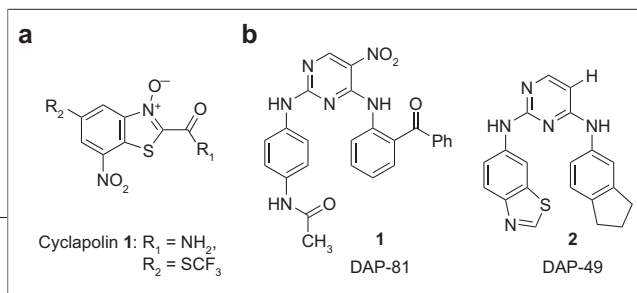


Figure 1. Chemical structures of Plk1 inhibitors. a) Cyclapolin 1, b) DAP-81 and DAP-49.

IC_{50} of 20–500 nM (Figure 1, panel a). Whether the most powerful of these inhibitors, cyclapolin 1, which acts in a non-ATP-competitive mode, interacts covalently with amino acids within the ATP binding cleft is currently the subject of further investigation by the same lab. The analysis of an additional panel of 38 protein kinases revealed that none are significantly inhibited by cyclapolin 1. The cellular effects of cyclapolin 1 were compared with the results of previous studies. Small interfering RNAs (siRNAs) were used systematically for Plk1 knockdown to study the mitotic phenotypes. In terms of consistency and threshold concentration, siRNAs are excellent tools for efficient silencing of Plk1(11–13). These investigations revealed that RNA interference (RNAi)-based depletion of Plk1 causes a mitotic arrest associated with a failure to form normal metaphase plates. This suggested that Plk1 is required for chromosome congression, anaphase initiation, and cytokinesis. Moreover, almost all RNAi Plk1 cells lack focused poles and show a high frequency of monopolar spindles. These Plk1-depleted cells also contain two centriole pairs that often are separated from each other but frequently are distributed randomly in the area of the spindle or even outside of it. Because γ -tubulin recruitment to the centrosomes depends on Plk1 (12), these results suggest that Plk1 is required for the stable association of centrosomes with spindle microtubules. Furthermore, the presence of Mad2 at many kinetochores in these knockdown cells suggests that they fail to enter anaphase because of the activation of the spindle checkpoint. Addition of Hesperadin, a small molecule that suppresses the spindle checkpoint by inhibiting Aurora B kinase, did not induce anaphase chromosome movements in Plk1-depleted cells. This suggests that Plk1 activity is

required for the generation of the spindle pulling

forces that stabilize the microtubule–kinetochore interactions. Interestingly, cyclapolin-1-induced mitotic effects differ from the cellular phenotype associated with Plk1 depletion via RNAi. Treatment of HeLa cells with cyclapolin 1 induces less severe mitotic defects in terms of the proportion of cells with a tetraploid chromosome content, with monopolar spindles and unfocused poles observed (Figure 2). Because of a collapse or an elongation failure in cyclapolin-1-treated cells, most spindles are shortened to <50% of the lengths observed in the control cells. The treatment of *Drosophila melanogaster* S2 cells with cyclapolin 1 induces a phenotype resembling the original *polo*¹ mutant cells with reduced catalytic activity. Whether microtubule nucleation from the chromosomes is involved in the formation of abnormal spindles is currently unknown. However, the use of an *in vitro* system for the

analysis of microtubule nucleation from the centrosomes has revealed convincingly that the inhibition of Plk1 in human cells and Polo in *Drosophila* cells by cyclapolin 1 perturbs the function and organization of the mitotic centrosome.

A promising alternative approach to the activity-based screen described above has been proposed by Peters *et al.* (6), who have probed the cell-division space by using a phenotype-based screen of chemical libraries to further elucidate the mechanisms underlying the regulation of cell division. The analysis of the structure–activity profiles of the diaminopyrimidines (DAPs) revealed that these compounds are “privileged scaffolds” that target different enzymes, including kinases. To explore and to span the cell-cycle phenotype space, the authors designed a collection of 100 DAPs that carry amine substituents (alkyl, cycloalkyl, aromatic, and heteroaromatic) at positions 2 and 4 of the core pyrimidine. This chemistry achieves a maximum level of diversity by altering the structure, polarity,

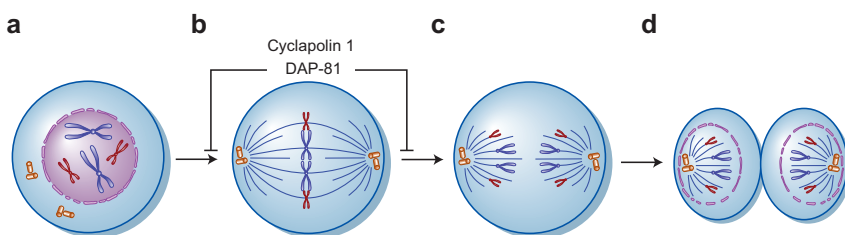


Figure 2. Plk1-related pharmacology of mitosis. a) During early prophase, centrosomal duplication occurs. Subsequently, centrosomes migrate to opposite poles, accompanied by a breakdown of the nuclear membrane during late prophase/early prometaphase. b) The centrosomes are at opposite poles of the cell. During metaphase, the chromosomes, now at their most highly coiled and condensed, become arranged on a plane equidistant from the two poles known as the metaphase plate. For each chromosome, the kinetochores of the sister chromatids face the opposite poles, and each is attached to a kinetochore microtubule coming from that pole. c) Anaphase begins when the duplicated centromeres of each pair of sister chromatids separate, and the now-daughter chromosomes begin moving toward opposite poles of the cell because of the action of the spindle. At the end of anaphase, a complete set of chromosomes has assembled at each pole of the cell. In cyclapolin-1-treated cells, the mitotic spindle remained short in anaphase or collapsed and then recovered with less-focused microtubules at one spindle pole. DAP-81 induces the formation of monopolar spindles in human cells. This suggests that cells reach a prometaphase-like state; that is, cells might have failed to form normal metaphase plates, as previously observed in RNAi Plk1-depleted cells. d) Cytokinesis is the process whereby the cytoplasm of a single cell is divided into two daughter cells.

Plk1 is required for the maintenance of the mitotic spindle as the addition of an inhibitory compound induces spindle collapse.

and hydrogen-bonding capacities of the molecules. BS-C-1 (African green monkey epithelial kidney) cells were treated with each DAP at a concentration of 20 μM , and compounds were selected that disrupt cell division but not microtubule organization during interphase. A small number of DAPs were found to induce a broad spectrum of cell-cycle phenotypes, an indication that compounds covering a limited chemical space can span a large cell-cycle phenotype space. Mitotic phenotypes were selected that already could be correlated to the function of specific enzymes to identify the target proteins of the analyzed compounds. For example, the inhibition or depletion of Plk1 was previously shown to correlate with a “monopolar spindle phenotype”. Significantly, the proportion of cells with monopolar spindles was increased in a dose-dependent manner by DAP-81, which inhibits Plk1 *in vitro* with an IC_{50} of 0.9 μM (Figure 1, panel b). Also, the structurally related compound DAP-49 induces the same phenotype but cannot inhibit Plk1, even at concentrations of up to 100 μM (Figure 1, panel b).

To examine the properties and the specificity of this novel compound, the study by Peters *et al.* (6) investigated whether (i) DAP-81 recapitulates the central phenotypic features of Plk1-depleted cells, (ii) DAP-81 inhibits the phosphorylation of specific Plk1 substrates, and (iii) the concentrations of DAP-81 that induce mitotic defects and inhibit the phosphorylation of Plk1 substrates also abrogate related kinases. Treatment of the human osteosarcoma cell line U2OS with 25 μM DAP-81 for 4 h was found to suppress the recruitment of γ -tubulin to the centrosomes and induce monopolar spindles in a high percentage of treated cells, a phenotype often observed after Plk1 depletion by RNAi (Figure 2). Centrosomes were also often found outside of the microtubule arrays, as previously observed in siRNA-treated cells (13). Moreover, in quantitative immunofluorescence assays, a dose-dependent reduction of phospho-Cdc25C

with saturating concentrations at 6 μM was observed. DAP-81 at lower concentrations induces perturbed bipolar mitotic spindles with reduced spindle–pole separation.

Because DAP-81, like many other ATP-competitive inhibitors, targets the nucleotide pocket of kinases, it was speculated that it might not be specific to a particular Plk-family member. Thus, in Peters *et al.* (6), the effects of DAP-81 on the Aurora kinases and cyclin-dependent kinases were evaluated. Phosphorylation of Ser10 on histone H3, a marker for Aurora kinase activity, was reduced by 20% after treatment with 6.3 μM DAP-81. In addition, whereas Aurora inhibitors such as Hesperadin can cause an exit from mitosis (14), DAP-81-treated PTK α T (rat kangaroo kidney) cells remain in mitosis at concentrations that induce the formation of monopolar spindles. Whether this off-target effect of DAP-81 is mitotically relevant to other cell lines should be addressed. This study (6) provides new information on the microtubule and chromosome dynamics of live cells by clearly showing that Plk1 is required for the maintenance of the mitotic spindle as the addition of an inhibitory compound induces a spindle collapse; that is, the two poles approach each other and the kinetochore fibers become shorter. Astral microtubules that extend from the spindle poles to the cellular cortex were also affected by the reversible DAP-81-mediated inhibition of Plk1 by showing increasing thickness. The study implies that the spindle collapse seen in these experiments seems to involve altered forces that keep the poles apart but not increasing forces that pull the chromosomes to the poles.

The error-free distribution of chromosomes to the daughter cells requires the precise coordination of multiple cell-cycle processes. The compounds reported by McInnes *et al.* (5) and Peters *et al.* (6) add promising novel candidates to the existing small family of Plk inhibitors comprising a benzthiazole *N*-oxide, a styryl benzylsulfone (ON 01910), and a dihydropteridinone

derivative (BI2536) (9, 15–18). Both novel reports also describe approaches to the identification of compounds that may become powerful tools for the elucidation of the molecular mechanisms regulating mitotic spindle dynamics and for cancer drug development. Homology modeling followed by *in silico* docking of known inhibitors by McInnes and colleagues (5) proved to be a valid strategy for identifying derivatives of a benzthiazole *N*-oxide core structure as highly selective inhibitors of Plk. Treatment of cells with cyclapolin 1 induces the formation of transient monopolar structures that reconstitute bipolar spindles and form short spindles. In a parallel study, DAP-81 was identified as another inhibitor of Plk1 by using a phenotype-based screen aimed at the identification of probes that target the phenotypic space of the cell cycle. Exploration of the chemical space around a known bioactive scaffold also enabled the identification of several compounds with intriguing properties. In cells treated with DAP-81, a collapse of mitotic spindles into monopolar structures was also observed, but under these conditions the reestablishment of bipolarity does not seem to occur. These observations suggest that the efficient inhibition of Plk1 by small molecules such as DAP-81 that target the nucleotide binding site of the enzyme seems to be sufficient to induce a mitotic phenotype, which was previously observed in Plk1-depleted cells after RNAi treatment (12, 13). The presence of a DAP-81-bound, enzymatically inactive Plk1 with a functional polo-box does not contribute enough to warrant proper spindle formation.

These new data on small molecules targeting Plk1 in a rapid manner further our understanding of the cell-cycle phenotype space considerably. Little was previously known about the impact of Plk1 on microtubules extending from the spindle poles to the cell periphery. In DAP-81-treated cells, these microtubules seem to be bundled into fibers that gain intensity when examined

confocally. Both studies described here raise the question of how Plk1 regulates the stabilization and destabilization of components of the mitotic spindle by phosphorylation, such that precise spatial and temporal dynamics can be achieved: how does Plk1 regulate the equilibrium between polymerization and depolymerization of γ -tubulin subunits during microtubule dynamics? The use of novel inhibitors of this kinase will likely enable researchers to further dissect the force-generating mechanisms that operate between the cellular cortex, centrosomes, and kinetochores along the time axis of mitotic progression. Both studies demonstrate that centrosomes lose their ability to nucleate microtubules and are often mislocalized in the context of the mitotic spindle when Plk1 is lost. Inactivation of Plk1 inhibits γ -tubulin recruitment to the centrosomes and the loss of mitosis-specific phosphoprotein monoclonal antibody 2 (MPM2) activity, which might be partially reflected also by the dephosphorylation of aspartic acid (abnormal spindle protein). Because the loss of the γ -tubulin ring complex from centrosomes correlates with the exit from mitosis, these novel inhibitors will also allow us to decipher the cellular machinery that triggers the centrosomal transition from mitosis to interphase.

According to recent estimates, >25% of all pharmaceutical drug targets are protein kinases. This has precipitated many new drug discovery programs designed to search for new chemical scaffolds that may potentially also target protein kinases. Increasing our fundamental knowledge of the mechanisms underlying the proliferation of cancer cells is also leading to the identification of targets that can be therapeutically manipulated to treat cancer cells. Many scientists have proposed Plk1 as an attractive target for anticancer drug development (4). Hence, the small-molecule inhibitors of Plk1 described by McInnes *et al.* (5) and Peters *et al.* (6) might provide the basis for the development of novel anticancer agents. It will be impor-

tant therefore to determine the effects of these new inhibitors upon other members of the Plk family that have highly homologous kinase domains. For example, inhibition of Plk3 by expression of a kinase-defective form can also induce the disorganization of microtubule structures, followed by a G2/M arrest (19). Furthermore, partial inhibition of Plk4 leads to tumor formation in transgenic mice (20). Thus, a careful evaluation of the specificity of any novel compounds that target kinases such as Plk1 is needed to address their suitability for the clinic.

Acknowledgments: I thank Yves Matthes and Sven Kappel for their assistance.

REFERENCES

- Glover, D. M., Hagan, I. M., and Tavares, A. A. (1998) Polo-like kinases: a team that plays throughout mitosis, *Genes Dev.* **12**, 3777–3787.
- Barr, F. A., Sillje, H. H., and Nigg, E. A. (2004) Polo-like kinases and the orchestration of cell division, *Nat. Rev. Mol. Cell. Biol.* **5**, 429–440.
- van de Weert, B. C., and Medema, R. H. (2006) Polo-like kinases: a team in control of the division, *Cell Cycle* **5**, 853–864.
- Strebhardt, K., and Ullrich, A. (2006) Targeting polo-like kinase 1 for cancer therapy, *Nat. Rev. Cancer* **6**, 321–330.
- McInnes, C., Mazumdar, A., Mezna, M., Meades, C., Midgley, C., Scaerou, F., Carpenter, L., Mackenzie, M., Taylor, P., Walkinshaw, M., Fischer, P. M., and Glover, D. (2006) Inhibitors of Polo-like kinase reveal roles in spindle-pole maintenance, *Nat. Chem. Biol.* **2**, 608–617.
- Peters, U., Cherian, J., Kim, J. H., Kwok, B. H., and Kapoor, T. M. (2006) Probing cell-division phenotype space and Polo-like kinase function using small molecules, *Nat. Chem. Biol.* **2**, 618–626.
- Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain, *Cell* **115**, 83–95.
- Cheng, K. Y., Lowe, E. D., Sinclair, J., Nigg, E. A., and Johnson, L. N. (2003) The crystal structure of the human polo-like kinase-1 polo box domain and its phospho-peptide complex, *EMBO J.* **22**, 5757–5768.
- McInnes, C., Mezna, M., and Fischer, P. M. (2005) Progress in the discovery of polo-like kinase inhibitors, *Curr. Top. Med. Chem.* **5**, 181–197.
- Wu, S. Y., McNae, I., Kontopidis, G., McClue, S. J., McInnes, C., Stewart, K. J., Wang, S., Zheleva, D. I., Marriage, H., Lane, D. P., Taylor, P., Fischer, P. M., and Walkinshaw, M. D. (2003) Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop, *Structure (Cambridge, MA, U.S.)* **11**, 399–410.
- Liu, X., and Erikson, R. L. (2002) Activation of Cdc2/cyclin B and inhibition of centrosome amplification in cells depleted of Plk1 by siRNA, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8672–8676.
- Spankuch-Schmitt, B., Bereiter-Hahn, J., Kaufmann, M., and Strebhardt, K. (2002) Effect of RNA silencing of polo-like kinase-1 (PLK1) on apoptosis and spindle formation in human cancer cells, *J. Natl. Cancer Inst.* **94**, 1863–1877.
- Sumara, I., Gimenez-Abian, J. F., Gerlich, D., Hirota, T., Kraft, C., de la Torre, C., Ellenberg, J., and Peters, J. M. (2004) Roles of polo-like kinase 1 in the assembly of functional mitotic spindles, *Curr. Biol.* **14**, 1712–1722.
- Hauf, S., Cole, R. W., LaTerra, S., Zimmer, C., Schnapp, G., Walter, R., Heckel, A., van Meel, J., Rieder, C. L., and Peters, J. M. (2003) The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint, *J. Cell. Biol.* **161**, 281–294.
- Gumireddy, K., Reddy, M. V., Cosenza, S. C., Boominathan, R., Baker, S. J., Papathi, N., Jiang, J., Holland, J., and Reddy, E. P. (2005) ONO1910, a non-ATP-competitive small molecule inhibitor of Plk1, is a potent anticancer agent, *Cancer Cell* **7**, 275–286.
- Steehmaier, M., Baum, A., Solca, F., Peters, J. M., Grauert, M., and Hoffmann, M. (2005) BI 2536, a potent and highly selective inhibitor of Polo-like kinase 1 (Plk1), induces mitotic arrest and apoptosis in a broad spectrum of tumor cell lines, *Clin. Cancer Res.* **11** (Suppl.), 9147.
- Baum, A., Garin-Chesa, P., Quant, J., Colbatzky, F., Munzert, G., Hoffmann, M., and Steegmaier, M. (2005) In vivo activity of BI 2536, a potent and selective inhibitor of the mitotic kinase Plk1, in a range of human cancer xenograft models, *Clin. Cancer Res.* **11** (Suppl.), 9146.
- Mross, K., Steinbild, S., Frost, A., Hedborn, S., Rentschler, J., Kaiser, R., Trommehauser, D., Stehle, G., and Munzert, G. (2005) A phase I single dose escalation study of the Polo-like kinase 1 (Plk1) inhibitor BI 2536 in patients with advanced solid tumors, *Clin. Cancer Res.* **11** (Suppl.), 9032.
- Wang, Q., Xie, S., Chen, J., Fukasawa, K., Naik, U., Traganos, F., Darzynkiewicz, Z., Jhanwar-Uniyal, M., and Dai, W. (2002) Cell cycle arrest and apoptosis induced by human Polo-like kinase 3 is mediated through perturbation of microtubule integrity, *Mol. Cell. Biol.* **22**, 3450–3459.
- Ko, M. A., Rosario, C. O., Hudson, J. W., Kulkarni, S., Pollett, A., Dennis, J. W., and Swallow, C. J. (2005) Plk4 haploinsufficiency causes mitotic infidelity and carcinogenesis, *Nat. Genet.* **37**, 883–888.

Mechanism-Based Protein Cross-Linking Probes To Investigate Carrier Protein-Mediated Biosynthesis

Andrew S. Worthington, Heriberto Rivera, Jr., Justin W. Torpey, Matthew D. Alexander, and Michael D. Burkart*

Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0358

The carrier protein is a highly conserved, small (~10 kDa), acidic domain found in fatty acid, polyketide, and nonribosomal peptide synthases. Whether incorporated into a large multidomain megasynthase (type I systems) or as isolated proteins functioning independently (type II systems), carrier proteins are highly flexible proteins (1, 2). Because of inherent technical hurdles posed by the cloning and expression of large, multidomain synthases, carrier protein participation within these systems remains largely uncharacterized, and those systems that have been studied show disordered carrier protein domains (3, 4). During polyketide, fatty acid, and nonribosomal peptide biosynthesis, the carrier protein serves as a scaffold to tether the building blocks and growing products as the component pieces are assembled and modified. The carrier protein must therefore interact with all enzyme domains responsible for loading carbon units, condensing these units, modifying the condensation product, and cleaving the final product from the synthase (Figure 1). Metabolic engineering, a major goal within the biosynthetic community, is premised upon the assumption that domains within the modular framework of these synthases may be stitched together to produce viable assembly lines, yet our understanding of basic interactions between these protein domains remains

incomplete. Progress has been hindered by the lack of structural data for many of these systems, and recent structures of these megasynthases contain substantial gaps (3–5). In such megasynthases, the acyl carrier protein (ACP) domain interacts with a variety of partner domains that can span great distances (up to ~100 Å). Sequence alignment and mutagenesis experiments have informed some features of carrier protein binding by these domains (6–8).

Modular biosyntheses require the activity of condensing enzymes to extend the molecular backbone of their products; in fatty acid and polyketide synthesis, these enzymes are ketosynthases. Prior to catalyzing the condensation reaction, the active site cysteine of the extension ketosynthases (KASI and KASII of *Escherichia coli* fatty acid synthase) accepts the growing product chain from an upstream carrier protein. The carbon backbone is then extended by decarboxylative condensation of the downstream carrier protein-tethered malonate or methylmalonate. In fatty acid synthesis, each subsequent step in condensation only occurs after the β -ketone of the growing chain bound to an upstream carrier protein becomes fully reduced by the stepwise activity of the ketoreductase, dehydratase, and enoyl reductase domains (Figure 1). Each of the enzymes in this pathway must recognize the identity of both the carrier protein and tethered substrate in order for

ABSTRACT Fatty acid, polyketide, and nonribosomal peptide biosynthetic enzymes perform structural modifications upon small molecules that remain tethered to a carrier protein. This manuscript details the design and analysis of cross-linking substrates that are selective for acyl carrier proteins and their cognate condensing enzymes. These inactivators are engineered through a covalent linkage to fatty acid acyl carrier protein *via* post-translational modification to contain a reactive probe that traps the active site cysteine residue of ketosynthase domains. These proteomic tools are applied to *Escherichia coli* fatty acid synthase enzymes, where KASI and KASII selectively cross-link ACP-bound epoxide and chloroacrylate moieties. These mechanism-based, protein–protein fusion reagents also demonstrated cross-linking of KASI to type II polyketide ACPs, while nonribosomal peptide carrier proteins showed no reactivity. Similar investigations into protein–protein interactions, proximity effects, and substrate specificities will be required to complete the mechanistic understanding of these pathways.

*Corresponding author, mburkart@ucsd.edu.

Received for review September 15, 2006 and accepted October 19, 2006.

Published online December 1, 2006

10.1021/cb6003965 CCC: \$33.50

© 2006 by American Chemical Society

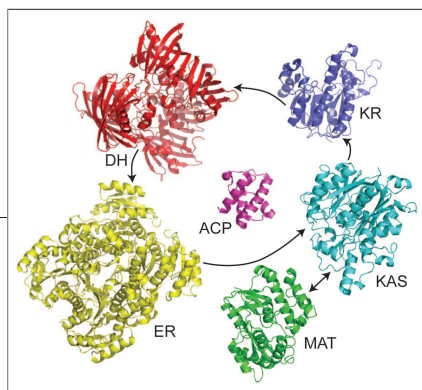


Figure 1. ACP partner domains in *E. coli* fatty acid synthesis. ACP is loaded with a malonyl moiety by malonyl-CoA:ACP transacylase (MAT). Fatty acid chain elongation at the β -ketoacyl-ACP synthase (KAS) produces an extended acyl chain bound to ACP, which is subsequently reduced to the saturated alkyl chain by the sequential action of β -ketoacyl-ACP reductase (KR), β -hydroxyacyl-ACP dehydratase (DH), and enoyl-ACP reductase (ER). The fully reduced acyl-ACP then loads the ketosynthase for another round of chain extension, or the product is incorporated into other metabolic pathways.

accurate processivity to occur. Our initial studies focus on the *E. coli* fatty acid ketosynthase elongation enzymes, KASIII, which catalyzes the initial condensation reaction between acetyl-coenzyme A (CoA) and acyl-ACP; KASI, which extends the fatty acid chain from C4 to C16; and KASII, which catalyzes the subsequent elongations. The crystal structure of each ketosynthase has been solved (9–11), but the interactions with ACP have been illustrated only from modeling studies of KASIII and ACP (12). Post-translational modification of carrier proteins *via* phosphopantetheinyl-transferase (PPTase) is required for all modular synthase activity. We have previously demonstrated the selective use of CoA analogues to modify carrier proteins within *in vitro* and *in vivo* contexts (13, 14). To study the interaction between the ketosynthase domain and ACP, we synthesized pantetheine analogues containing terminal moieties that serve as irreversible cross-linking reagents.

We began by modeling our target on the activity of cerulenin (**1**), a fungal metabolite well-known to serve as a mechanism-based inactivator of fatty acid and polyketide ketosynthase domains (Figure 2, panel a) (15–17). This inhibition occurs when the active site cysteine of the ketosynthase attacks the α -amidoepoxide moiety of ceru-

lenin, forming an irreversible covalent adduct with the enzyme. We reasoned that a similar moiety could be covalently attached to a carrier protein domain *via* promiscuous CoA metabolic uptake and PPTase-mediated carrier protein modification. This system would serve as a tool to probe protein–protein interactions between carrier proteins and ketosynthases. Utilizing our ability to attach almost any synthetic moiety to the terminus of a modified carrier protein, we could synthetically insert a mechanism-based inactivating functionality with selective distance from the modified serine residue of a carrier protein domain. We began with the design of a simple epoxide that could be installed through the addition of allylamine to pantothenic acid, or vitamin B5. The synthesis of the epoxy-pantetheine analogue **2** (Figure 2, panel a) was achieved by allylamine coupling with isopropylidene-protected pantothenic acid **3** using standard peptide coupling conditions, followed by acidic deprotection and epoxidation by dimethyldioxirane (DMDO) in acetone, to give **2** in an unoptimized 14% overall yield.

Initial analysis of compound **2** indicated that the epoxide moiety hydrolyzed slowly under aqueous-buffered conditions. This led us to follow examples of rationally designed cysteine protease inhibitors; we also chose to design Michael acceptors appended to the pantetheine backbone as more stable probes of ketosynthase activity (18). Simple acrylamide pantetheine analogues were identified as potential thiol traps *via* 1,4-conjugate addition. Here, we chose analogues that would exhibit slow reactivity with nonspecific nucleophiles and selective reactivity with activated cysteine nucleophiles. In an effort to eliminate retroaddition of the inactivated complex, β -chloroacrylamide-containing pantetheine analogues **4** and **5** were chosen as targets for the ability to undergo tandem 1,4-conjugate addition and β -chloro-elimination to yield irreversible complexes. The *cis* and *trans* analogues were synthesized from the previously

described amine **6** (19) *via* coupling with chloroacrylate under standard conditions and acidic deprotection to give **4** and **5**, each in an unoptimized yield of 23% for the two steps (Figure 2, panel b).

The pantetheine analogues thus created were analyzed for activity with CoaA (Pank), the first enzyme in the CoA biosynthetic pathway that serves as the gatekeeper and carries out the rate-limiting step in the biosynthesis of CoA (20). Should compounds **2**, **4**, and **5** serve as acceptable substrates with CoaA, we can assume that reactions with the partner enzymes for attaching 4'-phosphopantetheine to carrier proteins will proceed smoothly *in vitro* (13). In a coupled assay that monitors the consumption of ATP, these substrates showed $k_{\text{cat}}/K_{\text{m}}$ values comparable to that of pantetheine, but much lower than the natural substrate pantothenate (see Supplementary Table 1). These arise from similar turnover numbers (k_{cat}) but elevated binding constants (K_{m}) compared to those of pantothenate. Importantly, no enzyme inactivation by the electrophile-containing pantetheine analogues was detected. This established that **2**, **4**, and **5** should be acceptable substrates for *in vitro* conversion to CoA analogues and subsequent PPTase attachment onto carrier proteins. Pantetheine analogues **2**, **4**, and **5** were then coupled with purified *E. coli* fatty acid apo-ACP *via* a one-pot CoA enzymatic synthesis and PPTase transfer through the tandem activity of recombinant *E. coli* CoaA, CoaD, CoaE, and *Bacillus subtilis* Sfp (13, 21). To solutions of modified *crypto*-ACPs **7**, **8**, and **9**, *E. coli* recombinant KASI, KASII, and KASIII were added (Figure 3). Where each enzyme performs acyltransferase activity with acyl-ACP, we anticipated the formation of a fusion complex through the mechanism-based inactivation of each ketosynthase from active site cysteine attack of the epoxide in **7** to give **10** or tandem Michael addition/ β -elimination with **8** or **9** to give **11** (Figure 3). KASIII performs its substrate loading, or acyltrans-

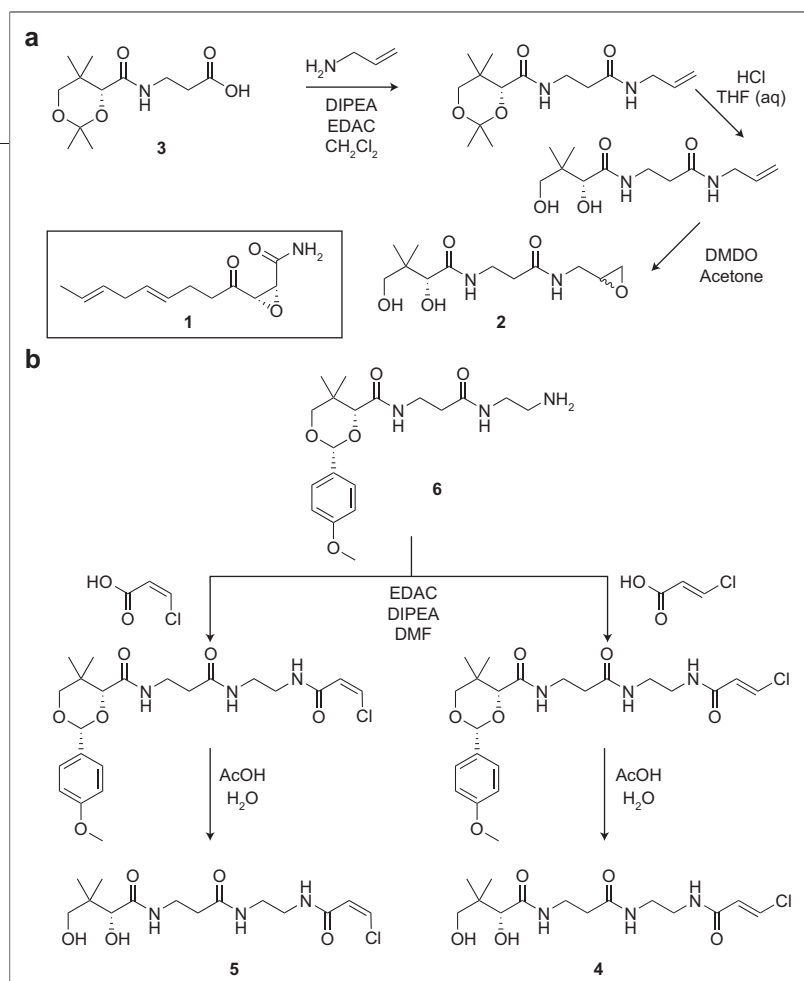


Figure 2. Pantetheine analogue synthesis. a) Synthesis of epoxide-functionalized pantetheine **2**. Cerulenin **1** serves as our model for ketosynthase inactivation. Isopropylidene-protected pantothenic acid **3** was coupled to allylamine. Subsequent deprotection and epoxidation with DMDO afforded **2**. b) Synthesis of β -chloroacrylate-pantetheine analogues **4** and **5**. Compound **6** was coupled to *cis*- and *trans*-3-chloroacrylic acid, followed by deprotection, to afford **4** and **5**.

ferase, step with acetyl-CoA as the acyl donor, in contrast to KASI and KASII, which utilize carrier protein-bound moieties. Therefore, KASIII would not be expected to form a fusion construct with these tools. However, KASI and KASII, which perform their acyltransferase step with acyl-ACP, should be cross-linked to ACP. The resulting complexes KASI-ACP and KASII-ACP would therefore contain an irreversibly blocked ketosynthase activity, thus abrogating fatty acid processivity.

The one-pot incubation of **2**, **4**, or **5** with the CoA biosynthetic enzymes, the PPTase *Sfp*, ACP, and either of the ketosynthases KASI or KASII generates irreversible covalent cross-linking between ACP and the ketosynthase domain (Figure 4). These products resulted in observed gel shifts of the keto-

synthase band from an observed mass ~ 50 kDa for the ketosynthase to ~ 80 kDa for the ACP-KAS complex when analyzed by SDS-PAGE (Figure 4, panel b). Extra bands seen in the regions of KASI and ACP-KAS of our negative control (Figure 4, panel b, lanes 1c and 2c) were contaminants that eluted with our His-tagged proteins during metal affinity chromatography. The large gel shift of the complex ACP-KAS would be expected, as the acidic ACP is known to run at an observed mass greater than ~ 20 kDa on SDS-PAGE (22). Nevertheless, the unusually large gel shift of the ketosynthase upon ACP binding merited further investigation, discussed below. SDS-PAGE analysis confirmed that no other enzymes in the mixture interacted with the modified *crypto*-ACP **7**,

8, or **9**. In addition, no interaction between modified ACP and KASIII was detected (data not shown). While each of the compounds gave the same cross-linking effect, the chloroacrylamide compounds **4** and **5** gave superior efficiency in cross-linking than the epoxide **2** in their interactions with both KASI and KASII. This is likely due to slow hydrolysis of the epoxide prior to interaction with the ketosynthase active site cysteine. However, it is also possible that the extended reactive centers of **4** and **5** favor ACP interaction. The site of nucleophilic attack for both **4** and **5** is two carbon lengths longer than the normal acylated pantetheine, while the terminal epoxide of **2** is one carbon length shorter. This chain length preference by the ketosynthase is currently being investigated. Nevertheless, these results suggest rather permissive acyl-ACP substrate loading by the *E. coli* fatty acid ketosynthases. In addition, there is a small but observable preference by both ketosynthases for the *trans*-isomer **4** over the *cis*-compound **5**, as visualized by band intensity. This preference has been observed in previous studies of structurally related cysteine proteases (18).

To investigate the unusually large gel shift of the ketosynthase band on SDS-PAGE, cross-linked proteins **10** and **11** were analyzed using in gel digestion followed with MALDI-TOF tandem mass spectrometry to verify the identity of each enzyme partner as well as the site-specificity of the cross-linking reaction. Considerable portions of both the ACP and each ketosynthase were detected (see Supplementary Figure 1). Importantly, active site residues Ser36 for ACP, Cys163 for KASI, and Cys164 for KASII were not identified in three independent analyses, suggesting the covalent modification had formed as designed in a mechanism-based manner. In addition, the lack of interaction between modified ACP and other enzymes in the one-pot reaction further supported the specific nature of the cross-linking between Ser36 of ACP and the active

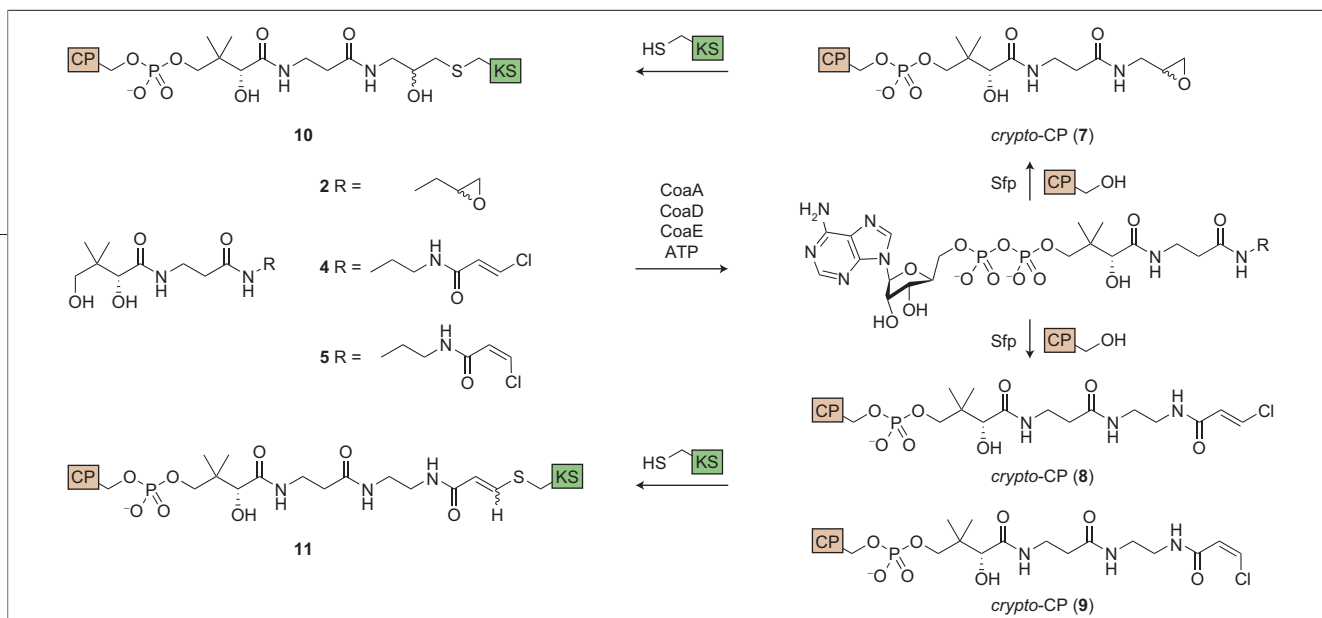


Figure 3. The one-pot formation of carrier protein–ketosynthase (CP–KS) complexes *in vitro*. Pantetheine analogues **2**, **4**, or **5** are incubated with CoaA, CoaD, and CoaE in the presence of ATP to generate CoA derivatives. Apo-CP is modified by Sfp to incorporate the functionalized pantetheine arms of these derivatives into *crypto*-CPs **7**, **8**, and **9**. Reaction of *crypto*-CP with the active site cysteine of a ketosynthase domain generates the covalent cross-link between carrier protein and ketosynthase.

site cysteine of KASI and KASII. Because of the nature of the proteomic analysis used, we were unable to observe a peptide fragment representative of the region incorporating both active sites. We have chosen to address this issue using protein crystallographic studies, which are forthcoming. Finally, the absence of reactivity between ACP and the conserved active site cysteine of KASIII suggests that the ACP binding motif of KASI and KASII is essential for substrate loading, while KASIII, which does bind ACP subsequent to priming by acetyl-CoA, may undergo a conformational change upon substrate loading. It is also probable that KASIII reacted with residual CoA analogues formed *in situ* within the one-pot reaction.

Because KASIII utilizes acetyl-CoA as an acyl donor, the absence of cross-linking by KASIII with these carrier protein analogues indicates that the cross-linking activity may be used to visualize protein–protein interactions between enzymes that demonstrate partner reactivity (Figure 1). To establish this potential, we investigated the use of pantetheine analogue **4** for specificity of KASI with recombinant carrier protein domains active in natural product biosyntheses. These included two carrier proteins from nonribosomal peptide synthetases, VibB (from *Vibrio cholerae*) and EntB (from *E. coli*), as well as two carrier proteins from type II polyketide synthetases, FrenACP (from *Streptomyces roseofulvus*) and OtcACP

(from *Streptomyces rimosus*) (Figure 4, panel c). Cross-linking with KASI was seen only with FrenACP and OtcACP, while VibB and EntB remained unmodified. Clear preference for *E. coli* ACP could be seen by comparison of intensity in the cross-linked bands. This result reflects the well-known sequence and activity-based homology between type II fatty acid ACPs and type II polyketide synthetase ACPs. Indeed, type II polyketide ACPs are known to catalyze transformations with type II fatty acid synthetase machinery from *E. coli* and streptomycete hosts (23–25). A lack of cross-linking between KASI and nonribosomal peptide carrier proteins indicates a distinct specificity in protein–protein interaction that is not satisfied with these pairings. The details of these interactions may be better understood through structural studies. Efforts to further understand the effect of specific residues upon protein–protein interactions in these modular synthetases remain essential toward understanding their connected activity.

Here, we have shown the utility of chemo-enzymatic synthesis of post-translationally modified carrier proteins to generate a mechanism-based cross-linking reagent. This approach has the potential for general applicability to all systems involving carrier protein-mediated acyl transfer. These molecules may prove to be potent mechanism-based inactivators of type II fatty acid biosynthetic machinery *in vivo*. We also foresee

the application of similar cross-linking experiments to probe substrate selectivity of ACP-mediated synthetases *via* modified carrier proteins.

METHODS

Synthetic protocols, compound characterization, and additional experimental procedures may be found in the Supporting Information.

Kinetics. Kinetic analysis was performed according to the protocol of Strauss and Begley (26).

Cross-Linking of Recombinant ACP and Ketosynthase. To compare **2**, **4**, and **5** by SDS-PAGE, we used the following procedure. To a 50 mM potassium phosphate, pH 7.0, buffer with 12.5 mM Mg₂Cl₂ and 8 mM ATP were added ACP (10 μg, 5-fold excess), ketosynthase (10 μg), CoaA (1 μg), CoaD (1 μg), CoaE (1 μg), and *B. subtilis* Sfp (10 μM). To the mixture, we added **2**, **4**, and **5** (2 mM), and the mixture was incubated at 37 °C for a minimum of 1 h. Negative controls contained no pantetheine analogue. Reaction times ranging from 1 h to overnight were tested for these reactions, and we found that a 1 h incubation time was sufficient. Samples were run on a 9% SDS-PAGE, where cross-linked product was detected by staining with Coomassie blue stain.

To compare fatty acid, polyketide, and nonribosomal peptide carrier proteins with SDS-PAGE, we used the following procedure. To a 50 mM potassium phosphate, pH 7.0, buffer with 12.5 mM Mg₂Cl₂ and 8 mM ATP were added carrier protein (5–10 μg, 3-fold excess), KASI (10 μg), CoaA (0.1 μg), CoaD (0.1 μg), CoaE (0.1 μg), and *B. subtilis* Sfp (10 μM). To the mixture, we added **4** (4 mM), and the mixture was incubated at 37 °C for a minimum of 1 h. Samples were run on a 10% SDS-PAGE, where cross-linked product was detected by staining with Coomassie blue stain.

Acknowledgment: Funding was provided by the University of California, San Diego, Department of Chemistry and Biochemistry, NIH R01GM075797, and ACS PRF 42158-G4. H.R. was supported as an NIH NIGMS PREP scholar. We thank Betsy Komives, Jim La Clair, Joe Noel, and Mike Austin for helpful

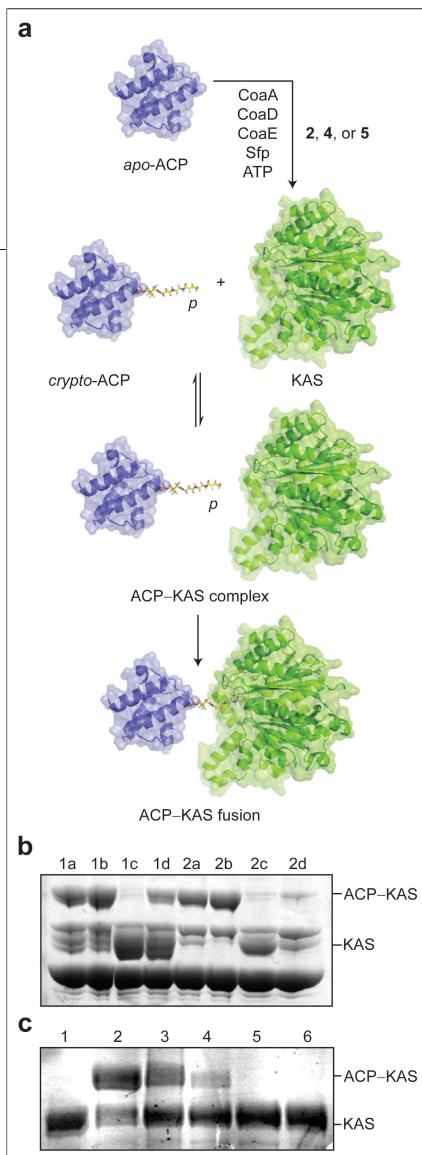


Figure 4. Application of the method for CP-KAS formation to the *E. coli* fatty acid synthase pathway. **a)** Apo-ACP is incubated with CoaA, CoaD, CoaE, and Sfp in the presence of ATP and 2, 4, or 5 to generate crypto-ACP with a pantetheine (p) analogue sidearm. Crypto-ACP reacts with the active site cysteine of the ketosynthase to form the CP-KAS fusion product. **b)** SDS-PAGE analysis of one-pot reactions to form ACP-KAS complexes. ACP was reacted with KASI in lanes 1a–d and with KASII in lanes 2a–d. Lanes designated “a” were incubated with compound 5, “b” with 4, and “d” with 2; lanes “c” were negative controls with no pantetheine analogue. **c)** SDS-PAGE analysis of one-pot reactions to form CP-KAS complexes using compound 4. KASI was reacted against water (as a negative control), *E. coli* ACP, FrenACP, OtcACP, EntB, and VibB in lanes 1–6, respectively.

discussions. Additionally, we thank Jordan Meier for providing compound 6 and H. Mori from the Japanese *E. coli* consortium for providing genes from the ASKA plasmid ORF library (27).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Parris, K. D., Lin, L., Tam, A., Mathew, R., Hixon, J., Stahl, M., Fritz, C. C., Seehra, J., and Somers, W. S. (2000) Crystal structures of substrate binding to *Bacillus subtilis* holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites, *Structure* 8, 883–895.
- Qiu, X., and Janson, C. A. (2004) Structure of apo acyl carrier protein and a proposal to engineer protein crystallization through metal ions, *Acta Crystallogr., Sect. D* 60, 1545–1554.
- Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006) Architecture of a fungal fatty acid synthase at 5 Å resolution, *Science* 311, 1263–1267.
- Maier, T., Jenni, S., and Ban, N. (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution, *Science* 311, 1258–1262.
- Tang, Y., Kim, C. Y., Mathews, I. L., Cane, D. E., and Khosla, C. (2006) The 2.7-Ångstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase, *Proc. Natl. Acad. Sci. U.S.A.* 103, 11124–11129.
- Marshall, C. G., Burkart, M. D., Meray, R. K., and Walsh, C. T. (2002) Carrier protein recognition in siderophore-producing nonribosomal peptide synthetases, *Biochemistry* 41, 8429–8437.
- Weissman, K. J., Hong, H., Popovic, B., and Meersman, F. (2006) Evidence for a protein-protein interaction motif on an acyl carrier protein domain from a modular polyketide synthase, *Chem. Biol.* 13, 625–636.
- Lai, J. R., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2006) Localized protein interaction surfaces on the EntB carrier protein revealed by combinatorial mutagenesis and selection, *J. Am. Chem. Soc.* 128, 11002–11003.
- Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., Lindqvist, Y., and Larsen, S. (1999) The X-ray crystal structure of beta-ketoacyl [acyl carrier protein] synthase I, *FEBS Lett.* 460, 46–52.
- Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G., and Lindqvist, Y. (1998) Crystal structure of beta-ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes, *EMBO J.* 17, 1183–1191.
- Qiu, X., Janson, C. A., Kostantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis, *J. Biol. Chem.* 274, 36465–36471.
- Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III, *J. Biol. Chem.* 274, 8231–8238.
- Worthington, A. S., and Burkart, M. D. (2006) One-pot chemoenzymatic synthesis of reporter-modified proteins, *Org. Biomol. Chem.* 4, 44–46.
- Clarke, K. M., Mercer, A. C., La Clair, J. J., and Burkart, M. D. (2005) In vivo reporter labeling of proteins via metabolic delivery of coenzyme A analogues, *J. Am. Chem. Soc.* 127, 11234–11235.
- D’Agnolo, G., Rosenfeld, I. S., Awaya, J., Omura, S., and Vagelos, P. R. (1973) Inhibition of fatty acid synthesis by the antibiotic cerulenin, *Biochim. Biophys. Acta* 326, 155–166.
- Omura, S., and Takeshima, H. (1974) Inhibition of the biosynthesis of leucomycin, a macrolide antibiotic, by cerulenin, *J. Biochem. (Tokyo)* 75, 193–195.
- Funabashi, H., Kawaguchi, A., Tomoda, H., Omura, S., Okuda, S., and Iwasaki, S. (1989) Binding site of cerulenin in fatty acid synthetase, *J. Biochem. (Tokyo)* 105, 751–755.
- Govardhan, C. P., and Abeles, R. H. (1996) Inactivation of cysteine proteases, *Arch. Biochem. Biophys.* 330, 110–114.
- Meier, J. L., Mercer, A. C., Rivera, H., and Burkart, M. D. (2006) Synthesis and evaluation of bioorthogonal pantetheine analogues for in vivo protein modification, *J. Am. Chem. Soc.* 128, 12174–12184.
- Karasawa, T., Yoshida, K., Furukawa, K., and Hosoki, K. (1972) Feedback inhibition of pantothenate kinase by coenzyme A and possible role of the enzyme for the regulation of cellular coenzyme A level, *J. Biochem. (Tokyo)* 71, 1065–1067.
- Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases, *Biochemistry* 37, 1585–1595.
- Rock, C. O., and Cronan, J. E. (1979) Re-evaluation of the solution structure of acyl carrier protein, *J. Biol. Chem.* 254, 9778–9785.
- Reville, W. P., Bibb, M. J., and Hopwood, D. A. (1996) Relationships between fatty acid and polyketide synthases from *Streptomyces coelicolor* A3(2): characterization of the fatty acid synthase acyl carrier protein, *J. Bacteriol.* 178, 5660–5667.
- Crosby, J., Byrom, K. J., Hitchman, T. S., Cox, R. J., Crump, M. P., Findlow, I. S., Bibb, M. J., and Simpson, T. J. (1998) Acylation of *Streptomyces* type II polyketide synthase acyl carrier proteins, *FEBS Lett.* 433, 132–138.
- Arthur, C., Cox, R. J., Crosby, J., Rahman, M. M., Simpson, T. J., Soulas, F., Spogli, R., Szafarska, A. E., Westcott, J., and Winfield, C. J. (2002) Synthesis and characterisation of acyl carrier protein bound polyketide analogues, *ChemBioChem* 3, 253–257.
- Strauss, E., and Begley, T. P. (2002) The antibiotic conversion of *N*-pentylpantothenamide results from its conversion to ethyldethia-coenzyme A, a coenzyme A antimetabolite, *J. Biol. Chem.* 277, 48205–48209.
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library, *DNA Res.* 12, 291–299.

Construction and Engineering of Positive Feedback Loops

Daniel J. Sayut, Yan Niu, and Lianhong Sun*

Department of Chemical Engineering, University of Massachusetts Amherst, 686 North Pleasant Street, Amherst, Massachusetts 01003

ABSTRACT Artificial positive feedback loops (PFLs) have been used as genetic amplifiers for enhancing the responses of weak promoters and in the creation of eukaryotic gene switches. Here we describe the construction and directed evolution of two PFLs based on the LuxR transcriptional activator and its cognate promoter, P_{luxI} . The wild-type PFLs are completely activated by 10 nM of 3-oxo-hexanoyl-homoserine lactone (OHHL). Directed evolution of LuxR increased the sensitivity of the feedback loops, resulting in systems that are completely activated at OHHL concentrations of 5 nM, or ~ 3 molecules per cell. The responses of the PFLs can also be modulated by adjusting inducer concentrations. These highly sensitive yet regulatable PFLs can be used to construct larger artificial genetic networks to gain understanding of the design principles of complex biological systems and are expected to find various applications in industrial fermentation and gene therapy.

There are many examples of positive feedback in biological networks, including the nuclear factor κ B and p53 signaling pathways, both of which regulate a diverse range of physiological functions (1, 2). Most genetic positive feedback loops (PFLs) exhibit switchlike “all-or-none” bistability, though a binary response is not the only possible output of positive feedback (3, 4). For the response of a PFL to be bistable, the system must contain some degree of ultrasensitivity (5), and its components must be balanced (6). In this regard, the combining of experimental results with stochastic models has proven to be valuable in predicting and characterizing the behavior of artificial positive feedback constructs (7, 8). The bistable responses that result from properly designed PFLs allow for their use as genetic switches to regulate gene expression. In one such example, positive feedback has been used to enhance the transcriptional activity of cell- and tissue-specific promoters, providing a method for the enhancement of gene expression in gene therapy (9, 10). The desired properties of PFLs used for controlling gene expression are that they are simple, tightly regulated, easily inducible, and nontoxic (11, 12). Because of these requirements, it is generally not effective to construct artificial PFLs using natural regulatory elements, as these elements are optimized for survival and reproduction rather than gene expression (13). As a result, an engineering strategy is frequently necessary to optimize the function of artificial genetic circuits (14). Demon-

strating this principle, we have used directed evolution to enhance the properties of PFLs constructed from bacterial quorum-sensing components.

Bacterial quorum-sensing systems are typically very sensitive to their inducers, activating gene expression at inducer concentrations $< 1 \mu\text{M}$ (15). Despite these innate sensitivities, we anticipated that more sensitive systems could be obtained by constructing PFLs based on quorum-sensing components. To construct our PFLs, we utilized the LuxR transcriptional activator and P_{luxI} promoter isolated from the quorum-sensing system of the marine bacterium *Vibrio fischeri* (16, 17). The LuxI–LuxR quorum-sensing system has been the focus of many studies due to its natural properties and possible applications in synthetic biology (18–20). In this system, LuxR activates the P_{luxI} promoter in the presence of high concentrations of 3-oxo-hexanoyl-homoserine lactone (OHHL) (21). Using these components, we constructed the first PFL, PFL1, by using P_{luxI} to regulate expression of LuxR (Figure 1, panel a). A *gfpuv* gene upstream of *luxR* determines the output of the circuit. An alternative PFL, PFL2, which is similar to PFL1 but with an additional constitutively expressed LuxR, was also constructed (Figure 1, panel b). In both of these systems, exogenous addition of OHHL causes activation of the P_{luxI} promoter by LuxR, resulting in expression of *luxR* and *gfpuv*. The increased expression of LuxR further enhances the activity of the P_{luxI} promoter, resulting in positive feedback. In

*Corresponding author,
lsun@ecs.umass.edu.

Received for review October 11, 2006
and accepted November 3, 2006.

Published online December 1, 2006
10.1021/cb6004245 CCC: \$33.50

© 2006 by American Chemical Society

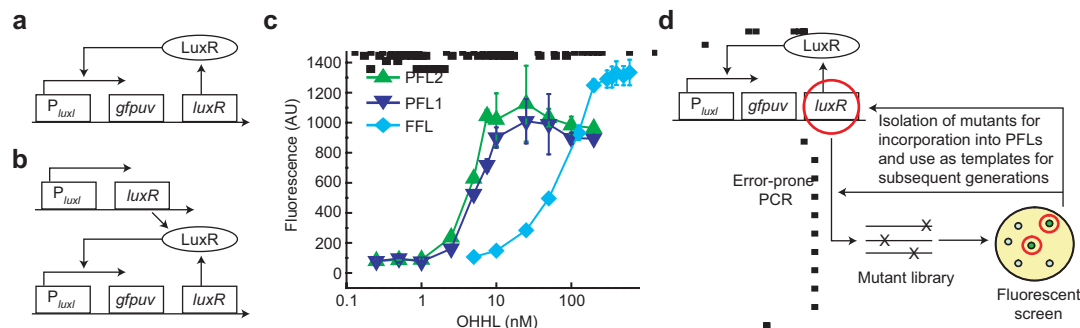


Figure 1. Design and response of PFLs. Schematic diagrams of a) PFL1 and b) PFL2. In PFL1, basal levels of LuxR are activated upon addition of OHHL, resulting in activation of the P_{luxI} promoter and establishment of positive feedback. In PFL2, intracellular concentrations of LuxR are increased by addition of a second LuxR gene constitutively expressed from a $P_{lac/ara}$ promoter. c) OHHL dose-responses of wild-type PFL1, PFL2, and reference FFL. Data measured using GFPuv fluorescence. d) Schematic diagram of directed evolution procedure. See main text for discussion.

designing PFL1, we assumed that basal expression of LuxR from the P_{luxI} promoter would be high enough to result in system activation upon addition of OHHL. In addition to the two PFLs, a reference feedforward loop (FFL) composed of a GFPuv protein regulated by a P_{luxI} promoter was also constructed to characterize the PFLs.

In populations governed by positive feedback, population heterogeneity is determined by the level of noise in the regulatory loop and the strength of the feedback (8, 22,

23). The population distributions of cells in the “on” and “off” states leads to an observed graded response at the population level (23, 24). Consequently, we characterized the PFLs and FFLs by measuring their dose-responses (25) (Figure 1, panel c) and determining $[OHHL]_{50}$, which we defined as the OHHL concentrations required by each circuit for half-maximal activation (Table 1). Both wild-type PFLs exhibit improved responses to OHHL over the FFL with a 10-fold decrease in $[OHHL]_{50}$. Complete acti-

vation of PFL1 and PFL2 occurred at OHHL concentrations of 10 nM, corresponding to ~ 6 molecules per cell (the volume of an *Escherichia coli* cell is $\sim 1 \times 10^{-15}$ L), a sensitivity that is rarely observed for inducible genetic systems. The responses of the two PFLs are similar, with PFL2 showing slightly increased activity at OHHL concentrations > 1 nM, presumably due to an increased level of LuxR in PFL2. Activation of PFL1 in the presence of OHHL indicates that basal expression from P_{luxI} allows for sufficient accumulation of LuxR for system activation, which is consistent with the native biological functions of the promoter (16). Despite the basal activity of the P_{luxI} promoter, the background activities of the two systems as measured by the GFPuv signal are minimal, and OHHL is required for activation. Overall, the high sensitivities of the feedback loops coupled with the ability to stringently control their response make them attractive alternatives to previously constructed genetic circuits for the efficient regulation and amplification of gene expression (12, 26, 27).

To determine if the sensitivity of the PFLs could be increased, we used directed evolution to enhance the activities of the LuxR transcriptional activator (Figure 1, panel d). In a series of papers, directed evolution has been used to alter LuxR’s specificity for its cognate signal molecule in order to allow for the creation of unique LuxR proteins that can be used for the construction of synthetic circuits with minimal cross talk (25, 28). Mutant LuxR proteins that are hypersensi-

TABLE 1. Characteristics of mutant LuxR proteins^a

	Nucleotide substitutions	Amino acid substitutions	$[OHHL]_{50,PFL1}$ (nM)	$[OHHL]_{50,PFL2}$ (nM)
Wild type	N/A	N/A	5.05 ± 0.30	4.48 ± 0.34
Mut64	A503T	N168I	2.08 ± 0.13	1.65 ± 0.04
Mut616	T138C	I46I	0.54 ± 0.04	0.38 ± 0.02
	A193G	K65E		
	A311C	K104T		
	T606A	D202E		
Mut620	A343G	T115A	1.41 ± 0.06	0.96 ± 0.04
	G484A	V162I		
Mut627	T90C	S30S	1.85 ± 0.04	1.33 ± 0.06
	T147A	H49Q		
	A355T	I119F		

^aThe $[OHHL]_{50}$ value for the FFL is 79.9 ± 5.53 .

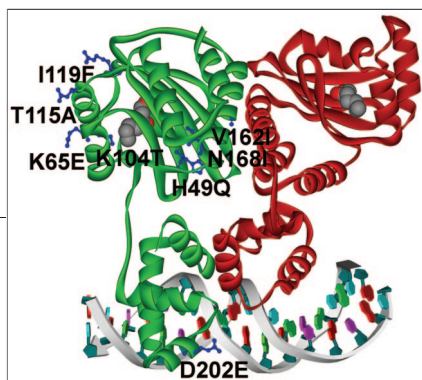


Figure 2. Mapping of identified LuxR amino acid substitutions in the crystal structure of TraR. The TraR dimer, target DNA, and OHHL signaling molecules are shown. The structures of the TraR residues that align with the amino acid substitutions in LuxR are labeled. Five of the identified mutants (H49Q, K65E, K104T, T115A, and I119F) cluster around the signal binding pocket. Two of the mutants (V162I and N168I) occur at the interface between the TraR dimers. Mutation D202E aligns to the DNA binding domain. While the amino acid sequence homology of LuxR and TraR is <20%, the overall structure homology of the two proteins is predicted to be high.

tive to OHHL were also obtained although the libraries were not screened for this property (20, 25). Therefore, we predicted that by using a similar genetic screen in the directed evolution of LuxR, we could identify more mutant LuxR proteins with hypersensitive responses to OHHL.

The mutant LuxR library was created using error-prone polymerase chain reaction (PCR) (29). Putative positive mutants were identified using a fluorescence screen of cells transformed with FFLs containing the mutant LuxR proteins (25). Screening of ~40,000 colonies identified 6 putative positive mutants. Subsequent fluorescence quantification of the mutant responses confirmed 4 of the 6 putative positive mutants as having enhanced responses over the wild-type LuxR. The amino acid substitutions for the LuxR mutants are given in Table 1. Most of the identified mutations are positioned in the amino terminal domain and presumably alter the binding ability of the OHHL signal molecule (30–32). While all of the identified mutations are novel, most of them are located near regions that were determined to be critical for LuxR function or near residues predicted to be part of the OHHL binding pocket. In particular, the K65E mutation is within three residues of two of the amino acids expected to make up part of

the OHHL binding pocket (31), and the I119F mutation is adjacent to a residue that is critical for LuxR function (32). One of the mutations, D202E, is positioned in a suspected helix–turn–helix motif (residues 200–224) and likely alters DNA binding. Alignment of the mutant sequences with the crystal structure of the LuxR homologue TraR supports the hypothesized role of the mutations in affecting OHHL binding (Figure 2).

Using the mutants identified in our screen, we constructed PFLs as described above and determined their responses (Figure 3). All of the mutant PFL1 and PFL2 systems had higher sensitivities to OHHL with $[OHHL]_{50}$ values that varied between 10- and 2.5-fold less than those of their respective wild-type feedback loops (Table 1). Except for Mut64 all of the mutant PFL2 PFLs exhibited OHHL responses that were significantly more sensitive than the mutant PFL1 PFLs, as well as improved saturation responses. The saturated responses of the PFL2 mutant systems were similar to the reference wild-type PFL2 systems. In the absence of OHHL, all mutant and wild-type PFLs had similar background level responses. Because of the highly sensitive responses of the first round of mutants, further evolution or recombination of the recovered mutants was not pursued.

Determining the exact mechanism for the increased sensitivity of the mutant PFLs is complicated because of large array of processes involved in LuxR controlled gene expression (33). To activate gene expression, the LuxR protein must bind the OHHL signal molecule and then undergo conformational changes that allow for dimerization, DNA binding, and interaction with RNA polymerase to activate transcription (17). In addition, the responses will be affected by the intercellular concentrations of the mutant LuxR proteins, which can vary from the wild type because of changes in expression, folding, stability, or solubility (25). The presence of the constitutively expressed LuxR or mutant LuxR gene in the PFL2 cir-

cuits allows for the effects of increased intracellular concentrations of the transcription factor to be examined by comparing the responses of the PFL2 circuits to the responses of the PFL1 circuits. The increased responses

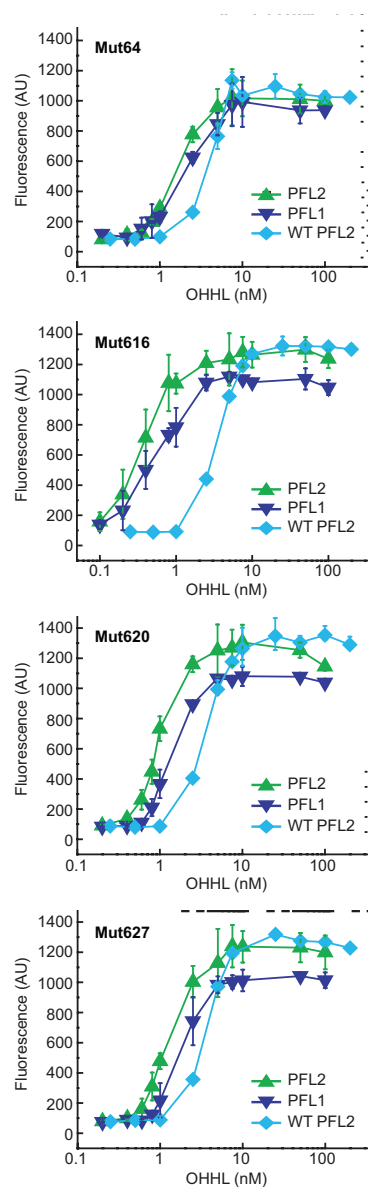


Figure 3. OHHL dose-responses of mutant PFL1 and PFL2 circuits compared against reference wild-type PFL2 circuits. WT = wild type. Data measured using GFPuv fluorescence.

of the Mut616, Mut620, and Mut627 PFL2 circuits in comparison to their respective PFL1 circuits show that increased intracellular concentrations of the LuxR protein increases the sensitivities of the PFLs and results in higher saturation levels. This suggests that the expression levels for the wild-type and mutant PFL2 circuits are similar, as there is no observable difference in the saturated responses of the mutant and wild-type PFL2 circuits. Therefore, the mutations identified in this study most likely increase the sensitivity of the LuxR protein by altering its interactions with OHHL or the regulatory elements of the P_{luxI} promoter.

Directed evolution is a powerful tool for the alteration of protein properties and has been used to tune the responses of artificial genetic circuits (14, 34). Here we have demonstrated the use of directed evolution for the enhancement of two simple PFLs. The resultant PFLs showed at least 2-fold improvements in sensitivity over the wild-type PFL, leading to complete activation of the mutant systems at low nanomolar concentrations of inducer. To obtain these engineered systems, we isolated the LuxR transcriptional activator and evolved its sensitivity to the OHHL inducer. This modular approach simplified the evolution procedure, allowing the overall sensitivity of the feedback loops to be modified using established evolutionary techniques. Similar methods that target the activators present in other positive feedback mechanisms and genetic circuits should allow for the engineering of a broad range of systems (9, 20).

Despite the extremely high sensitivities of the engineered PFLs, their responses can still be modulated by altering the OHHL concentration in the induced cultures. This property makes the constructed feedback loops attractive for applications that require regulatable gene expression, such as metabolic engineering and gene therapy (27, 35–37). A significant disadvantage of using inducible gene expression systems in large-scale fermentations is the high cost of the

chemical inducers used for activation of gene expression (38). Compared to other inducible gene regulation methods that require millimolar concentrations of inducer to completely activate gene expression (4, 12, 39), the constructed PFLs are fully activated at nanomolar concentrations of OHHL, increasing their cost effectiveness in industrial fermentations. Autonomous system activation can also be achieved by coexpression of the LuxI OHHL synthase (20, 31), bypassing the need to exogenously add OHHL. The resulting systems would regulate gene expression in a cell-density-dependent manner with many promising biotechnological applications (40, 41). Finally, quorum-sensing components homologous to those used to construct the PFLs have been used to construct functional genetic switches in mammalian and human cells (42, 43), suggesting potential application of these loops in gene therapy and tissue engineering (27, 44).

METHODS

Bacterial Strains, Plasmids, Media, and Culture Conditions.

E. coli strain DH5 α was used for all cloning and plasmid construction. *E. coli* strain Top10F' was used for expression experiments. *E. coli* strains were cultured in Luria–Bertani (LB) media at 37 °C with shaking or on LB agar plates. Phosphate-buffered saline was used for fluorescence measurements. Media were supplemented with kanamycin (50 $\mu\text{g mL}^{-1}$) and chloramphenicol (100 $\mu\text{g mL}^{-1}$) as appropriate. OHHL (Sigma) was used for the LuxI–LuxR system activation. All restriction enzymes are obtained from New England Biolabs.

Plasmid Construction. The plac-LuxRI and plux-GFPuv plasmids were obtained from the Arnold Group at Caltech (25). Plasmid plac-LuxR (Supplementary Figure 1, panel a) was constructed from the plac-LuxRI plasmid by PCR using the LuxR Δ I_{For} (5'-GGATCCATAAACACGCTGTAGTCAAGG-3') and LuxR Δ I_{Rev} (5'-TTTCTACAATAAGTCCCCTC-3') primers. These primers were designed to amplify the plac-LuxRI plasmid excluding a large portion of *luxI*. After amplification, the plasmid was recircularized by digestion with *Bam*HI followed by ligation with T4 DNA Ligase (New England Biolabs). Plasmid plux-GFPuv-LuxR (Supplementary Figure 1, panel b) was constructed by ligating *Pvu*II and *Not*I digested plux-GFPuv with *luxR* amplified from the plac-LuxR plasmid using primers LuxR(RBS)_{For} (5'-GAGCGATCGTTAAAGAGGAGAAAGG-3') and LuxR(RBS)_{Rev} (5'-ATAGCGGCCGACTTAATTTTTAAAGTATGG-3'). Plux-GFPuv-LuxR mutant plasmids containing the

mutant LuxR genes were constructed in an identical manner.

Library Construction and Screening. LuxR mutants were constructed by error-prone PCR using *Taq* polymerase (Eppendorf) with 7 mM MnCl₂ to increase mutation rates and increased dCTP and dTTP concentrations to counter *Taq* mutational bias as described (29). The primers LuxR(ePCR)_{For} (5'-GAAAGGTACCCATGAAAAACATAAAT-3') and LuxR(ePCR)_{Rev} (5'-CGGGGATCCCCTACTTAATT-3') were used to amplify the LuxR gene using plac-LuxR plasmid as the template. The library was constructed by ligating *Kpn*I- and *Bam*HI-digested plac-LuxR with the amplified mutants using T4 DNA ligase. Screening of the library was performed using a slightly modified protocol described in ref 25. Ligated mutant plasmids were transformed into Top10F' *E. coli* containing the plux-GFPuv plasmid and grown on LB plates containing 1 mM IPTG, 10 nM OHHL, and selection antibiotics at 37 °C for 18 h. The plates were then incubated at RT and visually checked for fluorescence every hour using a UV-transilluminator. Colonies becoming fluorescent before wild type were isolated and rescreened to ensure activity.

Fluorescent Quantification of LuxR Mutants and Positive Feedback Mechanisms.

The methods used to quantify the LuxR mutant proteins and the constructed PFLs were identical. The plac-LuxR mutant plasmids were transformed into Top10F' *E. coli* (Invitrogen) containing the plux-GFPuv signal plasmid for mutant screening, or the plux-GFPuv-LuxR mutant plasmids for establishment of mutant PFL2s. Mutant PFL1s were established by transforming Top10F' *E. coli* with plux-GFPuv-LuxR mutant plasmids. Transformed cells were grown on LB plates, and single colonies were used to grow overnight cultures. Log phase cultures (OD₅₉₅ \approx 0.3) were obtained by diluting overnight cultures 200-fold in fresh LB media containing 1 mM IPTG and growing at 37 °C for 3 h. After this growth period, cultures were induced with OHHL and incubation was continued for an additional 4 h. Cultures were then resuspended in phosphate-buffered saline and transferred to a clear bottom 96-well plate. A microtiter plate reader (Molecular Devices, SpectraMax M5) was used to measure the fluorescence of the GFPuv signal with excitation of 395 nm and emission of 509 nm. Fluorescence intensities were normalized to optical density at 505 nm. Each data point on response curves represents fluorescence data from two independently grown cultures.

Acknowledgments: We thank F. Arnold (Caltech) for providing genes and plasmids, and for her critical reading of the manuscript. This work was supported by a startup fund provided by the University of Massachusetts Amherst.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

1. Harris, S. L., and Levine, A. J. (2005) The p53 pathway: positive and negative feedback loops, *Oncogene* 24, 2899–2908.

2. Hoffmann, A., and Baltimore, D. (2006) Circuitry of nuclear factor kappa B signaling, *Immunol. Rev.* **210**, 171–186.
3. Becskei, A., Seraphin, B., and Serrano, L. (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion, *EMBO J.* **20**, 2528–2535.
4. Maeda, Y. T., and Sano, M. (2006) Regulatory dynamics of synthetic gene networks with positive feedback, *J. Mol. Biol.* **359**, 1107–1124.
5. Ferrell, J. E., and Xiong, W. (2001) Bistability in cell signaling: how to make continuous processes discontinuous, and reversible processes irreversible, *Chaos* **11**, 227–236.
6. Kramer, B. P., and Fussenegger, M. (2005) Hysteresis in a synthetic mammalian gene network, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9517–9522.
7. Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., Elston, T. C., and Collins, J. J. (2006) A bottom-up approach to gene regulation, *Nature* **439**, 856–860.
8. Isaacs, F. J., Hasty, J., Cantor, C. R., and Collins, J. J. (2003) Prediction and measurement of an autoregulatory genetic module, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7714–7719.
9. Emilusen, L., Gough, M., Bateman, A., Ahmed, A., Voellmy, R., Chester, J., Diaz, R. M., Harrington, K., and Vile, R. (2001) A transcriptional feedback loop for tissue-specific expression of highly cytotoxic genes which incorporates an immunostimulatory component, *Gene Ther.* **8**, 987–998.
10. Vilaboa, N., Fenna, M., Munson, J., Roberts, S. M., and Voellmy, R. (2005) Novel gene switches for targeted and timed expression of proteins of interest, *Mol. Ther.* **12**, 290–298.
11. Szulc, J., Wiznerowicz, M., Sauvain, M.-O., Trono, D., and Aebischer, P. (2006) A versatile tool for conditional gene expression and knockdown, *Nat. Methods* **3**, 109–116.
12. Fussenegger, M. (2001) The impact of mammalian gene regulation concepts on functional genomic research, metabolic engineering, and advanced gene therapies, *Biotechnol. Prog.* **17**, 1–51.
13. Arnold, F. H. (2001) Combinatorial and computational challenges for biocatalyst design, *Nature* **409**, 253–257.
14. Yokobayashi, Y., Weiss, R., and Arnold, F. H. (2002) Directed evolution of a genetic circuit, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16587–16591.
15. Thomas, M., and van Tilburg, A. (2000) Overexpression of foreign proteins using the *Vibrio fischeri* lux control system, *Methods Enzymol.* **305**, 315–329.
16. Engebrecht, J., Nealson, K., and Silverman, M. (1983) Bacterial bioluminescence - isolation and genetic-analysis of functions from *Vibrio fischeri*, *Cell* **32**, 773–781.
17. Fuqua, C., and Greenberg, E. P. (2002) Listening in on bacteria: acyl-homoserine lactone signalling, *Nat. Rev. Mol. Cell Biol.* **3**, 685–695.
18. You, L. C., Cox, R. S., Weiss, R., and Arnold, F. H. (2004) Programmed population control by cell-cell communication and regulated killing, *Nature* **428**, 868–871.
19. Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., and Collins, J. J. (2004) Programmable cells: Interfacing natural and engineered gene networks, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8414–8419.
20. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R. (2005) A synthetic multicellular system for programmed pattern formation, *Nature* **434**, 1130–1133.
21. Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Oppenheimer, N. J. (1981) Structural identification of autoinducer of photobacterium-fischeri luciferase, *Biochemistry* **20**, 2444–2449.
22. Kaern, M., Elston, T. C., Blake, W. J., and Collins, J. J. (2005) Stochasticity in gene expression. from theories to phenotypes, *Nat. Rev. Genet.* **6**, 451–464.
23. Smits, W. K., Kuipers, O. P., and Veening, J. W. (2006) Phenotypic variation in bacteria: the role of feedback regulation, *Nat. Rev. Microbiol.* **4**, 259–271.
24. Ferrell, J. E. (1999) Building a cellular switch: more lessons from a good egg, *Bioessays* **21**, 866–870.
25. Collins, C. H., Arnold, F. H., and Leadbetter, J. R. (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones, *Mol. Microbiol.* **55**, 712–723.
26. Karig, D., and Weiss, R. (2005) Signal-amplifying genetic enables in vivo observation circuit of weak promoter activation in the Rhl quorum sensing system, *Biotechnol. Bioeng.* **89**, 709–718.
27. Weber, W., and Fussenegger, M. (2006) Pharmacologic transgene control systems for gene therapy, *J. Gene Med.* **8**, 535–556.
28. Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR, *Nat. Biotechnol.* **24**, 708–712.
29. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR, *PCR Methods Appl.* **3**, S136–S140.
30. Trott, A. E., and Stevens, A. M. (2001) Amino acid residues in LuxR critical for its mechanism of transcriptional activation during quorum sensing in *Vibrio fischeri*, *J. Bacteriol.* **183**, 387–392.
31. Koch, B., Lijefors, T., Persson, T., Nielsen, J., Kjelleberg, S., and Givskov, M. (2005) The LuxR receptor: the sites of interaction with quorum-sensing signals and inhibitors, *Microbiology-SGM* **151**, 3589–3602.
32. Shadel, G. S., Young, R., and Baldwin, T. O. (1990) Use of regulated cell-lysis in a lethal genetic selection in *Escherichia coli* - identification of the autoinducer-binding region of the LuxR protein from *Vibrio Fischeri* Atcc-7744, *J. Bacteriol.* **172**, 3980–3987.
33. Dunlap, P. V., and Greenberg, E. P. (1988) Control of *Vibrio-fischeri* Lux gene-transcription by a cyclic-amp receptor protein LuxR protein regulatory circuit, *J. Bacteriol.* **170**, 4040–4046.
34. Arnold, F. H. (1998) Design by directed evolution, *Acc. Chem. Res.* **31**, 125–131.
35. Walz, D., and Caplan, S. R. (1995) Chemical oscillations arise solely from kinetic nonlinearity and hence can occur near equilibrium, *Biophys. J.* **69**, 1698–1707.
36. Toniatti, C., Bujard, H., Cortese, R., and Ciliberto, G. (2004) Gene therapy progress and prospects: transcription regulatory systems, *Gene Ther.* **11**, 649–657.
37. Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12678–12683.
38. Glick, B. R., and Pasternak, J. J. (2006) *Molecular biotechnology: principles and applications of recombinant DNA*, ASM Press, Washington, DC.
39. Bayer, T. S., and Smolke, C. D. (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression, *Nat. Biotechnol.* **23**, 337–343.
40. Chen, M. T., and Weiss, R. (2005) Artificial cell-cell communication in yeast *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis thaliana*, *Nat. Biotechnol.* **23**, 1551–1555.
41. Bulter, T., Lee, S. G., Woitl, W. W. C., Fung, E., Connor, M. R., and Liao, J. C. (2004) Design of artificial cell-cell communication using gene and metabolic networks, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2299–2304.
42. Weber, W., Malphettes, L., de Jesus, M., Schoenmakers, R., El-Baba, M. D., Spielmann, M., Keller, B., Weber, C. C., van de Wetering, P., Aubel, D., Wurm, F. M., and Fussenegger, M. (2005) Engineered streptomyces quorum-sensing components enable inducible siRNA-mediated translation control in mammalian cells and adjustable transcription control in mice, *J. Gene Med.* **7**, 518–525.
43. Weber, W., Schoenmakers, R., Spielmann, M., El-Baba, M. D., Folcher, M., Keller, B., Weber, C. C., Link, N., van de Wetering, P., Heinzen, C., Jolivet, B., Sequin, U., Aubel, D., Thompson, C. J., and Fussenegger, M. (2003) Streptomyces-derived quorum-sensing systems engineered for adjustable transgene expression in mammalian cells and mice, *Nucl. Acids Res.* **31**, e71.
44. Weber, W., and Fussenegger, M. (2004) Approaches for trigger-inducible viral transgene regulation in geno-based tissue engineering, *Curr. Opin. Biotechnol.* **15**, 383–391.

Phosphopeptide Modification and Enrichment by Oxidation–Reduction Condensation

Mangalika Warthaka, Paulina Karwowska-Desaulniers, and Mary Kay H. Pflum*

Department of Chemistry, Wayne State University, 5101 Cass Avenue, Detroit, Michigan 48202

Phosphorylation is an important post-translational modification that influences protein function and cell signaling (1). Because of its role in regulating protein activity, characterizing and monitoring phosphorylation events is a fruitful area of proteomics research. Although traditional methods for monitoring phosphorylated proteins include *in vivo* ^{32}P -labeling and gel electrophoresis, mass spectrometric (MS) analysis has emerged as a useful tool for phosphoproteomics applications (2). To increase the probability of detecting phosphopeptides in complex peptide mixtures, MS analysis is typically preceded by a phosphopeptide enrichment step, such as immobilized metal affinity chromatography (IMAC) (2, 3). Unfortunately, the reliance of IMAC methods on noncovalent interactions for phosphopeptide capture results in contamination with unphosphorylated peptides (4).

Towards more effective enrichment strategies, phosphopeptides have been attached to a solid phase *via* covalent bonds. For example, multiple strategies rely on phosphate β -elimination followed by thiol conjugate addition for covalent solid-phase capture (4–9). After the solid phase is washed, phosphopeptides are recovered by cleavage from the resin. A limitation of the strategy is that β -elimination occurs with unphosphorylated peptides, including serine, threonine, and glycosylated amino acids (10, 11), which contaminate the enriched mixture and complicate the MS analysis. In addition, only phosphoserine- and phosphothreonine-containing peptides are susceptible to β -elimination conditions,

which excludes analysis of phosphotyrosine-containing peptides.

As an alternative to β -elimination strategies, two direct phosphate modification reactions have been previously explored for phosphopeptide capture. In the first approach, a phosphotyrosine-containing peptide was alkylated using an α -diazo carbonyl functionalized resin for capture and release (12). In the second approach, phosphopeptides from cell lysates were enriched using carbodiimide condensation (13, 14). Like the β -elimination approach, a limitation of the two covalent chemistries is that their compatibility with phosphoserine, phosphothreonine, and phosphotyrosine peptides has not been established. Specifically, the alkylation chemistry has only been validated with phosphotyrosine-containing peptides, while the carbodiimide condensation strategy required initial enrichment of phosphotyrosine-containing peptides using antibodies. Because covalent phosphopeptide enrichment has the potential to significantly aid phosphoproteomics research, new strategies that enrich all phosphopeptides and avoid contamination by unphosphorylated peptides are needed.

In search of a new chemical strategy to enhance the chemical selectivity of phosphopeptide capture, we considered the oxidation–reduction condensation, where a phosphine and disulfide act to activate a phosphate for subsequent coupling with an alcohol or amine (Scheme 1). Oxidation–reduction condensation was originally reported for alkylation of phosphates to

ABSTRACT Many cellular processes are regulated by the reversible phosphorylation of proteins. Despite the importance of monitoring protein phosphorylation, available methods to modify and enrich phosphopeptides from complex mixtures for subsequent mass spectrometric analysis are challenging. Here the oxidation–reduction condensation was shown for the first time to directly modify the phosphate of phosphopeptides and phosphoproteins. By coupling with a solid-phase resin, the oxidation–reduction condensation was validated for capture and recovery of phosphoserine-, phosphothreonine-, and phosphotyrosine-containing peptides from a peptide mixture. In addition, full-length phosphoproteins or phosphopeptides from a protein digestion were captured and recovered using the oxidation–reduction condensation, demonstrating its compatibility with protein mixtures. The strategy modifies all phosphopeptides, maintains high chemical selectivity, requires only two steps, and relies on commercially available reagents, suggesting that the oxidation–reduction condensation has the potential to enhance phosphopeptide enrichment methods and encourage development of efficient biochemical and proteomics tools targeting phosphorylation.

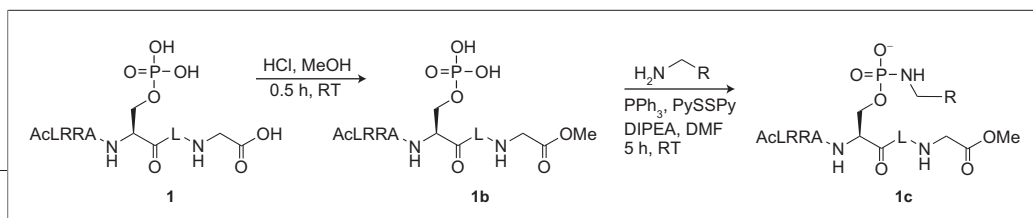
*Corresponding author,
pflum@chem.wayne.edu.

Received for review August 16, 2006
and accepted October 22, 2006.

Published online December 1, 2006

10.1021/cb6003564 CCC: \$33.50

© 2006 by American Chemical Society



Scheme 1. The two-step protection/oxidation–reduction condensation of phosphoserine-containing peptide **1** with benzyl amine.

create mixed phosphate diesters and phosphoramidates (15). Subsequently, oxidation–reduction condensation of nucleic acids was exploited to create various phosphate-modified DNA and RNA analogues (16–18). In addition, the oxidation–reduction condensation was employed for amino acid coupling in peptide synthesis (19). Despite the reported efficiency for phosphate modification and compatibility with peptides, the oxidation–reduction condensation has not been exploited with phosphopeptides.

We initially tested the oxidation–reduction condensation for phosphopeptide modification using synthetic peptides to allow a rigorous assessment of chemical selectivity. Phosphokemptide, (peptide **1**, Table 1), a natural enzymatic product of cAMP-dependent protein kinase (20), was generated. In addition, peptides where the phosphoserine of phosphokemptide was replaced by serine, cysteine, or aspartic acid (peptides **2**, **3**, and **4**, Table 1) were synthe-

sized to assess chemical selectivity.

Because the carboxyl terminus and carboxylic acid side chain of each peptide would react under conditions of oxidation–reduction condensation, the acids were selectively protected in the presence of phosphate esters, as previously reported (Scheme 1) (12, 14). MALDI-TOFMS data of the peptide products after protection were consistent with conversion of the carboxyl termini of peptides **1–3** to methyl esters (Table 1). In addition, aspartic acid-containing peptide **4** displayed a mass consistent with protection of the carboxyl terminus and carboxylic acid side chain. The protected peptides were used in the oxidation–reduction condensation reactions without purification.

The condensation of peptides **1–4** was initially explored with benzyl amine (Scheme 1, where R = Ph) because earlier work demonstrated an 80% conversion of phosphate to phosphoramidates (15). The oxidation–reduction condensation was carried out with

the protected phosphoserine-containing peptide **1b** (Scheme 1), and MALDI-TOFMS data (Table 1) indicated conversion to a single product consistent with benzylamine condensation

was observed (Table 1). The data indicate that oxidation–reduction condensation with benzylamine is dependent on the presence of a phosphate group. To assess the efficiency of the two-step phosphate modification reaction, the product mixture after oxidation–reduction condensation of peptide **1b** with benzyl amine was separated using reverse-phase HPLC (Supplementary Figure 1). While starting protected peptide **1b** was present in the reaction mixture, the predominant peptide peak corresponded to the condensation product **1c**. Quantification of the peptide peaks indicated an average of 89% conversion to condensation product.

To test the reactivity of the other phosphorylated peptides under condensation conditions, we synthesized peptides **5**, **6**, **7**, and **8** by replacing the phosphoserine of peptide **1** with phosphothreonine, threonine, phosphotyrosine, and tyrosine, respectively (Table 1). Oxidation–reduction condensation was carried out with the protected peptides, and MALDI-TOFMS data of the reaction products indicated that only the phosphopeptides were modified under the reaction conditions; unphosphorylated peptides were unreactive. HPLC analysis indicated an average of 90% conversion of the phosphothreonine peptide **5** to condensation product (Supplementary Figure 2), demonstrating that the condensation reaction equally modifies phosphoserine- and phosphothreonine-containing peptides. In the case of phosphotyrosine peptide **7**, (1c). In contrast, HPLC analysis revealed an average of 58% conversion to condensation product (Supplementary Figure 3). In total, the data indicate that the oxidation–reduction condensation reaction modifies all phosphate-containing peptides, with phosphoserine and phosphothreonine modified to a greater extent. In addition, phosphopeptides but not unphosphorylated peptides are susceptible to the reaction conditions.

TABLE 1. MALDI-TOFMS data of peptides **1–8** after carboxylic acid protection and oxidation–reduction condensation with benzylamine

Peptides	Protected peptides		Condensation products	
	Calcd ^a	Obsd ^b	Calcd ^a	Obsd ^b
1 (AcLRRApSLG)	908.46	908.47	997.53	997.58
2 (AcLRRASLG)	828.50	828.47	917.56	828.47 ^d
3 (AcLRRACLG)	844.47	844.51	917.56	844.51 ^d
4 (AcARRADLG)	828.46 ^c	828.44	903.56	828.44 ^d
5 (AcLRRApTLG)	922.45	922.55	1011.54	1011.61
6 (AcLRRATLG)	842.52	842.52	931.58	842.52 ^d
7 (AcLRRApYLG)	984.49	984.46	1073.54	1073.68
8 (AcARRAYLG)	862.48	862.71	951.57	862.71 ^d

^aCalculated values are masses [M + H]⁺ of protected peptides and expected condensation products. ^bObserved MALDI-TOFMS data were collected in the positive ion mode. ^cCalculated mass of peptide protected at the C-terminus and aspartic acid side chain. ^dMass of protected peptide, indicating no reaction under the condensation conditions.

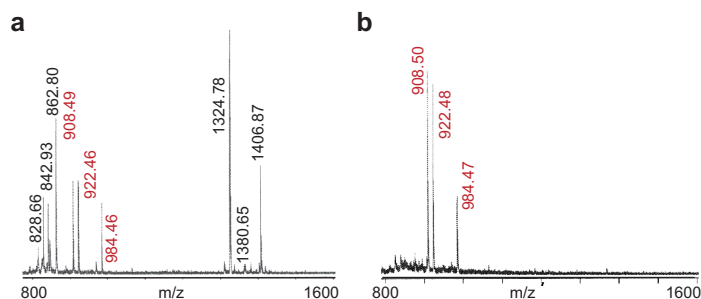


Figure 1. Enrichment of phosphopeptides from a peptide mixture. **a, b** MALDI-TOFMS analysis of phosphopeptides **1, 5, and 7** (shown in red) captured on a solid phase in the presence of six unphosphorylated peptides: peptides **2, 6, and 8**, angiotensin (DRVYIHPFHL, 1324.78 *m/z*), glucagon (AQDFVQWLMNT, 1380.65 *m/z*), and Abl substrate peptide (AcEAIYAAPFAKKK, 1406.87 *m/z*). Shown are reaction mixtures before solid-phase capture (**a**) and after cleavage from the resin (**b**).

The next step toward development of a selective phosphopeptide enrichment strategy was testing the oxidation–reduction condensation for solid-phase phosphopeptide capture. A mixture of synthetic peptides was employed initially to assess the compatibility and selectivity of the condensation in solid-phase capture and release. Phosphopeptides **1, 5, and 7** were subjected to the two-step reaction conditions in the presence of glycine-conjugated Wang resin (Scheme 1, where R = glycine-Wang resin) to test chemical compatibility. To simultaneously assess chemical selectivity, six competing unphosphorylated peptides were also included in the peptide mixture: peptides **2, 6, and 8**, angiotensin, glucagon, and an Abl substrate peptide (Table 1 and Figure 1). These unphosphorylated peptides were selected because they contain a random pool of naturally occurring amino acids. After solid-phase capture and resin washing, the bound peptides were released under acidic conditions. Whereas all nine peptides were observed as a mixture in solution before solid-phase capture (Figure 1, panel a), only the three phosphopeptides were recovered after release from the resin (Figure 1, panel b). Significantly, no trace of unphosphorylated peptide was detected after recovery. The data indicate that the two-step protection/condensation reaction strategy successfully captures all three phosphate-containing peptides in the presence of possible contaminating unphosphorylated peptides. In addition, the fact that the six unphosphorylated pep-

tides contained common functional groups present in all proteins—carboxylic acids, alcohols, thioethers, amines, amides, guanidines, imidazoles, and indoles—indicates that the reaction strategy is selective for modification and capture of phosphate-containing peptides.

To assess the limits of detection with the solid-phase capture and recovery by oxidation–reduction condensation, decreasing quantities of phosphoserine peptide **1** were subjected to bead binding and cleavage. The phosphopeptide was recovered and identified by MALDI-TOFMS with starting quantities as little as 100 pmol (data not shown). Because detection limits as low as 100 pmol have proven reliable for capture of phosphopeptides in cell lysates (14), the data indicate that the oxidation–reduction condensation is compatible with the concentrations of phosphopeptides in biological samples.

With the high selectivity and limits of detection of the oxidation–reduction condensation established using synthetic peptides, we sought to validate the reaction for solid-phase capture of a phosphopeptide from a full-length protein digestion. The β -casein protein, which contains a phosphoserine residue, was used previously to validate phosphopeptide enrichment for proteomics applications (13, 14). Similarly, the two-step oxidation–reduction condensation was performed with a trypsin digestion of the β -casein protein. Full-length β -casein was initially reduced and alkylated to improve solubility in organic solvents

before digestion with trypsin. Prior to enrichment, the digested mixture showed multiple peptide fragments, with only a small peak corresponding to the phosphopeptide (Figure 2, panel a). The peptide fragments were subjected to the two-step capture with the glycine-conjugated Wang resin. Following cleavage from the resin, the only peptide recovered corresponded to the phosphorylated peptide (Figure 2, panel b). The data with β -casein indicate that the oxidation–reduction condensation is compatible with peptide mixtures derived from a phosphoprotein for application to proteomics research.

Previous phosphopeptide enrichment using the carbodiimide condensation demonstrated 20–35% recovery of phosphopeptides from a β -casein digestion (13, 14). To directly compare the carbodiimide and oxidation–reduction condensation chemistries, we used quantitative MS analysis to assess the efficiency of the oxidation–reduction condensation for phosphopeptide capture on a solid phase, as previously reported (14). Quantitative MS indicated that an average of 37% of the phosphopeptide was recovered from the solid phase (Supplementary Figure 4), indicating that the phosphopeptide recovery with the oxidation–reduction condensation is comparable to the previous carbodiimide strategy. Coupled with the proven high selectivity and phosphopeptide compatibility, the experiments with β -casein establish that the oxidation–reduction condensation provides a needed alternative to previous chemical approaches.

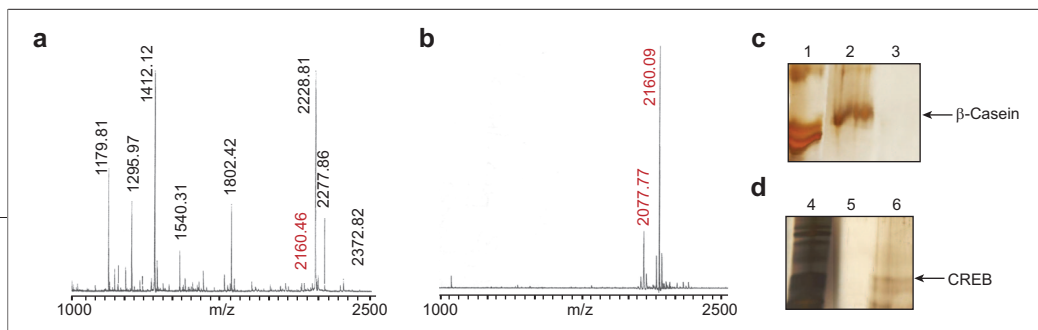


Figure 2. Enrichment of phosphopeptides from a protein digestion or full-length phosphoproteins. **a, b)** MALDI-TOFMS of trypsin digested β -casein before **(a)** and after **(b)** solid-phase phosphopeptide capture and release. The protected phosphopeptide (FQSEEQQTEDELQDK, 2160.09 m/z) is minimally deprotected (2077.77 m/z) under the cleavage conditions (shown in red). **c)** Full-length β -casein (lane 1) was recovered either untreated (lane 2) or after phosphatase treatment (lane 3). **d)** Full-length CREB partially purified after over-expression in bacteria (lane 4) was captured and recovered either untreated (lane 5) or after PKA phosphorylation (lane 6). The faster-migrating band beneath CREB corresponds to autophosphorylated PKA.

To further assess the chemical versatility of the oxidation–reduction condensation, we tested the capture and recovery of full-length phosphoproteins. Full-length β -casein (Figure 2, panel c, lane 1) was incubated with glycine-conjugated Wang resin under conditions of the oxidation–reduction condensation. Following cleavage from the solid phase, the phosphoprotein was recovered and visualized by silver staining (Figure 2, panel c, lane 2). Importantly, unphosphorylated β -casein generated by incubation with calf intestinal phosphatase was not recovered (Figure 2, panel c, lane 3), indicating that the capture is phosphate-dependent. Because no direct phosphate modification chemistry has been previously shown to enrich full-length proteins, the oxidation–reduction condensation provides a general method for solid-phase capture of phosphoproteins in addition to phosphopeptides.

Finally, to demonstrate general compatibility with multiple full-length proteins, the oxidation–reduction condensation was tested with the cyclic adenosine monophosphate response element binding (CREB) protein, which is phosphorylated by protein kinase A (PKA) at Ser133 (21). Full-length CREB was over-expressed in bacteria as a hexa-his tagged fusion protein and partially purified using Ni-nitrilotriacetic (NTA) resin (Figure 2, panel d, lane 4). The CREB protein was untreated or phosphorylated with PKA before capture and recovery *via* the protection/oxidation–reduction condensation. CREB was enriched from the protein mixture in a phosphorylation-dependent manner (Figure 2, panel d, lane 6); unphosphorylated CREB was absent after solid-phase

capture (Figure 2, panel d, lane 5). In addition, a protein band migrating at the same molecular weight as PKA was also observed (Figure 2, panel d, lane 6), indicating the capture of autophosphorylated PKA (22, 23). The experiments with full-length CREB and PKA validate the compatibility of phosphoprotein enrichment with multiple proteins. The fact that CREB maintains a single phosphorylated serine (21) indicates that the solid-phase capture method effectively enriches proteins with only one phosphorylated amino acid. Importantly, phosphorylated CREB and PKA were enriched from a mixture of proteins derived from cell lysates, suggesting that the chemistry is appropriate for phosphoproteomics applications.

We hypothesized that use of unexplored chemistry for phosphopeptide modification would have the potential to enhance the compatibility and selectivity of enrichment strategies. Here we demonstrate for the first time that an oxidation–reduction condensation reaction selectively converts the phosphate of phosphopeptides and phosphoproteins into phosphoramidates, allowing solid-phase enrichment. The highlights of the strategy include high chemical selectivity for phosphorylated peptides and proteins; compatibility with phosphoserine, phosphothreonine, or phosphotyrosine; and recovery of full-length phosphoproteins, distinguishing it from previous phosphate modification strategies. In total, the data presented here demonstrate that unexplored phosphate modification chemistry such as the oxidation–reduction condensation has the potential to enhance phosphopeptide and phosphoprotein enrichment strategies.

The spectrum of phosphopeptides identified from a cellular mixture is dependent on the type of enrichment chemistry employed (24). For example, different protocols for IMAC enrichment

identified overlapping but nonidentical phosphopeptides from mouse synaptic proteins (24). In addition, β -elimination and IMAC methods yielded a nonredundant group of phosphopeptides from mouse brain (4). These results suggest that multiple phosphopeptide enrichment approaches are needed to fully characterize the phosphoproteome. The oxidation–reduction condensation described here provides a general phosphopeptide enrichment approach with proven compatibility with phosphoserine, phosphothreonine, and phosphotyrosine peptides. In addition, its high chemical selectivity will avoid contamination by unphosphorylated peptides, including glycosylated peptides (10, 11), which will simplify MS analysis. Because covalent phosphopeptide enrichment methods have the potential to significantly aid phosphoprotein analysis, new strategies like the oxidation–reduction condensation will encourage development of efficient biochemical and proteomics tools targeting protein phosphorylation.

METHODS

Phosphopeptide Modification Reactions. Peptides 1–8 were synthesized using standard solid-phase fmoc chemistry (see Supporting Information). HPLC-purified peptide (0.04 μmol) was resuspended in 0.2 mL of anhydrous MeOH, and acetyl chloride (2 N, 0.4 mmol) was added to generate anhydrous HCl. After stirring for 30 min and solvent evaporation, the protected peptide precipitate was resuspended in 0.3 mL of anhydrous DMF under argon gas, and 2,2'-dithiodipyridine (0.4 μmol , 10 equiv), *N,N'*-diisopropylethylamine (DIPEA) (0.04 μmol , 1 eq), triphenylphosphine (0.4 μmol , 10 equiv), and benzyl amine (0.6 μmol , 15 equiv) were added with vigorous shaking. After 5 h, the reaction mixture was analyzed by MS and/or HPLC (see Supporting Information).

Capture of Phosphopeptides Using Glycine-Preloaded Wang Resin. The carboxylic acid groups of the peptide (0.04 μmol) or peptide mixture (0.32 μmol ; peptides **1**, **2**, **5**, **6**, **7**, and **8**, angiotensin, glucagons, and Abl kinase substrate peptide in a 1:1:1:1:1:3:3:2 molar ratio) were protected as described, and the peptide precipitate was resuspended in 0.1 mL of dry DMF. Fmoc-protected glycine-preloaded Wang resin (20 mg) that was presoaked in dry DMF (0.2 mL) was added under argon. The oxidation–reduction condensation reaction was initiated by addition of the reagents described above, with exception of benzyl amine. After shaking for 14 h, the resin was collected and washed with DMF (10 \times), dichloromethane (4 \times), 50% dichloromethane in ether (3 \times), and ether (3 \times). The washed beads were incubated with a mixture of 95% TFA, 2.5% TIS, and 2.5% water (0.4 mL) for 2 h to cleave the peptide. After solvent evaporation, the remaining peptide precipitate was resuspended in water before analysis by MALDI-TOFMS (see Supporting Information).

Trypsin Digestion of β -Casein. Bovine β -casein from milk (0.03 μmol , Sigma) was resuspended in trypsin buffer (50 mM ammonium carbonate, pH 8), reduced with 5 mM DTT at 60 $^{\circ}\text{C}$ for 30 min, and alkylated with 15 mM iodoacetamide at 25 $^{\circ}\text{C}$ for 30 min in the dark. Sequencing-grade trypsin (Promega) was added in 1:100 ratio (trypsin/protein, w/w) at 37 $^{\circ}\text{C}$ overnight. After lyophilization, the peptide fragments were protected in a solution of 100 μL of anhydrous MeOH and 500 μL acetyl chloride with vigorous shaking for 2 h at 12 $^{\circ}\text{C}$. Samples were lyophilized, and the oxidation–reduction condensation was initiated by addition of 10 mg of Fmoc-protected glycine-preloaded Wang bead resin (presoaked and deprotected as described previously), 2,2'-dithiopyridine (302 mM), triphenylphosphine (130 mM), and DIPEA (38 mM) in 0.3 mL of dry DMF for 15 h under argon, followed by washing of the beads and cleavage as described. MALDI-TOF analysis is described in the Supporting Information.

Capture of Full-Length Proteins. Full-length CREB was overexpressed in BL21 *E. coli* cells transformed with the T7-7 CREB plasmid, a kind gift of the Montminy lab (25), before purification with Ni-NTA (see Supporting Information). To phosphorylate the CREB-containing mixture, 0.014 μmol of total protein was incubated with 1500 units of PKA and 200 μM ATP in 1 \times kinase reaction buffer. To produce unphosphorylated β -casein protein, calf intestinal phosphatase (1 μL) was incubated with 0.03 μmol of protein for 1 h at 30 $^{\circ}\text{C}$. The CREB mixture (0.014 μmol) or β -casein protein (0.03 μmol) was used directly in solid-phase capture and release as described. After evaporation of the solvent, the proteins were separated using 12% SDS-PAGE before visualization by silver staining.

Acknowledgment Funding was provided by Research Corp. (Grant RI1006) and Wayne State University. We thank M. Montminy for the T7-7 CREB plasmid and E. Aubie, A. Bieliauskas, J. Flammer, K. Green, and S. Suwal for helpful comments.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science* **298**, 1912–1926.
- García, B. A., Shabanowitz, J., and Hunt, D. F. (2005) Analysis of protein phosphorylation by mass spectrometry, *Methods* **35**, 256–264.
- Corthals, G. L., Aebersold, R., and Goodlett, D. R. (2005) Identification of phosphorylation sites using microimmobilized metal affinity chromatography, *Mass spectrometry: Modified proteins and glycoconjugates*, Methods in Enzymology, Vol. 405, pp 66–81, Academic Press, Boston.
- Tseng, H., Ovaa, H., Wei, N. J. C., Ploegh, H., and Tsai, L. (2005) Phosphoproteomic analysis with a solid-phase capture-release-tag approach, *Chem. Biol.* **12**, 769–777.
- Oda, Y., Nagasu, T. T., and Chait, B. T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome, *Nat. Biotechnol.* **19**, 379–382.
- Knight, Z. A., Schilling, B., Row, R. H., Kensi, D. M., Gibson, B. W., and Shokat, K. M. (2003) Phosphospecific proteolysis for mapping sites of protein phosphorylation, *Nat. Biotechnol.* **21**, 1047–1054.
- Goshe, M. B., Veenstra, T. D., Panisko, E. A., Conrads, T. P., Angell, N. H., and Smith, R. D. (2002) Phosphoprotein isotope-coded affinity tags: Application to the enrichment and identification of low-abundance phosphoproteins, *Anal. Chem.* **74**, 607–616.
- Adamczyk, M., Gebler, J. C., and Wu, J. (2001) Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry, *Rapid Commun. Mass Spectrom.* **15**, 1481–1488.
- Rusnak, F., Zhou, J., and Hathaway, G. M. (2002) Identification of phosphorylated and glycosylated sites in peptides by chemically targeted proteolysis, *J. Biomol. Tech.* **13**, 228–237.
- Li, W., Backlund, P. S., Boykins, R. A., Wang, G., and Chen, H. (2003) Susceptibility of the hydroxyl groups in serine and threonine to β -elimination/Michael addition under commonly used moderately high-temperature conditions, *Anal. Biochem.* **323**, 94–102.
- Rusnak, F., Zhou, J., and Hathaway, G. M. (2004) Reaction of phosphorylated and O-glycosylated peptides by chemically targeted identification at ambient temperature, *J. Biomol. Tech.* **15**, 296–304.
- Lansdell, T. A., and Tepe, J. J. (2004) Isolation of phosphopeptides using solid phase enrichment, *Tetrahedron Lett.* **45**, 91–93.
- Zhou, H., Watts, J., and Aebersold, R. (2001) A systematic approach to the analysis of protein phosphorylation, *Nat. Biotechnol.* **19**, 375–378.
- Tao, W. A., Wollscheid, B., O'Brien, R., Eng, J. K., Li, X.-j., Bodenmiller, B., Watts, J. D., Hood, L., and Aebersold, R. (2005) Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry, *Nat. Methods* **2**, 591–598.
- Mukaiyama, T., and Hashimoto, M. (1971) Phosphorylation of alcohols and phosphates by oxidation–reduction condensation, *Bull. Chem. Soc. Jpn.* **44**, 196–199.
- Mukaiyama, T., and Hashimoto, M. (1972) Synthesis of oligothymidylates and nucleoside cyclic phosphates by oxidation–reduction condensation, *J. Am. Chem. Soc.* **94**, 8528–8532.
- Stepinski, J., Waddell, C., Stolarski, R., Darzynkiewicz, E., and Rhoads, R. E. (2001) Synthesis and properties of mRNAs containing the novel “anti-reverse” cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'-deoxy)GpppG, *RNA* **7**, 1486–1495.
- Prabakar, K. J., and Ferris, J. P. (1997) Adenine derivatives as phosphate-activating groups for the regioselective formation of 3',5'-linked oligoadenylates on montmorillonite: Possible phosphate-activating groups for the prebiotic synthesis of RNA, *J. Am. Chem. Soc.* **119**, 4330–4337.
- Mukaiyama, T., Matsueda, R., and Suzuki, M. (1970) Peptide synthesis via the oxidation–reduction condensation by the use of 2,2'-dipyridyldisulfide as an oxidant, *Tetrahedron Lett.* **11**, 1901–1904.
- Meyer, H. E., Hoffmann-Posorske, E., Korte, H., and Heilmeyer, L. M. (1986) Sequence analysis of phosphoserine-containing peptides: modification for picomolar sensitivity, *FEBS Lett.* **204**, 61–66.
- Gonzalez, G. A., and Montminy, M. R. (1989) Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133, *Cell* **59**, 675–680.
- Yonemoto, W., Garrod, S. M., Bell, S. M., and Taylor, S. S. (1993) Identification of phosphorylation sites in the recombinant catalytic subunit of cAMP-dependent protein kinase, *J. Biol. Chem.* **268**, 18626–18632.
- Steinberg, R. A., Cauthron, R. D., Symcox, M. M., and Shuntoh, H. (1993) Autoactivation of catalytic (C alpha) subunit of cyclic AMP-dependent protein kinase by phosphorylation of threonine 197, *Mol. Cell. Biol.* **13**, 2332–2341.
- Collins, M. O., Yu, L., Coba, M. P., Husi, H., Campuzano, I., Blackstock, W. P., Choudhary, J. S., and Grant, S. G. N. (2005) Proteomic analysis of in vivo phosphorylated synaptic proteins, *J. Biol. Chem.* **280**, 5972–5982.
- Parker, D., Ferreri, K., Nakajima, T., LaMorte, V. J., Evans, R., Koerber, S. C., Hoeger, C., and Montminy, M. R. (1996) Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism, *Mol. Cell. Biol.* **16**, 694–703.

HIV-1 Reverse Transcriptase Structure with RNase H Inhibitor Dihydroxy Benzoyl Naphthyl Hydrazone Bound at a Novel Site

Daniel M. Himmel[†], Stefan G. Sarafianos^{†,**}, Sanjeewa Dharmasena[‡], Mohammed M. Hossain[‡], Kessler McCoy-Simandle[‡], Tatiana Ilina[‡], Arthur D. Clark, Jr.[†], Jennifer L. Knight[§], John G. Julias[¶], Patrick K. Clark[¶], Karsten Krogh-Jespersen[§], Ronald M. Levy[§], Stephen H. Hughes^{||}, Michael A. Parniak[‡], and Eddy Arnold^{†,*}

[†]Center for Advanced Biotechnology and Medicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854-5627, [‡]Department of Medicine, Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261-0001, [§]Department of Chemistry and Chemical Biology and BIOMAPS Institute for Quantitative Biology, Rutgers University, Piscataway, New Jersey 08854-8066, [¶]Basic Research Program, SAIC-Frederick, Inc., Frederick, Maryland 21702-1201, ^{||}HIV Drug Resistance Program, NCI-Frederick, Building 539, Frederick, Maryland 21702-1201, ^{**}Current address, Christopher Bond Life Sciences Center, Department of Molecular Microbiology and Immunology, University of Missouri School of Medicine, Columbia, Missouri 65211-7310

Ⓜ This paper contains enhanced objects.

ABSTRACT The rapid emergence of drug-resistant variants of human immunodeficiency virus, type 1 (HIV-1), has limited the efficacy of anti-acquired immune deficiency syndrome (AIDS) treatments, and new lead compounds that target novel binding sites are needed. We have determined the 3.15 Å resolution crystal structure of HIV-1 reverse transcriptase (RT) complexed with dihydroxy benzoyl naphthyl hydrazone (DHBNH), an HIV-1 RT RNase H (RNH) inhibitor (RNHI). DHBNH is effective against a variety of drug-resistant HIV-1 RT mutants. While DHBNH has little effect on most aspects of RT-catalyzed DNA synthesis, at relatively high concentrations it does inhibit the initiation of RNA-primed DNA synthesis. Although primarily an RNHI, DHBNH binds >50 Å away from the RNH active site, at a novel site near both the polymerase active site and the non-nucleoside RT inhibitor (NNRTI) binding pocket. When DHBNH binds, both Tyr181 and Tyr188 remain in the conformations seen in unliganded HIV-1 RT. DHBNH interacts with conserved residues (Asp186, Trp229) and has substantial interactions with the backbones of several less well-conserved residues. On the basis of this structure, we designed substituted DHBNH derivatives that interact with the NNRTI-binding pocket. These compounds inhibit both the polymerase and RNH activities of RT.

*Corresponding author,
arnold@cabm.rutgers.edu.

Received for review July 17, 2006
and accepted November 21, 2006.

Published online December 15, 2006

10.1021/cb600303y CCC: \$33.50

© 2006 by American Chemical Society

Human immunodeficiency virus, type 1 (HIV-1), reverse transcriptase (RT) is essential for HIV replication. RT converts the single-stranded viral genomic RNA into a linear double-stranded DNA that can be integrated into the host chromosomes (reviewed in ref 1). The enzyme has two activities, (i) a DNA polymerase that can use either RNA or DNA as a template and (ii) an RNase H (RNH) that selectively degrades the RNA strand of an RNA–DNA heteroduplex. The RNH activity of RT is required for virus replication; cellular RNH cannot substitute for the retroviral enzyme (2). The RNH activity degrades the genomic RNA during first-strand (“minus-strand”) DNA synthesis, which allows the newly synthesized DNA to be used as a template for second-strand (“plus-strand”) DNA synthesis.

HIV-1 RT is a heterodimer consisting of 66 kDa (p66) and 51 kDa (p51) subunits. The two polypeptide chains have 440 N-terminal amino acid residues in common. These comprise four polymerase subdomains: the thumb, palm, fingers, and connection (3, 4). The C-terminus of p66 contains an additional 120 amino acid residues that form the bulk of the RNH domain. Despite having identical N-terminal sequences, the arrangement of the subdomains in the two subunits differs dramatically. The p66 subunit contains a large cleft formed by the fingers, palm, and thumb subdomains that can accommodate double-stranded nucleic acid template–primers (3–6). Although the p51 subunit contains the same four subdomains, it does not form a nucleic acid binding cleft.

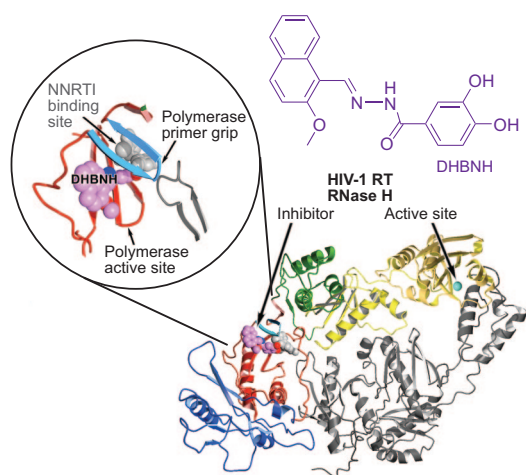


Figure 1. HIV-1 RT bound with DHBNH. Although DHBNH primarily inhibits the RNH activity, it binds >50 Å away from the RNH subdomain, at a site that partially overlaps the NNRTI-binding pocket. The subdomains of the p66 subunit are color-coded (fingers in blue, palm in red, thumb in green, connection in yellow, and RNH in gold). Upper left inset: a close-up of the DHBNH binding site. The inhibitor is shown in magenta. The position that would be occupied by an NNRTI is shown in gray (the NNRTI pocket is not occupied in the current structure). Upper right inset: chemical structure of DHBNH.

See a full 3D interactive version of this figure on the ACS Chemical Biology web page. Access to the full structure is available with the molecular visualization tool FirstGlance in Jmol (<http://molvis.sdsc.edu/fgij/index.htm>).

Because of its pivotal role in the HIV life cycle, HIV RT is a primary target for antiretroviral agents. All RT inhibitors currently approved for the treatment of acquired immune deficiency syndrome (AIDS) inhibit the polymerase activity of HIV-1 RT; there are no anti-AIDS drugs that specifically inhibit RNH. There are two major classes of anti-RT drugs: nucleoside/nucleotide RT inhibitors (both called NRTIs for simplicity), and non-nucleoside RT inhibitors (NNRTIs). NRTIs block reverse transcription because they lack a hydroxyl group at the 3'-position of the ribose ring and, when incorporated into viral DNA by RT, act as chain terminators. The NNRTIs, in contrast to NRTIs, bind in a hydrophobic pocket ~ 10 Å from the polymerase active site (Figure 1) and act noncompetitively. Binding an NNRTI does not prevent the binding of the nucleic acid or nucleoside triphosphate substrates to RT; rather, the NNRTI blocks the chemical step of the polymerization reaction (7, 8). Crystallographic studies

(9, 10) have shown that the binding of an NNRTI causes conformational changes near the polymerase active site of HIV-1 RT, including a displacement of the $\beta 12$ - $\beta 13$ - $\beta 14$ sheet that contains the polymerase primer grip (9–12), which is important for properly positioning the nucleic acid relative to the polymerase active site. Binding an NNRTI can also influence the geometry at the polymerase catalytic site (13–15). Many NNRTIs do not affect RNH activity; however, certain NNRTIs, rather than inhibit RNH activity, have been reported to increase the number of RNH cleavages and the rate of RNH activity under certain conditions (16–18).

The early successes of highly active antiretroviral therapy are now threatened by the emergence of drug-resistant viral variants, which arise from the rapid and error-prone replication of the virus (reviewed in ref 19). Because the virus can be suppressed but not eradicated in patients, drug treatments are life-long. This makes the toxicity of many of the existing drugs a significant problem in AIDS therapy. It is important to develop new inhibitors of HIV-1 that will target novel binding sites and inhibit essential viral functions that are not affected by existing drugs. Cross-resistance between such new inhibitors and existing drugs is unlikely. One such target is the RNH activity of HIV-1 RT.

Several classes of HIV-1 RNH inhibitors (RNHIs) have been reported, some of which have sub-micromolar activity. In contrast, most of the effective anti-AIDS drugs have IC_{50} values in the nanomolar or sub-nanomolar range, so the potency of the current RNHIs needs to be substantially improved. One of the most potent classes of RNHIs is the *N*-acyl hydrazone (NAH) analogues that are derivatives of *N*-(4-*tert*-butylbenzoyl)-2-hydroxy-1-naphthaldehyde hydrazone (BBNH) (20, 21). NAH compounds have been shown to inhibit either the polymerase or the RNH activity of RT and, in some cases, both (20, 21).

The development of effective RNHIs has been hampered by the lack of detailed knowledge of how the current lead compounds interact with HIV-1 RT. To better understand the mechanisms of the RNH inhibition and to help design improved inhibitors, we solved the crystal structure of HIV-1 RT in complex with the novel NAH analogue (*E*)-3,4-dihydroxy-*N'*-(2-methoxynaphthalen-1-yl)methylene)benzohydrazide (DHBNH) at 3.15 Å resolution.

RESULTS AND DISCUSSION

Inhibitory Activity of DHBNH. DHBNH inhibited the RNH activity of HIV-1 RT with an IC_{50} of 0.5 μ M (Table 1); this inhibition was noncompetitive with respect to the hybrid duplex nucleic acid substrate (data not shown). In contrast, DHBNH was unable to inhibit the RNA-dependent DNA polymerase (RDDP) activity of HIV-1 RT in standard processive RDDP assays using poly(rA)-oligo(dT) as template-primer. DHBNH was also \sim 40-fold less potent at inhibiting a catalytically active isolated HIV-1 RT-RNH domain (22, 23). This suggests that the binding pocket for interaction of DHBNH with RT may be outside the RNH domain of the enzyme.

Overall Protein Conformation. The RT/DHBNH complex crystallizes with an overall RT conformation similar to that observed in RT/NNRTI complexes. In this conformation, the cleft between the fingers and the thumb of p66 is wider than that in HIV-1 RTs complexed with either DNA-DNA (4, 5, 15) or RNA-DNA (6) template-primers. In the RT/DHBNH structure, the inhibitor does not bind in the vicinity of the RNH active site; instead, it binds to a novel site >50 Å away, between the NNRTI-binding pocket and the polymerase active site (Figures 1 and 2). The binding site, located in the palm of p66, is formed by the β 12- β 13 loop (including the polymerase primer grip, residues 229-231), Val108 of the β 6 strand, and the β 10 strand (including residues 186-188). The inhibitor binds within 3.5 Å of the catalytic YMDD sequence in the β 9- β 10 turn (p66 residues 183-186). DHBNH is oriented with its benzoyl ring partially entering the NNRTI pocket and the naphthyl ring system near the polymerase active site and the polymerase primer grip (Figure 2).

In structures of unliganded RT and RT complexed with nucleic acid, the side chains of Tyr181, Tyr188, and Trp229 fill the NNRTI pocket. Indeed, in the available structures, the pocket does not exist in the absence of an NNRTI. In the RT/DHBNH structure, the side chain of Trp229

is displaced from the pocket, as it is in RT/NNRTI structures; in the RT/DHBNH structure, the Trp229 side chain is positioned to interact with the benzoyl moiety of DHBNH. In structures of RT/NNRTI complexes, the side chains of Tyr181 and Tyr188 normally point toward the polymerase active site, helping to form the

hydrophobic pocket in which the NNRTI binds (3, 9, 10, 24, 25). By contrast, in the RT/DHBNH structure, the side chains of both tyrosines point away from the active site and are in positions similar to those in unliganded RT (9, 10) and RT complexed with nucleic acids (4-6, 15) (Figure 3). Although the side chains of Tyr181 and Tyr188 are in positions similar to those of unliganded RT and fill some of the NNRTI-binding pocket, an unoccupied cavity is present in the part of the pocket adjacent to the DHBNH hydroxyl on the 4-position of the benzoyl ring.

As in RT/NNRTI structures, the polymerase primer grip of the RT/DHBNH structure is substantially displaced from its position in unliganded RT, but its position in RT/DHBNH is significantly different from that seen in many RT/NNRTI complexes. When the NNRTI-binding pocket of the RT/DHBNH structure is superposed on a typical RT/NNRTI structure (the superposition is based on p66 residues 107-112 and 178-215), the position of the catalytic YMDD β 9- β 10 turn is similar in both

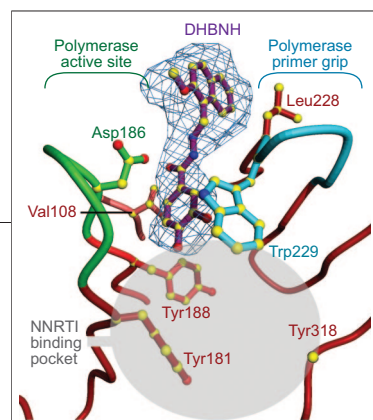


Figure 2. Electron density map of the bound inhibitor. A simulated-annealing $F_o - F_c$ omit map is shown at the 3σ contour level, generated with DHBNH omitted from the phase calculation. The binding site of DHBNH is adjacent to the polymerase active site (green), the polymerase primer grip (cyan), and the NNRTI-binding pocket (gray). The naphthyl ring appears to be positioned so that most of its contacts are made with Leu228. The carbonyl oxygen of Leu228 forms a hydrogen bond with the nitrogen of the DHBNH hydrazone directly adjacent to the benzoyl group. The DHBNH benzoyl ring sits snugly between Trp229 and Tyr188. The electron density appears to favor placement of a hydroxyl at both *meta* positions, suggesting that the benzoyl ring may adopt two alternative conformations.

See a full 3D interactive version of this figure on the ACS Chemical Biology web page. Access to the full structure is available with the molecular visualization tool FirstGlance in Jmol (<http://molvis.sdsc.edu/fgj/index.htm>).

TABLE 1. Some inhibitory properties of DHBNH

Parameter	IC_{50} (μ M) ^a
Inhibition of RT-RNH (intact enzyme)	0.5 ± 0.2 (noncompetitive)
Inhibition of RT RDDP activity	>25
Inhibition of p15-EC RNH fragment	18.5 ± 3.4

^aValues are the means ± standard deviation (SD) determined from at least five separate experiments, each carried out in duplicate.

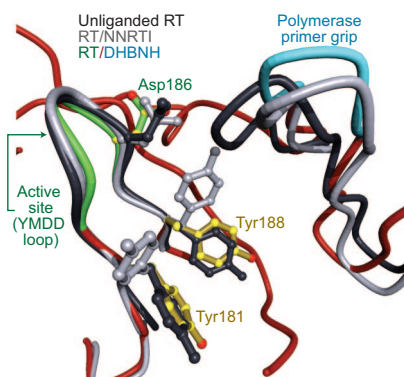


Figure 3. Inhibitor binding sites and conformational changes in the polymerase site. Shown are the superimposed structures of unliganded RT (**10**) (black), RT/TMC125-R165335 (**28**) (an RT/NNRTI complex, gray), and RT/DHBNH (in color). For clarity, the inhibitors are omitted from the diagram. Superposition is based on p66 residues 107–112 and 178–215. The positions and conformations of the polymerase primer grip differ in each structure. The primer grip of unliganded RT repositions to fill a major part of the NNRTI pocket. In the RT/DHBNH structure, although no NNRTI is present, the primer grip lifts up and away from the NNRTI pocket, leaving a cavity that is only partially filled by Tyr181 and Tyr188. The side-chain conformations of Tyr181 and Tyr188 are similar in unliganded RT and RT/DHBNH, whereas in RT/NNRTI complexes these tyrosine side chains typically rotate to form part of the NNRTI pocket. The active site YMDD loop assumes a similar conformation in RT/DHBNH and RT/NNRTI complexes.

See a full 3D interactive version of this figure on the ACS Chemical Biology web page. Access to the full structure is available with the molecular visualization tool FirstGlance in Jmol (<http://molvis.sdsc.edu/fgij/index.htm>).

structures, whereas the position of the polymerase primer grip in the RT/DHBNH structure is further from the active site by >1.4 Å compared with typical RT/NNRTI structures (Figure 3).

Protein–Ligand Interactions. DHBNH appears to have specific interactions with several amino acid residues, including the highly conserved residues Trp229 and Asp186. DHBNH interacts with other residues, including Val108, Leu187, Tyr188, Lys223, Phe227, and Leu228 (Figure 4, Supplementary Table 1). Contacts with most of these residues involve interactions with main-chain or C β atoms or both, suggesting that inhibitors that bind to the same site as DHBNH may be effective against viruses that carry most of the common mutations that give rise to NNRTI resistance. As expected, DHBNH retains full inhibitory potency against the RNH activities of HIV-1 RT mutants that are resistant to a variety of NNRTIs and NRTIs, including Tyr188Leu (Table 2).

The naphthyl ring of DHBNH is solvent-exposed. Most of the stabilizing contacts are with both the side chain and main chain of Leu228, including a possible hydrogen bond with the main-chain nitrogen (Figure 4, Supplementary Table 1). The main-chain carbonyl oxygen of Pro226 also appears to form a hydrogen bond with the naphthyl ring. It is possible that the side chain of Lys223 interacts with the naphthyl ring directly or indirectly (*i.e.*, through a water molecule), but the electron density is ambiguous. The position of the side chain of Lys223 appears to be stabilized by interactions with Glu224 and Pro226. Loop modeling using the program Prime suggests an alternative conformation for Lys223 in which a salt bridge is formed with Asp110. This lower-energy conformation also restricts the binding pocket and may provide additional stability to the bound DHBNH. $3F_o - 2F_c$ difference maps of the naphthyl region suggest the possibility that there may be alterna-

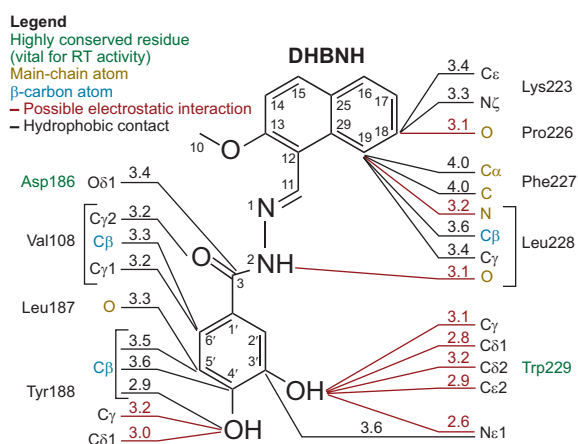


Figure 4. Possible protein–inhibitor contacts. Selected RT/DHBNH interactions are shown with contact distances (≤ 4.0 Å). Inhibitor atom numbering is indicated. All residues in contact with the inhibitor are from the p66 subunit of RT. Highly conserved residues that are vital for RT enzymatic activity are shown in green. Hydrogen bonds and other potential electrostatic interactions are designated by red lines, and hydrophobic interactions are depicted in black. The naphthyl ring has a large number of hydrophobic contacts with Leu228. Electrostatic interactions play a significant role in protein interactions with the rest of the inhibitor. Many RT-inhibitor contacts involve main-chain and β -carbon protein atoms. See Supplementary Table 1 for a comprehensive list of interactions.

TABLE 2. Inhibitory activity of DHBNH against drug-resistant HIV-1 RT variants

Virus/enzyme	IC ₅₀ (μM) ^a	
	DHBNH	Efavirenz (EFV)
Inhibition of RT-RNH		
Y181C RT	0.65 ± 0.1	^b
Y188L RT	1.2 ± 0.4	^b
V106A+Y181C RT	0.85 ± 0.25	^b
L100I+K103N RT	0.7 ± 0.15	^b
D67N+K70R+T215F+K219Q RT	0.5 ± 0.1	^b
Antiviral activity		
Wild type	5.5 ± 1.7	0.002 ± 0.0005
NVP-resistant (Y181C)	8.2 ± 2.5	0.032 ± 0.002
UC781-resistant (V106A+Y181C)	6.7 ± 1.4	0.2 ± 0.01
EFV-resistant (L100I+K103N)	7.7 ± 3.5	7.9 ± 0.3
AZT-resistant (D67N+K70R+T215F+K219Q)	5.6 ± 1.5	0.003 ± 0.001
Cytotoxicity (CC ₅₀)		
MT-2 cells	>100	^c
Peripheral blood mononuclear cells	>100	^c

^aValues are the means ± SD determined from at least three separate experiments.

^bNo inhibition. ^cNot determined.

tive binding modes for the naphthyl ring. The modest resolution produces a large electron density envelope around DHBNH, leading to some uncertainty about the geometry and position of the central (hydrazone) region of the inhibitor. In a crystal structure of the related NAH, BBNH, the free inhibitor was co-planar, except for substituents on the benzoyl ring (M. A. Parniak and G. I. Dmitrienko, unpublished data). Quantum mechanical calculations were performed to generate torsional energy profiles for rotation around each of the DHBNH core angles. Several structures from the Cambridge Structural Database (26) were identified that contained the same core motifs as DHBNH, and these structures were consistent with the energy minima determined for the core torsional angles. These energy profiles were used as guides to model DHBNH into the observed electron density. In the structure, the main-chain carbonyl of Leu228 is 3.1 Å away from the hydrazone nitrogen adjacent to the benzoyl group (Figure 4), which appears to be oriented appropriately to form a hydrogen bond.

Structure–Activity Relationship (SAR) Analysis. The benzoyl ring of DHBNH fits into a tunnel formed by Val108, Tyr188, Phe227, Leu228, and Trp229 that leads directly into the NNRTI-binding pocket. The hydroxyl at one of the *meta* positions on the benzoyl ring is <2.9 Å away from the indole ring of Trp229. The electron density for this portion of DHBNH suggests that there may be partial occupancy for a hydroxyl at the second *meta* position. This would imply that when bound to HIV-1 RT, the benzoyl ring of DHBNH may exist in two

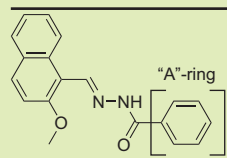
conformations. The second conformation would require a small adjustment in the position of the inhibitor, because there are close contacts with Trp229 and Tyr188, and this adjustment may account for the relatively broad electron density envelope for DHBNH.









The structure predicts that DHBNH derivatives with bulky substitutions at the *para* position of the benzoyl ring would access part of the NNRTI-binding pocket, which could cause increased inhibition of the RT polymerase activity while retaining the ability to inhibit RNH (Figure 2). To test this possibility, we prepared a series of NAHs with increasingly bulky substituents at the *para* position of the benzoyl ring (“A”-ring, Table 3). As predicted, the ability of NAH to inhibit RT DNA polymerase activity is substantially enhanced when the size of the *para* substituent is increased. In contrast, these same substitutions do not significantly affect the ability of the compounds to inhibit the RNH activity of HIV-1 RT.

A recently published structure for RT in a complex with the NNRTI CP-94,707 (27) suggests the idea that there is an opportunity to develop DHBNH derivatives that have NNRTI-like activity. The RT/CP-94,707 structure is similar to the RT/DHBNH structure in that the side chains of Tyr181 and Tyr188 are in the conformation seen in unliganded RT, and the overall conformations of the polymerase active site and NNRTI-binding pocket are very similar to those of RT/DHBNH (Figure 5). However, the binding of CP-94,707 differs from that of DHBNH in that CP-94,707 binds well inside the NNRTI-binding pocket with its benzo-thiazolidinone ring between Trp229 and Tyr188. The benzo-thiazolidinone ring of CP-94,707 appears to play a role similar to the benzoyl ring of DHBNH, which could account for the similarity in the overall conformations of the two structures. When the polymerase active sites of the two structures are superimposed (Figure 5), only the benzo-thiazolidinone ring of CP-94,707 overlaps the benzoyl ring of DHBNH. This superposition suggests that it may be possible to develop inhibitors that contact both the DHBNH and NNRTI binding sites and that such compounds would inhibit both the RNH and polymerase activities of HIV-1 RT.

A “dual inhibitor” could have both disadvantages and advantages. A DHBNH-like inhibitor with substituents that form contacts in the NNRTI pocket may be subject to cross-resistance from mutations that cause resistance to NNRTIs. A properly designed inhibitor,

TABLE 3. Effect of "A"-ring size on inhibitory potency of NAH



"A"-ring	Connolly molecular area (Å ²)	IC ₅₀ (μM) RT polymerase	IC ₅₀ (μM) RT RNH
	100.165	>50	15
	118.819	>50	15
	126.407	>50	12
	146.252	>50	2
	163.147	2	3
	169.97	0.5	5
	180.448	0.4	7
	201.066	1	7.5

however, could avoid this pitfall. For example, flexibility can be built into key positions on the NNRTI-like substituent so that the drug could assume multiple conformations in response to mutations in the NNRTI pocket. Flexibility of the compound is believed to contribute to the resilience of certain diarylpyrimidine-series NNRTIs, which show considerable promise in clinical trials against many NNRTI-resistant mutant strains (28–31). In addition, the DHBNH-like binding contacts could help to compensate for the loss of protein–inhibitor interactions that occur when residues in the NNRTI-binding pocket are mutated. An inhibitor with both DHBNH-like and NNRTI-like contacts could be highly specific for HIV-1 RT. Ongoing studies should help us determine the benefit of substitutions at the *para* position of the benzoyl ring of DHBNH. It should be noted that DHBNH derivatives without substitutions at this position are effective against a variety of NNRTI-resistant HIV-1 mutants, including variants with multiple mutations in the NNRTI-binding pocket (Table 2 and ref 20).

Mechanism of Action.

DHBNH is a sub-micromolar inhibitor of the RNH activity of HIV-1 RT and a very weak inhibitor of RT polymerase activity as measured in standard RT RDDP assays using a DNA primer (IC₅₀ > 25 μM). The superposition of the RT/DHBNH coordinates with structures of RT complexed with either RNA–DNA (6) or DNA–DNA (15) suggests that DHBNH would not prevent the binding of the template–primer or the dNTP substrates. This is consistent with the weak inhibitory activity of DHBNH against RT-catalyzed DNA synthesis. DHBNH does, however, inhibit to some extent the initiation of reverse transcription during HIV replication *in vivo* (Figure 6, panel a) as well as RNA-primed RT-catalyzed DNA polymerization *in vitro* (Figure 6, panel b), although at concentrations substantially higher than those needed to inhibit RNH activity.

Nonetheless, real-time polymerase chain reaction (PCR) analysis suggests that inhibition of the initiation of reverse transcription might make a significant contribution to the antiviral activity of DHBNH. The basis for the inhibition of the initiation of viral DNA synthesis is presently unclear, and additional studies are needed.

The similarity in conformation between RT/DHBNH and RT/CP-94,707 (27) suggests that these two compounds may have a similar mode of action. It is possible, for example, that CP-94,707 might also inhibit RNH activity. CP-94,707 was originally identified as an inhibitor of initiation of transfer-RNA (tRNA) primed DNA synthesis (27), supporting the idea that DHBNH may effectively inhibit initiation of tRNA-primed synthesis of minus-strand DNA.

If DHBNH interferes with correct positioning of a DNA–RNA substrate (*e.g.*, by affecting the position of the polymerase primer grip), then this effect might be even more pronounced with an RNA–RNA substrate. The initiation of HIV RT-catalyzed DNA synthesis with an RNA–RNA

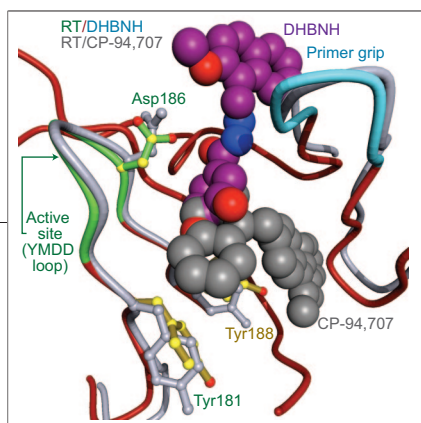


Figure 5. Comparison of binding modes for DHBNH and CP-94,707. The superimposed binding sites of RT complexed with DHBNH and CP-94,707 (27) are shown. In both complexes, Tyr181 and Tyr188 assume conformations similar to those in unliganded RT. However, CP-94,707 binds to a site that is distinct from the DHBNH binding site. CP-94,707 binds to a site closer to the typical NNRTI binding pocket than DHBNH.

See a full 3D interactive version of this figure on the ACS Chemical Biology web page. Access to the full structure is available with the molecular visualization tool FirstGlance in Jmol (<http://molvis.sdsc.edu/fgj/index.htm>).

duplex RNA is intrinsically more rigid than DNA–RNA and is expected to be more refractory to forming the $\sim 40^\circ$ bend between A- and B-forms of the template–primer as seen in complexes of RT with duplex nucleic acid (6, 15). Both of these factors may contribute to the substantially reduced levels of DNA synthesized in assays in which an RNA–RNA substrate is provided to RT in the presence of DHBNH.

Our current structural data for DHBNH are consistent with a mechanism of inhibition that involves the binding of DHBNH proximal to the polymerase active site (Figure 7). We believe this causes structural changes in the polymerase primer grip that may alter the trajectory of the template–primer between the polymerase and RNH domains, so that the sugar–phosphate backbone of the RNA template would not be properly positioned at the RNH active site and could not be cleaved. Changes in specificity and efficiency of RNA cleavage caused by changes in the polymerase domain of RT have been

reported previously. For example, mutations in the polymerase primer grip and thumb of p66 have been shown to dramatically alter the specificity of RNH cleavage of the RNA template strand (32–37). Furthermore, crystallographic studies suggest that the polymerase primer grip plays a role in positioning the RNA template relative to the RNH active site (6, 15). Moreover, there is excellent agreement between the structural data, the model, and the SAR that predicts which of the substituted compounds will interact with the NNRTI-binding pocket.

We cannot exclude the possibility that DHBNH also binds at a second site near the RNH active site. We previously suggested, based on kinetic analysis of the inhibition, that NAHs may inhibit RT polymerase and RNH activities by binding to two different sites on the enzyme (20, 21). A second DHBNH binding site is supported by the observation that DHBNH is equally effective at inhibiting RT RNH activity in the presence or the absence of 20 μM of the NNRTI nevirapine (data not shown), a concentration that would saturate the NNRTI-binding site. When nevirapine is bound to RT, the side chain of Tyr188 rotates toward the polymerase active site (3, 9,

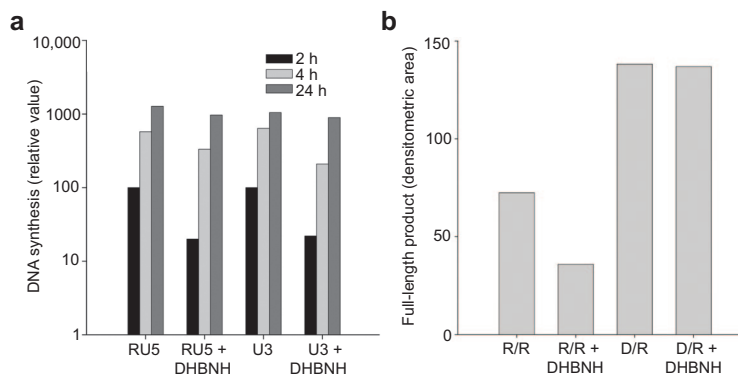


Figure 6. DHBNH inhibits RNA-primed initiation of reverse transcription. a) Real-time PCR analysis of intracellular reverse transcription. The amounts of RU5 and U3 DNA (see ref 6) were measured using real-time PCR to monitor the initiation of reverse transcription and minus-strand DNA transfer, respectively. The y -axis shows the relative amounts of RU5 and U3 DNA 2, 4, and 24 h after infection in the presence or absence of DHBNH. b) DHBNH partially inhibits RNA-primed but not DNA-primed RDDP activity *in vitro*. Reactions were carried out with a 42-nucleotide RNA template annealed to a 21-nucleotide RNA primer (R/R) or to the same sequence DNA primer (D/R) as described in Methods. The figure shows the amount of full-length 21-nucleotide extended DNA polymerization product formed by RT in 10 min. The final concentration of DHBNH in reactions containing the inhibitor was 10 μM . Only starting primer and full-length polymerization products were seen in these assays. The lack of intermediate products suggests that DHBNH was primarily affecting the initiation of reverse transcription and not subsequent elongation.

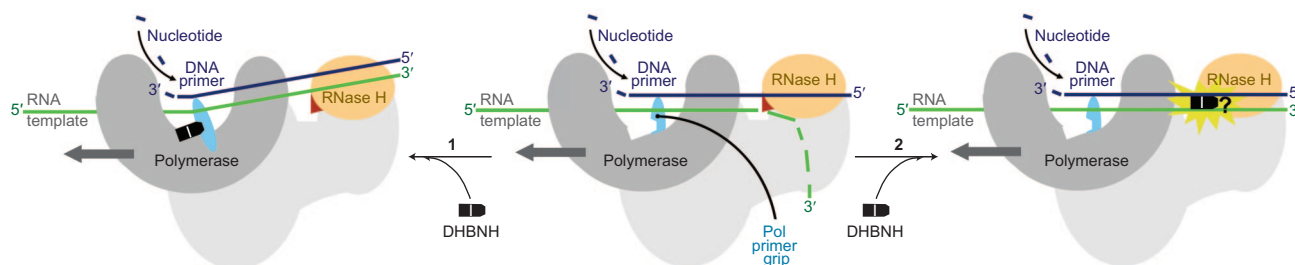


Figure 7. Possible mechanisms for inhibition by DHBNH. Two alternative modes of RNHI activity are represented schematically. 1) As seen in the current RT/DHBNH crystal structure, the inhibitor binds adjacent to the polymerase active site and induces a repositioning of the polymerase primer grip (cyan). This may redirect the trajectory of the substrate so that the RNA strand is not close enough to the RNH active site for cleavage to occur. 2) The inhibitor may bind at a second site close to the RNH domain to inhibit its activity. These two alternative modes of action would not necessarily be exclusive.

10, 24, 25) into a position that would block the binding of the benzoyl ring of DHBNH observed in the current structure. This hypothetical second binding site could be formed in part by the RNA–DNA substrate and would not be present in the crystals we prepared, because nucleic acid was not cocrystallized with the protein. We are trying to obtain a crystal structure of DHBNH bound to the RT–nucleic acid complex.

Conclusions. We have refined a 3.15 Å resolution crystal structure of HIV-1 RT complexed with DHBNH, an NAH-class RNHI. The crystal structure shows that DHBNH binds to a novel site adjacent to the polymerase active site and the NNRTI-binding pocket, >50 Å away from the RNH active site. Binding of DHBNH directly affects the position of the polymerase primer grip, as well as the thumb, which is located adjacent to the primer grip. This

finding is consistent with the possibility that the inhibitor perturbs the trajectory of the template–primer so that RNH cannot cleave the RNA strand of an RNA–DNA duplex. Preliminary SAR studies show that DHBNH derivatives with substituents on the benzoyl ring can interact with residues in the NNRTI-binding pocket to inhibit the polymerase activity, as predicted from the RT/DHBNH structure. The DHBNH binding site provides opportunities to develop new inhibitors that can inhibit the polymerase activity, the RNH activity, or both. More importantly, the prevalence of DHBNH interactions with main-chain and Cβ atoms suggests that inhibitors developed based on chemical modifications of DHBNH have the potential to be effective against a wide range of drug-resistant mutant strains of RT.

METHODS

2-Methoxy-1-naphthaldehyde was obtained from Sigma-Aldrich (St. Louis, MO). 3,4-Dihydroxybenzhydrazide and other acid hydrazides were obtained from TransWorld Chemical (Rockville, MD). [³H]-TTP and the homopolymeric template–primer poly(rA)–oligo(dT)_{12–18} were products of Amersham Biosciences. The oligonucleotides 5′-GAU CUG AGC CUG GGA GCU-fluorescein-3′ and 5′-dabcyl-AGC TCC CAG GCT CAG ATC-3′ were synthesized and provided as an annealed RNA–DNA duplex by TriLink Biotechnologies (San Diego, CA). TriLink Biotechnologies also provided all other oligonucleotides used in these studies.

Synthesis of NAHs. DHBNH and other NAHs were synthesized by condensation of the aromatic aldehyde with the corresponding acid hydrazide, essentially as described previously (38). As an example, DHBNH was prepared by the dropwise addition of 3,4-dihydroxybenzhydrazide (1.1 mM in 10 mL of ethanol) to a solution of 2-methoxy-1-naphthaldehyde (1 mM in 4% acetic acid in ethanol) with stirring while heating in a boiling water bath. Heating and stirring were continued for 20 min following completion of addition of the ethanolic acid hydrazide solution,

and then the mixture was allowed to cool to room temperature. The precipitate was collected by filtration, washed with cold ethanol and diethyl ether, and mass dried. Elemental and mass spectral analyses were consistent with the expected structure (for summary of DHBNH characterization data, see Supplementary Figures 1–3).

Protein Preparation and Purification. HIV-1 RT was prepared as described previously (39). The RT/DHBNH crystallization complex was prepared by mixing 7.4 μL of 20 mM inhibitor in dimethyl sulfoxide (DMSO) with 2.5 μL of 20% β-octyl glucopyranoside. Of this solution, 7.5 μL was combined with 65.0 μL of 40 mg mL⁻¹ RT and 57.5 μL of additional RT buffer (10 mM tris(hydroxymethyl)aminoethane (Tris), pH 8.0, 75 mM NaCl) on ice. A catalytically active chimeric isolated HIV-1 RNH domain protein containing an α-helical substrate-binding loop derived from *Escherichia coli* RNHI (22, 40) (termed p15-EC, a kind gift from C. Shaw-Reid, Merck Research Laboratories, West Point, PA) was expressed and purified as described previously (40).

Assay of RT RDDP Activity. HIV-1 RT RDDP activity was generally determined by a fixed time assay as previously described

(20). Briefly, reaction mixtures (100 μ L total volume) contained 50 mM Tris-HCl (pH 7.8, 37 $^{\circ}$ C), 60 mM KCl, 10 mM $MgCl_2$, 10–25 ng of p51/p66 RT, 0.5 units of template–primer, and 5 μ M [3H]-TTP substrate. Stock solutions of DHBNH and other NAHs were prepared in DMSO. Aliquots of these DMSO solutions containing the inhibitor were added such that the final DMSO concentration was <2%. Reaction assays were incubated at 37 $^{\circ}$ C for 10–20 min and then quenched with 500 μ L of cold 20 mM sodium pyrophosphate in 10% trichloroacetic acid (TCA). After 30 min on ice, the samples were filtered on Whatman 934-AH glass fiber filters and washed with 10% TCA and ethanol, and the radioactivity determined by liquid scintillation spectrometry.

RNA-primed and DNA-primed RT RDDP activity was evaluated using a 42-nucleotide RNA template of the sequence 5'-GGAAAUCUCUACGAGUGGCGCCCGAACAGGGACCUGACCAG-3' annealed to the complementary 21-nucleotide RNA primer 5'-CUGGUCAGGUCCUGUUCGGG-3' or to the complementary 21-nucleotide DNA primer 5'-CTGGTCAGGTCCTGTCGGG-3'. Reaction mixtures (15 μ L final volume) contained 50 nM template–primer, 250 nM p66/p51 RT, and 25 μ M of each dNTP ($[\alpha\text{-}^{32}P]$ -dCTP was included as tracer) in 50 mM Tris, pH 8.0, containing 60 mM KCl and 5 mM $MgCl_2$, in the presence or the absence of 10 μ M DHBNH. Briefly, RT and template–primer were incubated at 37 $^{\circ}$ C for 5 min, and then the reactions were started by the addition of the dNTPs and $MgCl_2$ to the final concentrations indicated. After a 10 min incubation at 37 $^{\circ}$ C, reactions were quenched by the addition of an equal volume of sequencing gel loading buffer (98% deionized formamide, 10 mM EDTA, and 1 mg mL^{-1} each of bromophenol blue and xylene cyanol) followed by heating at 95 $^{\circ}$ C for 5 min. Reaction products were resolved by denaturing gel electrophoresis, and the amount of full-length 42-nucleotide polymerization product was quantified by phosphorimaging.

Assay of RT RNH Activity. RT RNH activity was measured using a FRET-based microplate fluorescence assay that we have recently described (41). Briefly, 50 μ L of a 0.5 μ M solution of RNA–DNA hybrid duplex in 50 mM Tris, pH 8.0, containing 60 mM KCl, was added to individual wells of a 96-well microplate. Reactions were initiated by the addition of 50 μ L of 5 nM HIV-1 RT in 50 mM Tris, pH 8.0, containing 60 mM KCl and 10 mM $MgCl_2$, and allowed to proceed at 37 $^{\circ}$ C for 30 min. Reactions were quenched by the addition of 50 μ L of 0.5 M EDTA, pH 8.0. Fluorescence intensity was assessed using an excitation wavelength of 490 nm and an emission wavelength of 528 nm, with cutoff filter set to 515 nm. To assess the effect of inhibitors, 1 μ L of a DMSO inhibitor solution was added to the microplate well prior to the addition of substrate and RT solutions in order to ensure adequate mixing and suspension of the hydrazone in the aqueous reaction medium.

Cell Culture and Antiviral Assays. Antiviral assays were carried out by infection of MT-2 lymphoblastoid cells in the presence of varying concentrations of DHBNH essentially as previously described (20). Briefly, MT-2 cells (4×10^5 cells mL^{-1}) were incubated overnight in the absence or in the presence of varying concentrations of DHBNH and were then infected with HIV-1 (NL4-3, moi of 0.01). The extent of HIV-1 replication was evaluated 5 d post-infection by microscopic evaluation of HIV-induced cytopathic effect and by analysis of HIV-1 p24 antigen levels in cell-free culture supernatants. Cytotoxicity was evaluated in MT-2 cells and in peripheral blood mononuclear cells using (2,3-bis(2-methoxy-4-nitro-5-sulphophenyl)-5-[(phenylamino)carbonyl]-2H-tetrazolium hydroxide, XTT in a commercially available kit (Roche Diagnostics, Indianapolis IN) according to manufacturer's directions. Cells were incubated with varying concentrations of DHBNH (0–200 μ M) for 5 d prior to evaluation of XTT reactivity.

Real Time PCR Analysis of Intracellular Reverse Transcription

Products. 293 and HOS cells were maintained in Dulbecco's modified Eagle's medium (Life Technologies) supplemented with 5% fetal bovine serum, 5% newborn calf serum, and penicillin (50 units mL^{-1}) plus streptomycin (50 μ g mL^{-1} ; Quality Biological). VSV-g pseudotyped HIV vectors that undergo a single cycle of replication were generated by cotransfecting 293 cells with 5 μ g of pNLNcoMIVR-E-HSA and 3 μ g of pHCMV-g using the calcium phosphate method. The 293 cells were washed with 10 mL of phosphate-buffered saline (PBS) 8, 24, and 32 h after transfection. The 48-h supernatants were harvested, clarified by centrifugation, and filtered through a 0.45 μ m filter. The supernatants were treated with 100 units of RNase-free DNase I (Roche) for 30 min at room temperature and were then concentrated to 2 mL using the 300,000 MWCO Vivaspin 20 mL concentrators. The viruses were diluted to a total volume of 13 mL in complete media, and then 2 mL of diluted virus was used to infect 2×10^5 HOS cells. Stocks (10 mM) of DHBNH were prepared in DMSO and diluted to a final concentration of 10 μ M in the treated groups, which were incubated for 4 h with DHBNH prior to viral infection. HOS cells were washed with 2 mL of PBS 2 h after infection, and then fresh medium was added. Total DNA was isolated 2, 4, and 24 h after infection using the EZ-1 DNA Tissue protocol (Qiagen). Real-time PCR reactions were used to quantitate the amounts of DNA in the infected cells as previously described (42). Virions containing the D110E polymerase active site mutation were also generated and used to demonstrate that the plasmid DNA used in the transfection did not significantly contribute to the amount of viral DNA measured.

Crystallization and Data Collection. RT/DHBNH was crystallized by vapor diffusion in microseeded hanging drops containing equal volumes of protein (above) and reservoir solution (50 mM imidazole, pH 6.4, 100 mM ammonium sulfate, 15 mM magnesium sulfate, 5% glucose, 11.5% poly(ethylene glycol) (PEG) 8000) at 4.0 $^{\circ}$ C. RT/DHBNH crystals were transferred to a stabilization solution containing mother liquor and 15% PEG 8000. This was replaced stepwise at 5–10 min intervals with stabilization solutions containing increasing concentrations of sucrose in 5% increments until the sucrose concentration was 25%. The crystals were subsequently flash-cooled and stored in liquid N_2 . X-ray data were collected at 100 K at the National Synchrotron Light Source at Brookhaven National Laboratories, Beamline X25. The data were processed using HKL-DENZO-SCALEPACK (43). Crystallographic statistics are shown in Supplementary Table 2).

Computational Methods. Density functional theory calculations were carried out with the Gaussian 98 software package (44) using the B3LYP functional model (45, 46) and the 6-31G* basis set. Each of the four torsional angles in the DHBNH core was scanned in 15 $^{\circ}$ increments followed by geometry optimization. Loop modeling of p66 residues Gln222 through Leu228 was performed using Prime (Schrödinger, LLC) with the AGBNP solvation model (47, 48).

Structure Determination and Refinement. Phases were determined by molecular replacement with the program AMoRe (49) using the HIV-1 RT/R100943 structure (PDB accession number 1S6P) as an initial search model. Stepwise model building and torsional simulated annealing refinement protocols were conducted using the O graphics package (50) and CNS (51) with a bulk solvent correction. Ligand geometry was optimized using the IMPACT software package (Schrödinger, LLC). Restrained minimization of the entire complex, in which hydrogen atoms were added to the structure, side-chain hydroxyl groups were reoriented, and potential steric clashes were alleviated, was performed using Schrödinger's FirstDiscovery protein preparation facility (Schrödinger, LLC). Results from the energetic calculations performed in FirstDiscovery and Prime (above) were itera-

tively incorporated into the crystallographic refinement (unpublished procedure). Water molecules were added manually in the final stages of refinement and were only built in where they could be justified by hydrogen bonds and $F_o - F_c$ electron density at or above the 2.5σ contour level.

Accession Codes: The atomic coordinates for the refined structure of RT/DHBNH have been deposited in the Protein Data Bank (PDB accession number 2I5J).

Acknowledgments: This research was supported by National Institutes of Health (NIH) Grants AI 27690 (MERIT Award to E.A.) and P01 GM 066671 (to E.A. and M.A.P.), GM64375 (to R.M.L.), and F32 AI 060300 (NIH NRSA fellowship to D.M.H.). We are grateful to synchrotron staff members at Brookhaven National Light Source and other members of our laboratories for their assistance and helpful discussions. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-98CH10886. This publication has been funded in part with federal funds from the National Cancer Institute, NIH, under contract no. NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, as well as the National Institute of General Medical Sciences.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Coffin, J. M., Hughes, S. H., and Vamusi, H. E. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Tisdale, M., Schulze, T., Larder, B. A., and Moelling, K. (1991) Mutations within the RNase H domain of HIV-1 RT abolish virus infectivity, *J. Gen. Virol.* 72, 59–66.
- Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor, *Science* 256, 1783–1790.
- Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr, Ju, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA, *Proc. Natl. Acad. Sci. U.S.A.* 90, 6320–6324.
- Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: Implications for drug resistance, *Science* 282, 1669–1675.
- Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Jr, Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA, *EMBO J.* 20, 1449–1461.
- Rittinger, K., Divita, G., and Goody, R. S. (1995) Human immunodeficiency virus reverse transcriptase substrate-induced conformational changes and the mechanism of inhibition by nonnucleoside inhibitors, *Proc. Natl. Acad. Sci. U.S.A.* 92, 8046–8049.
- Spence, R. A., Kati, W. M., Anderson, K. S., and Johnson, K. A. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors, *Science* 267, 988–993.
- Rodgers, D. W., Camblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C., and Harrison, S. C. (1995) The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1, *Proc. Natl. Acad. Sci. U.S.A.* 92, 1222–1226.
- Hsiou, Y., Ding, J., Das, K., Clark, A. D., Jr., Hughes, S. H., and Arnold, E. (1996) Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: Implications of conformational changes for polymerization and inhibition mechanisms, *Structure* 4, 853–860.
- Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J., and Arnold, E. (1994) Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase: implications for mechanisms of drug inhibition and resistance, *J. Mol. Biol.* 243, 369–387.
- Das, K., Ding, J., Hsiou, Y., Clark, A. D., Jr., Moereels, H., Koymans, L., Andries, K., Pauwels, R., Janssen, P. A. J., Boyer, P. L., Clark, P., Smith, R. H., Jr., Smith, M. B. K., Michejda, C. J., Hughes, S. H., and Arnold, E. (1996) Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant, *J. Mol. Biol.* 264, 1085–1100.
- Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., and Stuart, D. I. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors, *Nat. Struct. Biol.* 2, 303–308.
- Ding, J., Das, K., Hsiou, Y., Zhang, W., and Arnold, E. (1997) Structural studies of HIV-1 reverse transcriptase and implications for drug design, in *Structure-based Drug Design* (Veerapandian, P., Ed.) pp 41–82, Marcel Dekker, Inc., New York.
- Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jr, Jacobo-Molina, A., Tantillo, C., Hughes, S. H., and Arnold, E. (1998) Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 Å resolution, *J. Mol. Biol.* 284, 1095–1111.
- Palaniappan, C., Fay, P. J., and Bambara, R. A. (1995) Nevirapine alters the cleavage specificity of ribonuclease H of human immunodeficiency virus 1 reverse transcriptase, *J. Biol. Chem.* 270, 4861–4869.
- Shaw-Reid, C. A., Feuston, B., Munshi, V., Getty, K., Krueger, J., Hazuda, D. J., Parniak, M. A., Miller, M. D., and Lewis, D. (2005) Dissecting the effects of DNA polymerase and ribonuclease H inhibitor combinations on HIV-1 reverse-transcriptase activities, *Biochemistry* 44, 1595–1606.
- Gopalakrishnan, V., and Benkovic, S. (1994) Effect of a thiobenzimidazolone derivative on DNA strand transfer catalyzed by HIV-1 reverse transcriptase, *J. Biol. Chem.* 269, 4110–4115.
- Sarafianos, S. G., Das, K., Hughes, S. H., and Arnold, E. (2004) Taking aim at a moving target: designing drugs to inhibit drug-resistant HIV-1 reverse transcriptases, *Curr. Opin. Struct. Biol.* 14, 716–730.
- Borkow, G., Fletcher, R. S., Bamard, J., Arion, D., Motakis, D., Dmitrienko, G. I., and Parniak, M. A. (1997) Inhibition of the ribonuclease H and DNA polymerase activities of HIV-1 reverse transcriptase by *N*-(4-*tert*-butylbenzoyl)-2-hydroxy-1-naphthaldehyde hydrate, *Biochemistry* 36, 3179–3185.
- Sluis-Cremer, N., Arion, D., and Parniak, M. A. (2002) Destabilization of the HIV-1 reverse transcriptase dimer upon interaction with *N*-acyl hydrazone inhibitors, *Mol. Pharmacol.* 62, 398–405.
- Keck, J. L., and Marqusee, S. (1995) Substitution of a highly basic helix/loop sequence into the RNase H domain of human immunodeficiency virus reverse transcriptase restores its Mn(2+)-dependent RNase H activity, *Proc. Natl. Acad. Sci. U.S.A.* 92, 2740–2744.
- Shaw-Reid, C. A., Munshi, V., Graham, P., Wolfe, A., Witmer, M., Danzeisen, R., Olsen, D. B., Carroll, S. S., Embrey, M., Wai, J. S., Miller, M. D., Cole, J. L., and Hazuda, D. J. (2003) Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid, *J. Biol. Chem.* 278, 2777–2780.

24. Ding, J., Das, K., Moereels, H., Koymans, L., Andries, K., Janssen, P. A. J., Hughes, S. H., and Arnold, E. (1995) Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse non-nucleoside inhibitors, *Nat. Struct. Biol.* **2**, 407–415.
25. Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., and Stuart, D. (1995) High resolution structures of HIV-1 RT from four RT-inhibitor complexes, *Nat. Struct. Biol.* **2**, 293–302.
26. Allen, F. H. (2002) The Cambridge Structural Database: A quarter of a million crystal structures and rising, *Acta Crystallogr. B* **58**, 380–388.
27. Pata, J. D., Stirtan, W. G., Goldstein, S. W., and Steitz, T. A. (2004) Structure of HIV-1 reverse transcriptase bound to an inhibitor active against mutant reverse transcriptases resistant to other non-nucleoside inhibitors, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10548–10553.
28. Das, K., Clark, A. D., Jr., Lewi, P. J., Heeres, J., de Jonge, M. R., Koymans, L. M. H., Vinkers, H. M., Daeyaert, F., Ludovici, D. W., Kukla, M. J., Corte, B. D., Kavash, R. W., Ho, C. Y., Ye, H., Lichtenstein, M. A., Andries, K., Pauwels, R., Béthune, M.-P., Boyer, P. L., Clark, P., Hughes, S. H., Janssen, P. A. J., and Arnold, E. (2004) Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants, *J. Med. Chem.* **47**, 2550–2560.
29. Das, K., Lewi, P. J., Hughes, S. H., and Arnold, E. (2005) Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors, *Prog. Biophys. Mol. Biol.* **88**, 209–231.
30. Guillemont, J., Pasquier, E., Palandjian, P., Vernier, D., Gaurrand, S., Lewi, P. J., Heeres, J., de Jonge, M. R., Koymans, L. M. H., Daeyaert, F. F. D., Vinkers, M. H., Arnold, E., Das, K., Pauwels, R., Andries, K., de Béthune, M.-P., Bettens, E., Hertogs, K., Wigerinck, P., Timmerman, P., and Janssen, P. A. J. (2005) Synthesis of novel diarylpyrimidine analogues and their antiviral activity against human immunodeficiency virus type 1, *J. Med. Chem.* **48**, 2072–2079.
31. Janssen, P. A. J., Lewi, P. J., Arnold, E., Daeyaert, F., de Jonge, M., Heeres, J., Koymans, L., Vinkers, M., Guillemont, J., Pasquier, E., Kukla, M., Ludovici, D., Andries, K., de Béthune, M.-P., Pauwels, R., Das, K., Clark, A. D., Jr., Frenkel, Y. V., Hughes, S. H., Medaer, B., De Knaep, F., Bohets, H., De Clerck, F., Lampo, A., Williams, P., and Stofels, P. (2005) In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoetenyl]-2,6-dimethylphenyl]amino]-2-pyrimidinyl]amino]-benzonitrile (R278474-rilpivirine), *J. Med. Chem.* **48**, 1901–1909.
32. Ghosh, M., Jacques, P. S., Rodgers, D. W., Ottman, M., Darlix, J.-L., and Le Grice, S. F. J. (1996) Alterations to the primer grip of P66 HIV-1 reverse transcriptase and their consequences for template-primer utilization, *Biochemistry* **35**, 8553–8562.
33. Ghosh, M., Williams, J., Powell, M. D., Levin, J. G., and Le Grice, S. F. (1997) Mutating a conserved motif of the HIV-1 reverse transcriptase palm subdomain alters primer utilization, *Biochemistry* **36**, 5758–5768.
34. Palaniappan, C., Wisniewski, M., Jacques, P. S., Le Grice, S. F., Fay, P. J., and Bambara, R. A. (1997) Mutations within the primer grip region of HIV-1 reverse transcriptase result in loss of RNase H function, *J. Biol. Chem.* **272**, 11157–11164.
35. Powell, M. D., Ghosh, M., Jacques, P. S., Howard, K. J., Le Grice, S. F., and Levin, J. G. (1997) Alanine-scanning mutations in the “primer grip” of P66 HIV-1 reverse transcriptase result in selective loss of RNA priming activity, *J. Biol. Chem.* **272**, 13262–13269.
36. Gao, H. Q., Boyer, P. L., Arnold, E., and Hughes, S. H. (1998) Effects of mutations in the polymerase domain on the polymerase, RNase H and strand transfer activities of human immunodeficiency virus type 1 reverse transcriptase, *J. Mol. Biol.* **277**, 559–572.
37. Powell, M. D., Beard, W. A., Bebenek, K., Howard, K. J., Le Grice, S. F., Darden, T. A., Kunkel, T. A., Wilson, S. H., and Levin, J. G. (1999) Residues in the alphaH and alpha helices of the HIV-1 reverse transcriptase thumb subdomain required for the specificity of RNase H-catalyzed removal of the polypurine tract primer, *J. Biol. Chem.* **274**, 19885–19893.
38. Edward, J. T., Gauthier, M., Chubb, F. L., and Ponka, P. (1988) Synthesis of new acylhydrazones as iron-chelating compounds, *J. Chem. Eng. Data* **33**, 538–540.
39. Clark, A. D., Jr., Jacobo-Molina, A., Clark, P., Hughes, S. H., and Arnold, E. (1995) Crystallization of human immunodeficiency virus type 1 reverse transcriptase with and without nucleic acid substrates, inhibitors and an antibody Fab fragment, *Methods Enzymol.* **262**, 171–185.
40. Carroll, S. S., Sardana, V., Yang, Z., Jacobs, A. R., Mizenko, C., Hall, D., Hill, L., Zugar-Murphy, J., and Kuo, L. C. (2000) Only a small fraction of purified hepatitis C RNA-dependent RNA polymerase is catalytically competent: Implications for viral replication and *in vitro* assays, *Biochemistry* **39**, 8243–8249.
41. Pamiak, M. A., Min, K. L., Budihis, S. R., Le Grice, S. F., and Beutler, J. A. (2003) A fluorescence-based high-throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity, *Anal. Biochem.* **322**, 33–39.
42. Julias, J. G., Ferris, A. L., Boyer, P. L., and Hughes, S. H. (2001) Replication of phenotypically mixed human immunodeficiency virus type 1 virions containing catalytically active and catalytically inactive reverse transcriptase, *J. Virol.* **75**, 6537–6546.
43. Otwinowski, Z., and Minor, W. (2001) DENZO and SCALEPACK, in *Crystallography of Biological Macromolecules* (Rossmann, M. G., and Arnold, E., Eds.) pp 226–235, Kluwer Academic Publishers, Boston.
44. Frisch, M. J., Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., Pople, J. A. (1998) *Gaussian 98*, revision A.9, Gaussian, Inc., Pittsburgh, PA.
45. Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.* **98**, 5648–5652.
46. Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B: Condens. Matter* **37**, 785–789.
47. Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., and Friesner, R. A. (2004) A hierarchical approach to all-atom loop prediction, *Proteins: Struct., Funct., Bioinf.* **55**, 351–367.
48. Gallicchio, E., and Levy, R. M. (2004) AGBNP: an analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling, *J. Comput. Chem.* **25**, 479–499.
49. Navaza, J. (1994) AMoRe: an automated package for molecular replacement, *Acta Crystallogr. A* **50**, 157–163.
50. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved experimental procedures for building protein models in electron-density maps and the location of errors in these models, *Acta Crystallogr. A* **47**, 110–119.
51. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr. D* **54**, 905–921.

Mechanism-Based Probe for the Analysis of Cathepsin Cysteine Proteases in Living Cells

Howard C. Hang^{†,*}, Joana Loureiro[†], Eric Spooner[†], Adrianus W. M. van der Velden[§], You-Me Kim[†], Annette M. Pollington[†], Rene Maehr[¶], Michael N. Starnbach[§], and Hidde L. Ploegh^{†,*}

[†]Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142,

[§]Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, ^{*}Current

address: Laboratory of Chemical Biology and Microbial Pathogenesis, The Rockefeller University, New York, New York

10021, [¶]Current address: Department of Molecular Cell Biology, Harvard University, Cambridge, Massachusetts 02142

The explosion of genomic sequences from many organisms has provided unprecedented opportunities for the global analysis of complex biological processes. While the advances in nucleotide microarrays and mass spectrometry (MS) have enabled large-scale comparative analyses of gene and protein expression, respectively, these technologies do not fully reveal the functional complexity of proteomes that are regulated by post-transcriptional/translational mechanisms. To address the functional complexity of proteomes, new chemical strategies have been developed for the targeted analysis of individual protein superfamilies (1). In particular, mechanism-based probes have been developed for covalent labeling of enzyme families through active site nucleophiles (*i.e.*, serine hydrolases (2) and cysteine proteases (3–5)) or by photochemical cross-linking (metalloproteases) (6, 7). These mechanism-based probes are typically appended with chemical or epitope tags such as ¹²⁵I, fluorophores, or biotin to visualize covalently labeled polypeptides in cell lysates or tissue samples after gel-based separation. Since these mechanism-based probes are designed to target the active site of enzymes, the extent of protein labeling with these chemical probes often reflects the amount of active enzyme present in the sample, which provides details of function not available by measuring transcript or protein abundance alone. These chemical approaches have been termed activity-based protein profiling and have provided insight into protein function previously unattainable by either gene or protein expression analyses (1). Mechanism-based probes are also providing new tools for drug discovery, small-molecule target identification, and the discovery of previously uncharacterized protein activities (1).

ABSTRACT Mechanism-based probes are providing new tools to evaluate the enzymatic activities of protein families in complex mixtures and to assign protein function. The application of these chemical probes for the visualization of protein labeling in cells and proteomic analysis is still challenging. As a consequence, imaging and proteomic analysis often require different sets of chemical probes. Here we describe a mechanism-based probe, azido-E-64, that can be used for both imaging and proteomics. Azido-E-64 covalently modifies active Cathepsin (Cat) B in living cells, an abundant cysteine protease involved in microbial infections, apoptosis, and cancer. Furthermore, azido-E-64 contains an azide chemical handle that can be selectively derivatized with phosphine reagents *via* the Staudinger ligation, which enables the imaging and proteomic analysis of Cat B. We have utilized azido-E-64 to visualize active Cat B during infection of primary macrophages with *Salmonella typhimurium*, a facultative intracellular bacterial pathogen. These studies demonstrated that active Cat B is specifically excluded from *Salmonella*-containing vacuoles, which suggests that inhibition of protease activity within bacteria-containing vacuoles may contribute to bacterial virulence.

*Corresponding authors,
ploegh@wi.mit.edu, hang@wi.mit.edu.

Received for review October 18, 2006
and accepted November 15, 2006.

Published online December 8, 2006

10.1021/cb600431a CCC: \$33.50

© 2006 by American Chemical Society

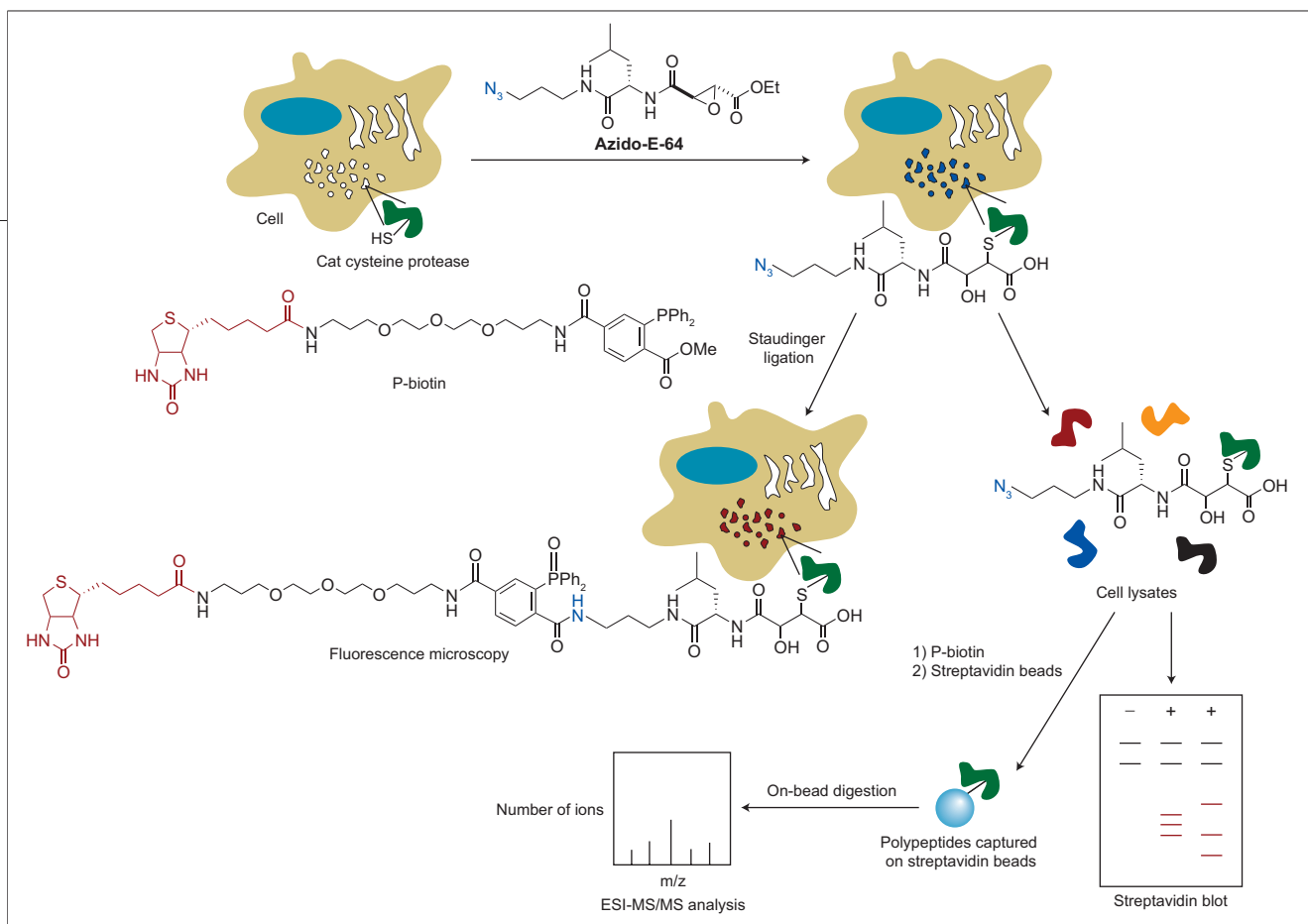


Figure 1. Mechanism-based probe for labeling active Cat cysteine proteases in living cells. Treatment of cells with azido-E-64 enables covalent modification of Cat cysteine proteases in living cells. Reaction of the azide chemical handle on azido-E-64 with p-biotin via the Staudinger ligation allows visualization of active proteases by fluorescence microscopy or streptavidin blot and also enables proteomic analysis of azido-E-64-labeled proteins by MS after affinity enrichment with streptavidin beads.

Microbial pathogens utilize complex mechanisms to infect host cells and cause disease, which requires new approaches to dissect these interactions for therapeutic intervention (8). Mechanism-based probes present novel chemical tools to understand microbial pathogenesis. For example, the application of mechanism-based probes to the *Plasmodium falciparum* life cycle suggested the cysteine protease falcipain 1 plays an essential role during parasite invasion of host cells and may be a target for antimalarial therapeutics (9). We recently used a mechanism-based probe for deubiquitinating enzymes to assess the role of these cysteine proteases during herpesvirus infection (3, 10). Our analysis uncovered a new family of virally encoded deubiquitinating enzymes embedded within the tegument protein that is conserved throughout the herpesviridae (11), the deubiquitinating activities of which were not evident from bioinformatic analysis of viral genomes. Furthermore, this mechanism-based probe has identified deubiquitinating enzymes from *Chlamydia trachomatis*, a Gram-negative bacterial pathogen with currently no available methods for genetic manipulation (12) as well as the parasite *P. falciparum* (13). Chemical probes are beginning to provide new insights into microbial pathogen-

esis; however, cell-permeable mechanism-based probes will be required to understand these complex interactions *in vivo*.

While mechanism-based probes have enabled the selective profiling of various enzyme families in complex mixtures *in vitro* (1), significant challenges still exist for the application of these mechanism-based probes to living cells and for proteomic analysis of labeled polypeptides. Large chemical tags, such as biotin or fluorophores utilized for the visualization or proteomic analysis of labeled polypeptides, may alter the specificity of protein labeling or passive diffusion into cells. These challenges have necessitated different sets of probes for proteomic analysis and for imaging. For example, the epoxysuccinate mechanism-based probe DCG-04 provides an excellent tool for the proteomic analysis of the Cathepsin (Cat) cysteine protease family (4, 14). However, the biotin moiety on DCG-04 prevents passive diffusion of this probe across cellular membranes, which precludes its application to living cells (4, 14). Fluorophore-modified derivatives of DCG-04 have also been generated for visualization of Cat cysteine proteases in live cells (15) and *in vivo* (16), but these have limitations due to background fluorescence of unbound

probes. Recently, the development of quenched activity-based probes (qABPs) by Blum *et al.* (17) provides an elegant method to visualize Cat cysteine proteases in cells. These qABPs are based on acyloxyl methyl ketone mechanism-based probes that are modified with a fluorophore and quencher. The latter departs following protein labeling to yield a fluorescently tagged enzyme (17). Collectively, biotinylated probes (4) and qABPs (17) afford two different sets of ABPs for proteomics and imaging, respectively. Ideally, a single ABP could function for both proteomic analysis and imaging in cells.

The development of two bioorthogonal labeling reactions, the Staudinger ligation (18) and the Huisgen [3 + 2] cycloaddition (19), decouples protein labeling from visualization and proteomic analysis, which provides exciting opportunities for the application of mechanism-based probes in living cells and *in vivo*. The Staudinger ligation allows the use of the azide as a small chemical tag that can be converted into a variety of functionalities, including biotin or fluorophores by reaction with appropriately designed phosphine reagents (Figure 1) (18). Likewise, the Huisgen [3 + 2] cycloaddition, or “click-chemistry”, allows the azide or alkynes to be used as small chemical tags for subsequent attachment of visualization or affinity tags (20). In fact, these two-step labeling approaches have been applied toward mechanism-based probes for the in-cell labeling of the proteasome (21) and serine hydrolases (20), using the Staudinger ligation and click chemistry, respectively. More recently, these bioorthogonal reactions have enabled the MS-based proteomic analysis of azide- (22) and alkyne-labeled proteins (23, 24).

Here we describe a single-mechanism-based probe, azido-E-64, that enables live cell labeling of Cat cysteine proteases and subsequent proteomic analysis as well as visualization of Cat cysteine proteases in cells by fluorescence microscopy (Figure 1). Azido-E-64 is equipped with a small azide chemical handle that enables diffusion of the probe through cellular membranes. Proteomic analysis and visualization of the covalently modified proteases are accomplished after bioorthogonal labeling with a phosphine reagent *via* the Staudinger ligation (18) (Figure 1). Using this novel mechanism-based probe, we demonstrate active Cat B is excluded from *Salmonella*-containing vacuoles (SCVs) in primary macrophages, suggesting that inhibition of endocytic protease activity may contribute to the survival of intracellular bacterial pathogens.

RESULTS AND DISCUSSION

Azido-E-64 Irreversibly Labels Active Cat Cysteine

Proteases in Living Cells. To monitor the activity of the Cat cysteine proteases in living cells, we synthesized a mechanism-based probe azido-E-64 (Figure 2), along with a free acid derivative azido-E-64-OH (Figure 2 and Supplementary Figure 1). The design of these mechanism-based probes was based on the natural product E-64 (25) and DCG-04 (4) (Figure 2, panel a), two well-characterized epoxysuccinate protease inhibitors that covalently modify the active site thiol of Cat cysteine proteases. The inhibitory activity of azido-E-64 and azido-E-64-OH was evaluated by incubation of cell lysates from the RAW264.7 macrophage cell line with increasing concentrations of the mechanism-based probes and assayed for DCG-04 labeling to visualize the remaining active sites not already modified (Figure 2, panel b). E-64 was also included in these assays for comparison (Figure 2). DCG-04 covalently modifies the active site thiol of Cat cysteine proteases, which is then detected by virtue of the biotin moiety in DCG-04 after SDS-PAGE followed by streptavidin–horseradish peroxidase (HRP) blotting (4, 14). Azido-E-64 was less potent at blocking DCG-04 labeling *in vitro* compared to E-64 or azido-E-64-OH (Figure 2, panel b). In contrast, preincubation of live macrophages with azido-E-64 abrogated DCG-04 labeling of cell lysates more efficiently than did E-64 or azido-E-64-OH and completely blocked labeling with DCG-04 at 13 μ M azido-E-64 (Figure 2, panel c). The greater potency of azido-E-64 compared to E-64 or azido-E-64-OH in cells is presumably due to the ethyl ester on azido-E-64, which is likely converted into the free acid by nonspecific esterases within cells. These observations are consistent with previous reports comparing E-64 derivatives (26).

To visualize cellular targets of azido-E-64 directly, cell lysates from macrophages treated with azido-E-64 were reacted with phosphine-biotin (p-biotin) for bioorthogonal labeling of azide-modified proteins (Figure 1). In the absence of azido-E-64, no p-biotin labeling was observed in the molecular mass range (20–37 kDa) expected for the Cat cysteine proteases (Figure 3). The intensity of azido-E-64 labeling increased in a dose-dependent manner (Figure 3, panel a). Azido-E-64 labeling was also time-dependent, was observable within 5 min of incubation, and reached saturation within 40 min (Figure 3, panel b). In addition, treatment of RAW264.7 macrophages with several known endocytic

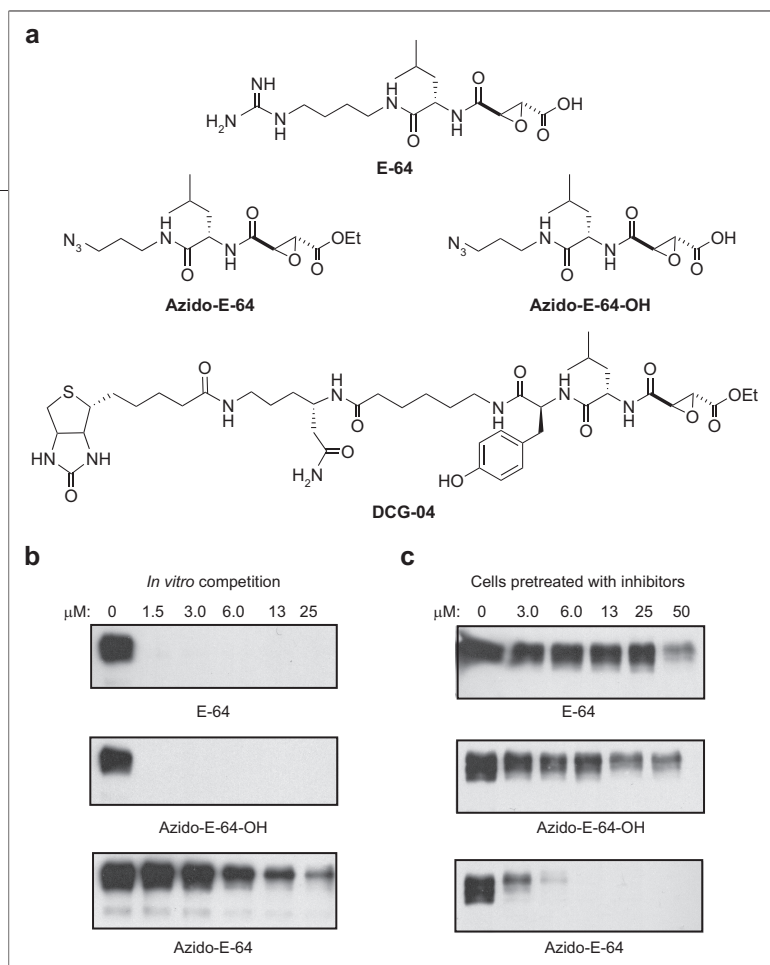


Figure 2. Azido-E-64 derivatives irreversibly inhibit Cat cysteine proteases *in vitro* and in cells. a) Mechanism-based irreversible inhibitors of Cat cysteine proteases based upon epoxysuccinate reactive group. b) Dose-dependent inhibition of DCG-04 labeling of RAW264.7 macrophage lysates treated with E-64, azido-E-64-OH, or azido-E-64. Cell lysates at pH 5.5 were labeled with DCG-04 (5 μ M) for 1 h in the presence of increasing concentrations of E-64, azido-E-64-OH, or azido-E-64. Reactions were terminated by acetone precipitation, and proteins were analyzed by streptavidin blotting. c) DCG-04 labeling of Cat cysteine proteases from RAW264.7 macrophages treated with E-64, azido-E-64-OH, or azido-E-64. Cells were incubated with increasing concentrations of E-64, azido-E-64-OH, or azido-E-64 for 1 h at 37 $^{\circ}$ C. Cell lysates at pH 5.5 were prepared and labeled with DCG-04 (5 μ M) for 1 h in the presence of increasing concentrations of E-64, azido-E-64-OH, or azido-E-64. Reactions were terminated by acetone precipitation, and proteins were analyzed by streptavidin blotting.

cysteine protease inhibitors prior to the addition of azido-E-64 inhibited the labeling of target proteins (Figure 3, panel c). Leupeptin, a broad-spectrum cysteine protease inhibitor, blocked the majority of azido-E-64 labeling, while the Cat B selective inhibitor (CA-074b-OMe) (27) primarily inhibited the labeling of a polypeptide at lower molecular weight (Figure 3, panel c), presumed to be Cat B (*vide infra*). Furthermore, addition of ammonium chloride or bafilomycin, a v-ATPase H^{+} pump inhibitor, both of which increase endosomal pH, also reduced azido-E-64 labeling (Figure 3, panel c). Conversely, the broad-spectrum aspartyl protease inhibitor pepstatin A had no effect on

azido-E-64 labeling (Figure 3, panel c). These data demonstrate that azido-E-64 targets active Cat cysteine proteases that function at low pH.

Azido-E-64 Selectively Targets Active Cat B in Macrophages. On the basis of the inhibitory activity of azido-E-64 measured by DCG-04 labeling (Figure 2), the molecular weight of the polypeptides labeled by azido-E-64, and their differential sensitivity to CA-074-OMe (Figure 3), the predominant protease targeted by azido-E-64 is most likely Cat B. To confirm that Cat B is indeed the most prominent target of azido-E-64, cell lysates from macrophages treated with or without azido-E-64 were labeled with p-biotin, incubated with streptavidin beads, and analyzed by streptavidin blot as well as by MS (Figures 1 and 4). As shown (Figure 4, panel a), Cat B was specifically recovered on streptavidin beads from azido-E-64 treated cell lysates following p-biotin labeling, as judged by anti-Cat B immunoblot. Moreover, polypeptides captured on streptavidin beads were denatured and then subjected to on-bead digestion with trypsin (Figure 1 and Figure 4, panel b). Tryptic peptides eluted from streptavidin beads were collected and sequenced by electrospray ionization (ESI) MS/MS. Comparison of peptides eluted from streptavidin beads revealed Cat B-specific peptides from azido-E-64 treated cell lysates, which were not present in control samples (Figure 4, panel b, and Supplementary Figure 2). To unequivocally establish Cat B as the major target for azido-E-64, bone marrow-derived macrophages (BMM ϕ s) prepared from wild-type and Cat B-deficient mice were labeled with azido-E-64. The major polypeptides specifically labeled by azido-E-64 in wild-type BMM ϕ s were identical in molecular weight compared to RAW264.7 macrophages (Figures 3 and 4). The major polypeptide labeled by azido-E-64 in wild-type BMM ϕ s is absent from Cat B-deficient BMM ϕ s. Importantly, the Cat B labeled by azido-E-64 in RAW macrophages and BMM ϕ s is the mature active enzyme (\sim 30 kDa), as no labeling of Cat B proform (\sim 39 kDa) was detected (Figures 3 and 4). We conclude that the primary target of azido-E-64 in macrophages is active Cat B.

Visualization of Active Cat B by Fluorescence Microscopy. Having established the specificity of azido-E-64 labeling for the active form of Cat B in macrophages, we investigated whether azido-E-64 labeling could be exploited for visualization of Cat B by fluorescence microscopy. BMM ϕ s were plated on cover-slips,

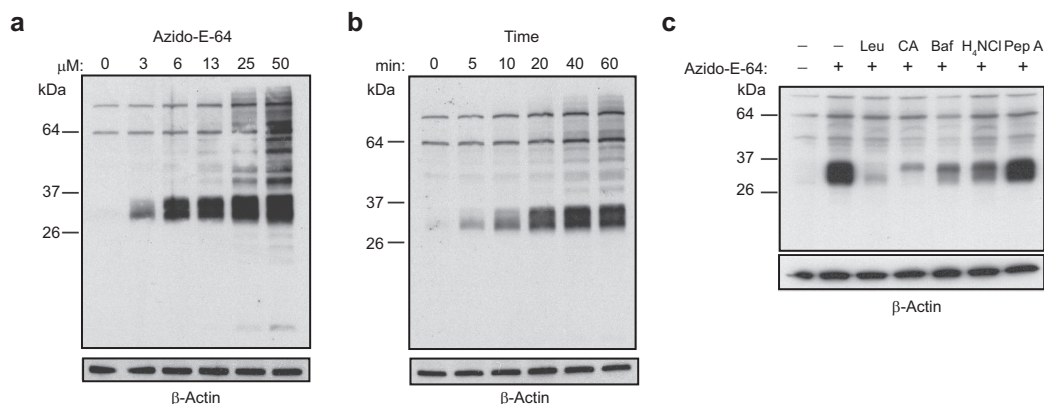


Figure 3. Visualization of Cat cysteine proteases labeled by azido-E-64 in living cells. **a)** Dose-dependent labeling of Cat cysteine proteases in RAW264.7 macrophages treated with azido-E-64. Cells were incubated with increasing concentrations of azido-E-64 for 1 h at 37 °C. Cells were harvested and lysed in the presence of protease inhibitor cocktail, and cell lysates were reacted with p-biotin (200 μ M) for 2 h at 37 °C. Reactions were terminated by acetone precipitation, and proteins were analyzed by streptavidin blotting. **b)** RAW264.7 macrophages were treated with azido-E-64 (20 μ M) for the time indicated. Visualization of azido-E-64 labeled proteins was performed as described above in panel a. **c)** RAW264.7 macrophages were treated with azido-E-64 (20 μ M) and lysosomal protease inhibitors Leupeptin (1 mM), CA-074-OMe (10 μ M), Bafilomycin (1 μ M), H₂NCl (50 mM), or Pepstatin A (10 μ M) for 1 h; harvested; and analyzed for azido-E-64 labeled proteins as described above in panel a.

treated with azido-E-64 or DMSO, fixed, and reacted *in situ* with p-biotin (Figures 1 and 5). The labeled cells were then stained with fluorescently conjugated streptavidin for imaging by spinning-disk confocal microscopy. To visualize endocytic compartments, BMM ϕ s were costained for lysosome-associated membrane glycoprotein 1 (LAMP-1) (Figure 5). Cells treated with azido-E-64 exhibited significant staining with fluorescent streptavidin compared to the DMSO control (Figure 5, top three panels), which demonstrates that azido-E-64 can be used to selectively visualize active Cat B in cells by fluorescence microscopy. The azido-E-64-specific staining resided within LAMP-1⁺ compartments, as expected for the labeling of endocytic proteases, and overlapped with Cat B protein that was visualized by anti-Cat B antibody staining (Figure 5). In contrast, azido-E-64 staining did not exhibit significant overlap with the distribution of protein disulfide isomerase (PDI), an endoplasmic reticulum (ER)-resident protein (Figure 5). These data demonstrate that azido-E-64 specifically labels Cat B in endocytic compartments where the active form of this protease is expected to reside. Similar results were also obtained with RAW264.7 macrophages (Supplementary Figure 3). We also analyzed monocyte-derived M ϕ s from class II MHC-eGFP knock-in mice (28) and observed spe-

cific azido-E-64 labeling that was blocked after pretreatment of cells with leupeptin or bafilomycin (Supplementary Figure 4). Collectively, these results establish that azido-E-64 labeling can be visualized not only by immunoblotting but also by fluorescence microscopy. The visualization of azido-E-64 labeling by fluorescence microscopy provides an additional measure of specificity as well as the means to monitor the activity of these proteases within subcellular compartments at the single-cell level.

Infection of Primary Macrophages with Live *Salmonella* Excludes Active Cat B from SCVs. The ability to specifically visualize active Cat cysteine proteases in living cells presented us with the unique opportunity to address the activity of these endocytic proteases during microbial infection of host cells. Intracellular bacterial pathogens have evolved sophisticated mechanisms to evade destruction by host cells to sustain a productive infection (29). *Salmonella* is an example of a facultative intracellular bacterial pathogen that exploits two type III secretion systems (TTSSs) to infect host cells and cause disease (29). These TTSSs form molecular syringes that inject protein substrates, termed “effectors”, into host cells to modulate cellular pathways for survival and replication (30). Upon invasion of host cells, *Salmonella* is

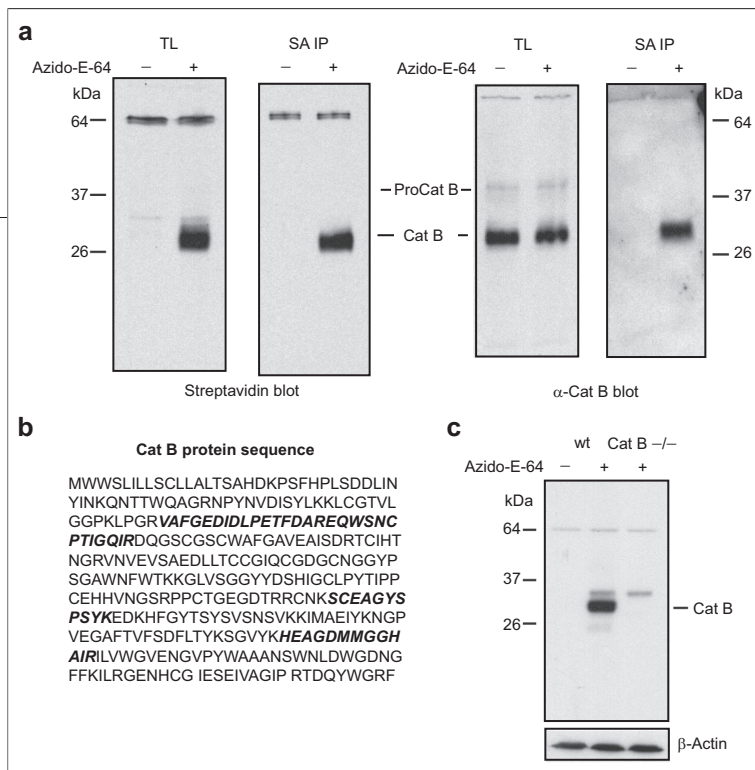


Figure 4. Azido-E-64 selectively labels Cat B in macrophages. **a)** Selective enrichment of Cat B from azido-E-64 treated macrophages after p-biotin labeling and incubation with streptavidin–agarose beads. TL = total lysate; SA IP = streptavidin–agarose immunoprecipitation **b)** Cat B protein sequence. Cat B peptides selectively identified from azido-E-64 treated samples after p-biotin labeling, streptavidin affinity enrichment, on-bead digestion, and MS sequencing are indicated in bold italic type. Cat B peptides recovered (MH⁺): **VAFGEDIDLPETFDAR** (1794.85), **EQWSNCPTIGQIR** (1588.78), **SCEAGYSPSYK** (1248.55), and **HEAGDMMGGHAIR** (1381.61). Representative MS/MS spectra of Cat B peptide is provided in Supplementary Figure 2. **c)** BMM ϕ s from wild-type and Cat B $-/-$ mice were treated with DMSO (–) or 20 μ M azido-E-64 (+). Streptavidin-HRP blots were inactivated with H₂O₂ and probed with anti- β -actin antibody to demonstrate equal levels of protein loading.

thought to inhibit phago-lysosome fusion and resides in an SCV, a unique endocytic compartment (31). The SCV is devoid of oxidative burst enzymes, such as NADPH oxidase (32) and inducible nitric oxide synthase (33) that enable the bacteria to avoid degradation by reactive radical species. Endocytic compartments are also rich in degradative enzymes that include the Cat family of aspartyl, cysteine, and serine proteases, which have been implicated in antigen processing and presentation (34), but their roles in *Salmonella* infection remain unclear.

The Cat families of proteases are synthesized as zymogens that are processed into their mature, active forms upon arrival in endocytic compartments. Here their activities are controlled by subcellular localization, the presence of endogenous inhibitors or activators, and pH (35). Reports that document the association of these endocytic proteases with *Salmonella* in host cells yield a mixed outcome (36). In epithelial cell lines infected with *S. typhimurium*, Cat D does not exhibit significant overlap with SCVs at early time points of infection as

assessed by immunocytochemistry but appears to be recruited at late stages of infection during *Salmonella*-induced filament formation (37–39). Cat D and L were also absent from SCVs in macrophage-like cell lines and were not recruited over the course of the infection (40, 41). In contrast, SCVs in primary BMM ϕ s were shown to associate with Cat L over the course of the *Salmonella* infection (42). Furthermore, sucrose-gradient purification of SCVs from a macrophage cell line recovered proforms of Cat D and L, as demonstrated by immunoblot analysis (43, 44), which suggests that *Salmonella* may reside in an endocytic compartment with inactive proteases. However, the recovery of calnexin, an ER-resident protein, from purified SCVs suggests that proforms of endocytic proteases may be recruited to SCVs prior to their arrival into endocytic compartments (44), possibly through an ER-phagosome fusion pathway (45, 46). While it is clear that avirulent or heat-killed bacteria are targeted to lysosomes for degradation, reports on the interactions between endocytic proteases and live *Salmonella* present conflicting conclusions, particularly the studies that pertain to macrophages. Because the activities of these endocytic proteases are controlled post-translationally, it is imperative to assess whether the proteases associated with intracellular bacteria are indeed enzymatically active.

To determine whether active endocytic proteases are present in SCVs of primary macrophages, we assayed azido-E-64 labeling of active Cat B in *Salmonella*-infected BMM ϕ s. BMM ϕ s plated on cover glass slips were infected with *S. typhimurium* at a multiplicity of infection (MOI) of \sim 100 for 1 h and labeled with azido-E-64 for 30 min. Following fixation of *S. typhimurium*-infected BMM ϕ s, azide-modified polypeptides were reacted with p-biotin and prepared for fluorescence microscopy. Azido-E-64 efficiently labeled LAMP-1⁺ compartments of *Salmonella*-infected BMM ϕ s, but azido-E-64 labeling was completely absent from SCVs as judged by costaining with anti-*Salmonella* antibody (Figure 6). The lack of azido-E-64 labeling in SCVs was reproducible over several independent infections of BMM ϕ s, where multiple *Salmonella*-infected macrophages were analyzed. Furthermore, the segregation of active Cat B from SCVs, as assessed by azido-E-64 labeling, was not a transient phenomenon and persisted over 4 h postinfection (data not shown). The azido-E-64 labeling was also specific for Cat B, as no detectable azido-E-64 labeling was observed in Cat B-deficient BMM ϕ s

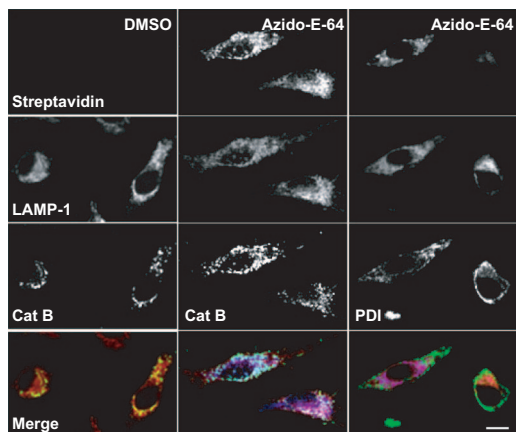


Figure 5. Immunofluorescence analysis of active Cat B in BMMφs labeled with azido-E-64. BMMφs were treated with DMSO or azido-E-64 (20 μ M), fixed, reacted with p-biotin, and stained with AlexaFluor647-conjugated streptavidin. LAMP-1 was visualized by staining with rat anti-mouse LAMP-1 followed by AlexaFluor595-conjugated anti-rat. Cat B or PDI was visualized by staining with rabbit polyclonal sera for Cat B and PDI, respectively, followed by AlexaFluor488-conjugated anti-rabbit. Images were acquired by spinning-disk confocal microscopy. For merged images, AlexaFluor647-conjugated streptavidin (blue) and LAMP-1 (red), Cat B (green), or PDI (green). Scale bar represents 10 μ m.

infected with *S. typhimurium* (Figure 6). Most bacteria were degraded when heat-killed *Salmonella* were used to infect BMMφs, but the bacteria that remained inside LAMP-1⁺ compartments of macrophages colocalized with active Cat B, which demonstrates the exclusion of active Cat B from SCVs required live *Salmonella* (Figure 6).

Exclusion of Active Cat B from Bacteria-Containing Vacuoles Is Specific to Salmonella. To determine whether the exclusion of active Cat B from bacteria-containing vacuoles in BMMφs was specific for *S. typhimurium* or simply a general property of intracellular bacteria, we generated monomeric-red fluorescent protein (mRFP)-labeled *S. typhimurium* and nonpathogenic *Escherichia coli* and compared their localization within BMMφs with that of active Cat B, as judged by azido-E-64 labeling. SCVs in BMMφs infected with mRFP *S. typhimurium* were also devoid of active Cat B (Figure 7), which is consistent with our results using anti-*Salmonella* polyclonal sera (Figure 6). In contrast, internalized mRFP *E. coli* targeted to LAMP-1⁺ compartments were mostly

degraded and colocalized with active Cat B (Figure 7). Together, these observations suggest that the exclusion of active Cat B from bacteria-containing vacuoles in BMMφs is specific to live *Salmonella* and not a general property exhibited by bacteria inside host cells (Figure 7).

The development of mechanism-based probes has yielded new chemical tools for dissecting the function of protein superfamilies in biology not accessible by genomics and proteomics (1). Neither transcriptional profiling nor immunocytochemical detection of the polypeptides can uncover regulatory non-template-encoded phenomena, such as ionic environment, or the presence of modulators of enzyme activity, either endogenous or specified by pathogens. These probes have enabled the profiling of enzyme activities in complex mixtures for several enzyme families and identified new enzymatic activities not predicted by bioinformatics (1). While it is clear that mechanism-based probes are making a contribution to biology, the application of mechanism-based probes to living cells is still challenging. Here we demonstrate that a single-mechanism-based probe, azido-E-64, not only can be used to visualize Cat B in living cells but also allows affinity enrichment for MS-based proteomic analysis with phosphine reagents *via* the Staudinger ligation (Figure 1).

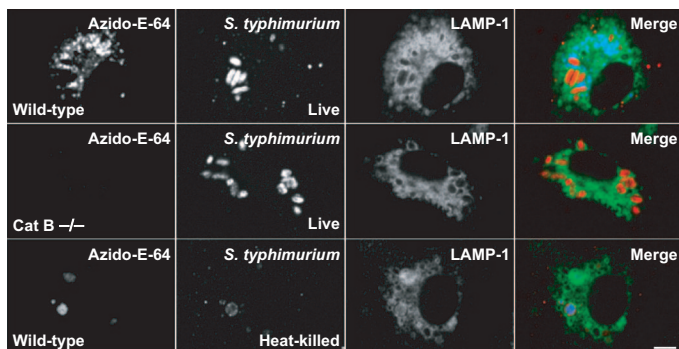


Figure 6. Visualization of active Cat B in *S. typhimurium*-infected BMMφs. BMMφs were infected with live or heat-killed *S. typhimurium* (MOI = 100) for 1 h, washed and labeled with azido-E-64 (20 μ M), and visualized as described in Figure 6. *S. typhimurium* rabbit was visualized by staining with anti-*S. typhimurium* polyclonal sera, followed by AlexaFluor568-conjugated anti-rabbit. LAMP-1 was visualized by staining with rat anti-mouse LAMP-1 followed by AlexaFluor488-conjugated anti-rat. Images were acquired by spinning-disk confocal microscopy. For merged images, AlexaFluor647-conjugated streptavidin (blue), *S. typhimurium* (red), and LAMP-1 (green). Scale bar represents 5 μ m.

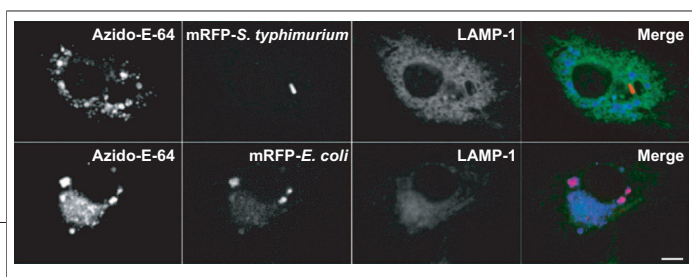


Figure 7. The exclusion of active Cat B from bacteria-containing vacuoles is specific to *S. typhimurium*-infected BMMø. BMMø were infected with mRFP *S. typhimurium* or mRFP *E. coli* (MOI = 100) for 1 h, washed and labeled with azido-E-64 (20 μ M), and visualized as described in Figure 6. LAMP-1 was visualized by staining with rat anti-mouse LAMP-1 followed by AlexaFluor488-conjugated anti-rat. Images were acquired by spinning-disk confocal microscopy. For merged images, AlexaFluor647-conjugated streptavidin (blue), mRFP *S. typhimurium* and mRFP *E. coli* (red), and LAMP-1 (green). Scale bar represents 5 μ m.

Many pathogens infect host cells through endocytic pathways to co-opt host nutrient and biosynthetic machinery for survival and replication (29). Within the endocytic compartment of host cells, pathogens face a broad arsenal of host factors and enzymes that are capable of destroying them. As a consequence, many successful intracellular pathogens have evolved sophisticated mechanisms to evade destruction. The interactions between pathogens and host cells are very complex and require new tools to dissect mechanisms of immune evasion. Cell-permeable mechanism-based probes provide new opportunities to investigate host-pathogen interactions not possible with genomics or proteomics. *Salmonella* in particular traffics into a unique endocytic compartment, the *Salmonella*-containing vacuole, which allows the bacterium to survive and replicate in epithelial cells and macrophages. Unlike *Mycobacterium tuberculosis*, which inhibits vacuole acidification (47), *Salmonella* actively decreases endosomal pH (~4.0–6.0) to enhance intracellular survival (48), conditions that are compatible with the pH optima of endocytic proteases (35). The Cat cysteine proteases are the most abundant proteolytic enzymes in endocytic compartments and are involved in many biological processes, ranging from bone resorption, tumor metastasis, antigen presentation, and viral infection, as well as transcription regulation or apoptosis (49). While a few studies suggest that these endocytic proteases are excluded (40, 41) or differentially processed (43, 44) in SCVs of macrophage-like cell lines, Cat cysteine proteases do associate with SCVs of primary macrophages (42). As a result, the activity of Cat cysteine proteases during bacterial infections is unclear.

To monitor the activities of the Cat cysteine proteases during *Salmonella* infection, we utilized azido-E-64 to visualize Cat B in macrophages after bacterial infection. Following infection of BMMø with *Salmonella*, labeling with azido-E-64 showed that active Cat B was absent from SCVs at early time points of infection, a trait that persisted for several hours after internalization of bacte-

ria. The exclusion of active Cat B from bacteria-containing vacuoles was specific to live *Salmonella*, as the distribution of heat-killed *Salmonella* or *E. coli* within BMMø completely overlapped with azido-E-64 labeling. These data are consistent with previous immunofluorescence studies that show Cat D and L polypeptides to be absent from SCVs in macrophage cell lines (40, 41). Moreover, the lack of azido-E-64 labeling within SCVs of primary BMMø is in agreement with the association of SCVs with inactive proforms of endocytic proteases recovered by sucrose-gradient purification (43, 44). These studies suggest *Salmonella* may encode specific gene products, possibly secreted into SCVs that inhibit endocytic protease activity or interfere with the conversion of inactive proteases into their mature active enzymes. In addition to modification of phago-lysosome fusion (50), *Salmonella* may also interfere with endocytic protease activity directly to survive and replicate inside macrophages. Future studies with *Salmonella* mutants should enable the dissection of specific genes that are responsible for the exclusion or inhibition of active endocytic proteases from SCVs.

In conclusion, the two-step labeling approach reported here provides a nonradioactive method to evaluate the activity of the Cat cysteine proteases in living cells and also allows affinity enrichment of labeled proteins for MS-based proteomic analysis. In addition, the ability to perform fluorescence microscopy with mechanism-based probes allows the visualization of enzyme activities within subcellular compartments of individual cells, the biochemical evaluation of which is often problematic because the organelles concerned are often difficult to obtain in pure form, particularly for complex series of events such as infection with a pathogen. Thus, the application of this chemical approach to other intracellular pathogens that reside in endocytic compartments should provide further insight into the mechanisms by which these pathogens manage to evade degradation by host cells. The extension of this approach to other mechanism-based probes should broaden their utility for studies in living cells and animal models. Collectively, cell-permeable mechanism-based probes should provide a powerful means to measure changes in enzyme activity *in vivo* and afford the possibility of examining biological activities that are controlled post-translationally.

METHODS

Cell Culture. RAW264.7 macrophages were cultured in DMEM supplemented with 10% fetal calf serum (FCS), 100 units mL⁻¹ penicillin, and 0.1 mg mL⁻¹ streptomycin, and cells were incubated in a 5% CO₂ humidified incubator at 37 °C. Bone marrow cells were harvested by flushing the femur and tibia of wild-type C57BL/6 mice using a 25-gauge needle. For BMMø, 2 × 10⁶ bone marrow cells were plated on 100 × 20 mm tissue culture plates in complete DMEM media supplemented with 10% FCS, 100 units mL⁻¹ penicillin, 0.1 mg mL⁻¹ streptomycin, and macrophage-colony stimulating factor derived from the supernatant of L929 cells. Fresh media was replenished on days 4 and 7.

Treatment of Mø with Inhibitors. Mø were seeded ~2.0 × 10⁶ cells per well in six-well polystyrene tissue culture plates with 1 mL of media. Cells were then supplemented with protease inhibitors (E-64 or azido-E-64) by adding the appropriate volume from 10 mM DMSO stock solutions. After incubation of cells for 37 °C for the times indicated, cells were washed twice with phosphate-buffered saline (PBS) and harvested.

DCG-04 Labeling of Cell Lysates. Cell pellets from ~2.0 × 10⁶ cells were resuspended in 100 µL of ice-cold pH 5.5 lysis buffer (0.1% Triton x-100, 0.5% CHAPS, 50 mM citrate, 5 mM dithiothreitol (DTT), pH 5.5) and centrifuged at 4 °C for 10 min at 20,000g. The supernatant (cell lysate) was collected, and the protein concentration was determined by Bradford assay (BioRad). Cell lysates (50 µg) were reacted with 5 µM of DCG-04 with or without (E-64 or azido-E-64) for 2 h at 37 °C in a final reaction volume of 50 µL. The reactions were terminated by addition of ice-cold acetone, incubated at -20 °C for 20 min, and centrifuged at 4 °C for 10 min at 20,000g to precipitate proteins. The supernatant was discarded, and the protein pellet was resuspended in SDS-protein loading buffer with 2-mercaptoethanol.

P-Biotin Labeling of Cell Lysates. Cell pellets from ~2.0 × 10⁶ cells were resuspended in 100 µL of ice-cold NP-40 lysis buffer (0.5% NP-40, 50 mM Tris, 5 mM MgCl₂, pH 7.4) with Complete Mini protease inhibitor cocktail and centrifuged at 4 °C for 10 min at 20,000g. The supernatant (cell lysate) was collected, and the protein concentration was determined by Bradford assay (BioRad). Staudinger ligation labeling of azide-modified proteins was performed by incubation of cell lysates (50 µg) with 250 µM p-biotin (5 mM in DMSO stock) and 5 mM DTT in total reaction volume of 50 µL for 2 h at 37 °C. The reactions were terminated by addition of ice-cold acetone (1 mL), and the product was incubated at -20 °C for 20 min and centrifuged at 4 °C for 10 min at 20,000g to precipitate proteins. The supernatant was decanted, and the protein pellet was resuspended in SDS-protein loading buffer with 2-mercaptoethanol.

Immunoblotting. P-biotin (~20 µg) and DCG-04 (~10 µg) labeled proteins were separated by SDS-PAGE (12.5% gel) and transferred to a polyvinylidene fluoride membrane, and the membrane was blocked with 5% nonfat dried milk in PBS/Tween 20 (PBST) (PBS, 0.1% Tween 20, pH 7.4) overnight at 4 °C or 1 h at RT. The membrane was washed with PBST (3 × 10 mL) and incubated with streptavidin-HRP (1:5000 in PBST) for 1 h, washed with PBST (3 × 25 mL), and developed using Western Lighting Chemiluminescence Reagent Plus (Perkin Elmer). To demonstrate equal levels of protein loading, streptavidin-HRP blots were inactivated with H₂O₂ and probed with anti-β-actin antibody followed by rabbit anti-mouse HRP. Anti-Cat B immunoblots were performed with rabbit anti-Cat B polyclonal and then probed with mouse anti-rabbit HRP.

Streptavidin Affinity Enrichment and Proteomic Analysis of Biotinylated Proteins. Cell lysates (5 mg) from macrophages treated with or without azido-E-64 were reacted with 25 µM p-biotin (5 mM in DMSO stock) and 5 mM DTT in a total reaction volume of 5 mL for 2 h at 37 °C with rocking. The reactions were

terminated by addition of ice-cold acetone (40 mL), and the resulting product was incubated at -20 °C for 20 min and centrifuged at 4 °C for 10 min at 4000g to precipitate proteins. The supernatant was decanted, and the protein pellet was resuspended in 5 mL of 0.2% SDS (50 mM Tris, pH 7.4) with sonication. Proteins were precipitated again with ice-cold acetone (40 mL) and resuspended in 5 mL of 0.2% SDS (50 mM Tris, pH 7.4) with sonication to remove any residue p-biotin. Pre-washed streptavidin-agarose beads (250 µL) were then added to protein lysates (1 mg mL⁻¹) and allowed to incubate at 4 °C for 1 h with rocking. Streptavidin-agarose beads were centrifuged for 5 min at 4000g at 4 °C and washed with 10 mL of 0.2% SDS (50 mM Tris, pH 7.4) four times. One-fifth of the streptavidin-agarose beads (~50 µL of slurry) were removed, resuspended in SDS-protein loading buffer with 2-mercaptoethanol, and analyzed by immunoblotting as described above. The remaining streptavidin-agarose beads (~200 µL) were washed with 1 mL of 50 mM NH₄CO₃, pH 7.4, twice, denatured, and reduced with 500 µL of 6 M urea, 10 mM tris(2-carboxyethyl)phosphine, 50 mM NH₄CO₃, pH 7.4, for 30 min at RT with rocking. Iodoacetamide from a stock solution of 1 M was added to streptavidin-agarose beads to give a final concentration of 20 mM iodoacetamide and allowed to react for 30 min at RT with rocking in the dark. The beads were then washed with 1 mL of 50 mM NH₄CO₃, pH 7.4, three times, resuspended in 500 µL of 50 mM NH₄CO₃, pH 7.4, with 2 µg of porcine trypsin, and allowed to digest overnight at 37 °C with rocking. The streptavidin beads were centrifuged for 1 min at 4000g, and the supernatants were collected and concentrated by SpeedVac. Recovered peptides were analyzed by reverse-phase LC ESI-MS using Waters nanoAquity-UPLC coupled to a Thermo LTQ linear ion-trap mass spectrometer. MS/MS spectra were searched by SEQUEST against the NCBI database (nr.fasta.hr 6/27/2006). SEQUEST results were analyzed with Bioworks Browser 3.2 and filtered with the following criteria: different peptides; minimum cross correlation coefficients (1, 2, 3 charge states) of 1.50, 2.00, 2.50; number different peptides of 2 per protein and Sp preliminary score of 500.

Fluorescence Microscopy. Mø (~2.0 × 10⁵ per well) were seeded in an eight-well chambered Lab-Tek II cover glass slides 16 h before analysis. Mø on cover glass slides were treated with DMSO or 20 µM azido-E-64 in 250 µL of media for 30 min at 37 °C, washed with PBS, and fixed with ice-cold methanol for 5 min at 4 °C. After fixation, cells were washed with PBS and reacted with 250 µM p-biotin in 200 µL of PBS for 2 h at 37 °C. The cells were then washed with PBS and blocked with 10% bovine serum albumin (BSA) in PBS for 16 h at 4 °C or 1 h at RT. The cells were stained with primary antibodies dissolved in permeabilization buffer (10% BSA, 0.5% saponin in PBS) for 1 h at RT, washed with permeabilization buffer (3 × 250 µL), and stained with fluorescently conjugated secondary antibodies or streptavidin dissolved in permeabilization buffer for 1 h at RT in the dark. Slides were washed with permeabilization buffer (2 × 250 µL), PBS, and samples were mounted with Fluoromount G (Southern Biotech) and analyzed using a Perkin-Elmer spinning disk confocal microscope with UltraView software.

Bacterial Infection of BMMø. BMMø (~2.0 × 10⁵ per well) were seeded in eight-well chambered Lab-Tek II cover glass slides the 16 h before analysis and infected with bacteria (MOI = ~100), centrifuged at 500g for 5 min, and incubated for 1 h at 37 °C. The cells were washed with PBS (250 µL × 2) and labeled with 20 µM azido-E-64 for 20 min at 37 °C. BMMø were fixed and stained for immunofluorescence and analyzed as described above.

Acknowledgments: H.C.H. acknowledges the Damon Runyon Cancer Research Foundation for a postdoctoral fellowship. J.L. thanks the Gulbenkian Ph.D. Program in Biomedicine and the Portuguese Science and Technology Foundation for a predoctoral fel-

lowship. A.W.M.v.d.W. acknowledges Philip Morris USA, Inc., for support. R.M. was supported by a Boehringer Ingelheim Fonds predoctoral fellowship. A.M.P. acknowledges the National Science Foundation for a graduate fellowship. Y.-M.K. acknowledges the Leukemia and Lymphoma Society for a postdoctoral fellowship. This work was funded by NIH Grants to M.N.S. (AI055962) and H.L.P. (5RO1A1034893-13).

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Evans, M. J., and Cravatt, B. F. (2006) Mechanism-based profiling of enzyme families, *Chem. Rev.* **106**, 3279–3301.
- Liu, Y., Patricelli, M. P., and Cravatt, B. F. (1999) Activity-based protein profiling: the serine hydrolases, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14694–14699.
- Borodovsky, A., Ovaa, H., Kolli, N., Gan-Erdene, T., Wilkinson, K. D., Ploegh, H. L., and Kessler, B. M. (2002) Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family, *Chem. Biol.* **9**, 1149–1159.
- Greenbaum, D., Medzihradszky, K. F., Burlingame, A., and Bogoy, M. (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools, *Chem. Biol.* **7**, 569–581.
- Kato, D., Boatright, K. M., Berger, A. B., Nazif, T., Blum, G., Ryan, C., Chehade, K. A., Salvesen, G. S., and Bogoy, M. (2005) Activity-based probes that target diverse cysteine protease families, *Nat. Chem. Biol.* **1**, 33–38.
- Chan, E. W., Chattopadhyaya, S., Panicker, R. C., Huang, X., and Yao, S. Q. (2004) Developing photoactive affinity probes for proteomic profiling: hydroxamate-based probes for metalloproteases, *J. Am. Chem. Soc.* **126**, 14435–14446.
- Saghatelian, A., Jessani, N., Joseph, A., Humphrey, M., and Cravatt, B. F. (2004) Activity-based probes for the proteomic profiling of metalloproteases, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10000–10005.
- Raskin, D. M., Seshadri, R., Pukatzki, S. U., and Mekalanos, J. J. (2006) Bacterial genomics and pathogen evolution, *Cell* **124**, 703–714.
- Greenbaum, D. C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K. F., Engel, J., DeRisi, J., Holder, A. A., and Bogoy, M. (2002) A role for the protease falcipain 1 in host cell invasion by the human malaria parasite, *Science* **298**, 2002–2006.
- Kattenhorn, L. M., Korbel, G. A., Kessler, B. M., Spooner, E., and Ploegh, H. L. (2005) A deubiquitinating enzyme encoded by HSV-1 belongs to a family of cysteine proteases that is conserved across the family *Herpesviridae*, *Mol. Cell* **19**, 547–557.
- Schlieker, C., Korbel, G. A., Kattenhorn, L. M., and Ploegh, H. L. (2005) A deubiquitinating activity is conserved in the large tegument protein of the *Herpesviridae*, *J. Virol.* **79**, 15582–15585.
- Misaghi, S., Balsara, Z. B., Catic, A., Spooner, E., Ploegh, H. L., and Stambach, M. N. (2006) Chlamydia trachomatis-derived deubiquitinating enzymes in mammalian cells during infection, *Mol. Microbiol.* **61**, 142–150.
- Artavanis-Tsakonas, K., Misaghi, S., Comeaux, C. A., Catic, A., Spooner, E., Duraisingh, M. T., and Ploegh, H. L. (2006) Identification by functional proteomics of a deubiquitinating/deNeddylating enzyme in *Plasmodium falciparum*, *Mol. Microbiol.*, in press.
- Lennon-Dumenil, A. M., Bakker, A. H., Maehr, R., Fiebiger, E., Overkleeft, H. S., Roseblatt, M., Ploegh, H. L., and Lagaudriere-Gesbert, C. (2002) Analysis of protease activity in live antigen-presenting cells shows regulation of the phagosomal proteolytic contents during dendritic cell activation, *J. Exp. Med.* **196**, 529–540.
- Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K. F., and Bogoy, M. (2002) Chemical approaches for functionally probing the proteome, *Mol. Cell. Proteomics* **1**, 60–68.
- Joyce, J. A., Baruch, A., Chehade, K., Meyer-Morse, N., Giraudo, E., Tsai, F. Y., Greenbaum, D. C., Hager, J. H., Bogoy, M., and Hanahan, D. (2004) Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis, *Cancer Cell* **5**, 443–453.
- Blum, G., Mullins, S. R., Keren, K., Fonovic, M., Jedeszko, C., Rice, M. J., Sloane, B. F., and Bogoy, M. (2005) Dynamic imaging of protease activity with fluorescently quenched activity-based probes, *Nat. Chem. Biol.* **1**, 203–209.
- Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction, *Science* **287**, 2007–2010.
- Wang, Q., Chan, T. R., Hilgraf, R., Fokin, V. V., Sharpless, K. B., and Finn, M. G. (2003) Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition, *J. Am. Chem. Soc.* **125**, 3192–3193.
- Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003) Activity-based protein profiling, *in vivo* using a copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition, *J. Am. Chem. Soc.* **125**, 4686–4687.
- Ovaa, H., Van Swieten, P. F., Kessler, B. M., Leeuwenburgh, M. A., Fiebiger, E., Van Den Nieuwendijk, A. M., Galardy, P. J., Van Der Marel, G. A., Ploegh, H. L., and Overkleeft, H. S. (2003) Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy, *Angew. Chem., Int. Ed.* **42**, 3626–3629.
- Sprung, R., Nandi, A., Chen, Y., Kim, S. C., Barma, D., Falck, J. R., and Zhao, Y. (2005) Tagging-via-substrate strategy for probing O-GlcNAc modified proteins, *J. Proteome Res.* **4**, 950–957.
- Sieber, S. A., Niessen, S., Hoover, H. S., and Cravatt, B. F. (2006) Proteomic profiling of metalloprotease activities with cocktails of active-site probes, *Nat. Chem. Biol.* **2**, 274–281.
- Speers, A. E., and Cravatt, B. F. (2005) A tandem orthogonal proteolysis strategy for high-content chemical proteomics, *J. Am. Chem. Soc.* **127**, 10018–10019.
- Hanada, K., Tamai, M., Yamaguchi, M., Ohmura, S., Sawada, J., and Tanaka, I. (1978) Isolation and characterization of E-64, a new thiol protease inhibitor, *Agric. Biol. Chem.* **42**, 523–528.
- Powers, J. C., Asgian, J. L., Ekici, O. D., and James, K. E. (2002) Irreversible inhibitors of serine, cysteine, and threonine proteases, *Chem. Rev.* **102**, 4639–4750.
- Buttle, D. J., Murata, M., Knight, C. G., and Barrett, A. J. (1992) CA074 methyl ester: a proinhibitor for intracellular cathepsin B, *Arch. Biochem. Biophys.* **299**, 377–380.
- Boes, M., Cerny, J., Massol, R., Op den Brouw, M., Kirchhausen, T., Chen, J., and Ploegh, H. L. (2002) T-cell engagement of dendritic cells rapidly rearranges MHC class II transport, *Nature* **418**, 983–988.
- Monack, D. M., Mueller, A., and Falkow, S. (2004) Persistent bacterial infections: the interface of the pathogen and the host immune system, *Nat. Rev. Microbiol.* **2**, 747–765.
- Galan, J. E. (2001) *Salmonella* interactions with host cells: type III secretion at work, *Annu. Rev. Cell Dev. Biol.* **17**, 53–86.
- Brumell, J. H., and Grinstein, S. (2004) *Salmonella* redirects phagosomal maturation, *Curr. Opin. Microbiol.* **7**, 78–84.
- Vazquez-Torres, A., Xu, Y., Jones-Carson, J., Holden, D. W., Lucia, S. M., Dinauer, M. C., Mastroeni, P., and Fang, F. C. (2000) *Salmonella* pathogenicity island 2-dependent evasion of the phagocyte NADPH oxidase, *Science* **287**, 1655–1658.
- Chakravorty, D., Hansen-Wester, I., and Hensel, M. (2002) *Salmonella* pathogenicity island 2 mediates protection of intracellular *Salmonella* from reactive nitrogen intermediates, *J. Exp. Med.* **195**, 1155–1166.
- Honey, K., and Rudensky, A. Y. (2003) Lysosomal cysteine proteases regulate antigen presentation, *Nat. Rev. Immunol.* **3**, 472–482.
- Turk, V., Turk, B., and Turk, D. (2001) Lysosomal cysteine proteases: facts and opportunities, *EMBO J.* **20**, 4629–4633.
- Holden, D. W. (2002) Trafficking of the *Salmonella* vacuole in macrophages, *Traffic* **3**, 161–169.

37. Brumell, J. H., Tang, P., Mills, S. D., and Finlay, B. B. (2001) Characterization of *Salmonella*-induced filaments (Sifs) reveals a delayed interaction between *Salmonella*-containing vacuoles and late endocytic compartments, *Traffic* 2, 643–653.
38. Garcia-del Portillo, F., and Finlay, B. B. (1995) Targeting of *Salmonella typhimurium* to vesicles containing lysosomal membrane glycoproteins bypasses compartments with mannose 6-phosphate receptors, *J. Cell Biol.* 129, 81–97.
39. Meresse, S., Steele-Mortimer, O., Finlay, B. B., and Gorvel, J. P. (1999) The rab7 GTPase controls the maturation of *Salmonella typhimurium*-containing vacuoles in HeLa cells, *EMBO J.* 18, 4394–4403.
40. Garvis, S. G., Beuzon, C. R., and Holden, D. W. (2001) A role for the PhoP/Q regulon in inhibition of fusion between lysosomes and *Salmonella*-containing vacuoles in macrophages, *Cell. Microbiol.* 3, 731–744.
41. Rathman, M., Barker, L. P., and Falkow, S. (1997) The unique trafficking pattern of *Salmonella typhimurium*-containing phagosomes in murine macrophages is independent of the mechanism of bacterial entry, *Infect. Immun.* 65, 1475–1485.
42. Oh, Y. K., Alpuche-Aranda, C., Berthiaume, E., Jinks, T., Miller, S. I., and Swanson, J. A. (1996) Rapid and complete fusion of macrophage lysosomes with phagosomes containing *Salmonella typhimurium*, *Infect. Immun.* 64, 3877–3883.
43. Hashim, S., Mukherjee, K., Raj, M., Basu, S. K., and Mukhopadhyay, A. (2000) Live *Salmonella* modulate expression of Rab proteins to persist in a specialized compartment and escape transport to lysosomes, *J. Biol. Chem.* 275, 16281–16288.
44. Mills, S. D., and Finlay, B. B. (1998) Isolation and characterization of *Salmonella typhimurium* and *Yersinia pseudotuberculosis*-containing phagosomes from infected mouse macrophages: *Y. pseudotuberculosis* traffics to terminal lysosomes where they are degraded, *Eur. J. Cell Biol.* 77, 35–47.
45. Jutras, I., and Desjardins, M. (2005) Phagocytosis: at the crossroads of innate and adaptive immunity, *Annu. Rev. Cell Dev. Biol.* 21, 511–527.
46. Touret, N., Paroutis, P., and Grinstein, S. (2005) The nature of the phagosomal membrane: endoplasmic reticulum versus plasma-membrane, *J. Leukocyte Biol.* 77, 878–885.
47. Sturgill-Koszycki, S., Schlesinger, P. H., Chakraborty, P., Haddix, P. L., Collins, H. L., Fok, A. K., Allen, R. D., Gluck, S. L., Heuser, J., and Russell, D. G. (1994) Lack of acidification in *Mycobacterium phagosomes* produced by exclusion of the vesicular proton-ATPase, *Science* 263, 678–681.
48. Coombes, B. K., Brown, N. F., Valdez, Y., Brumell, J. H., and Finlay, B. B. (2004) Expression and secretion of *Salmonella* pathogenicity island-2 virulence genes in response to acidification exhibit differential requirements of a functional type III secretion apparatus and SsaL, *J. Biol. Chem.* 279, 49804–49815.
49. Lecaille, F., Kaleta, J., and Bromme, D. (2002) Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design, *Chem. Rev.* 102, 4459–4488.
50. Buchmeier, N. A., and Heffron, F. (1991) Inhibition of macrophage phagosome-lysosome fusion by *Salmonella typhimurium*, *Infect. Immun.* 59, 2232–2238.

EDITOR-IN-CHIEF

Laura L. Kiessling
University of Wisconsin, Madison

BOARD OF EDITORS

Jennifer A. Doudna
University of California, Berkeley

Kai Johnsson
Ecole Polytechnique Fédérale de Lausanne

Anna K. Mapp
University of Michigan, Ann Arbor

Michael A. Marletta
University of California, Berkeley

Peter H. Seeberger
Eidgenössische Technische Hochschule

James R. Williamson
The Scripps Research Institute

EDITORIAL ADVISORY BOARD

Carolyn R. Bertozzi
University of California, Berkeley

Brian T. Chait
Rockefeller University

Tim Clackson
ARIAD Pharmaceuticals, Inc.

Jon C. Clardy
Harvard Medical School

Benjamin F. Cravatt
The Scripps Research Institute

Peter B. Dervan
California Institute of Technology

Rebecca W. Heald
University of California, Berkeley

Linda C. Hsieh-Wilson
California Institute of Technology

Tony Hunter
Salk Institute

Stephen C. Kowalczykowski
University of California, Davis

Richard H. Kramer
University of California, Berkeley

Thomas V. O'Halloran
Northwestern University

Hiroyuki Osada
RIKEN

Anna M. Pyle
Yale University

Ronald T. Raines
University of Wisconsin, Madison

Charles Sawyers
University of California, Los Angeles

Stuart L. Schreiber
Harvard University

Peter G. Schultz
The Scripps Research Institute

Michael P. Sheetz
Columbia University

H. Ulrich Stilz
Sanofi-Aventis, Frankfurt

Christopher T. Walsh
Harvard Medical School

A Year of Firsts for ACS Chemical Biology and the ACS

ACS Chemical Biology (ACS CB) completes its first year of publication with this issue. And what a year it has been! The ready acceptance and extraordinary usage of the journal by the community of scientists, readers, authors, and librarians have been breathtaking. When we commenced upon this new venture at the beginning of the year, our vision for the journal was to foster substantive collaboration between biologists and chemists. This first year has seen great collaboration and community-building, and it remains our ever-defining goal to further collaborate with and enhance the community as we go forward.

In launching a new web and print publication in chemical biology, the ACS made a commitment to bring together a diverse group of scientists to cover the range of subdisciplines in biology and chemistry. In doing so, we have created an innovative, “new” publication offering interactive features that engage the chemical biology research community and facilitate international communication across the various disciplines of chemistry and biology. In this, we have gone beyond our traditional journal model to offer such interactive features as the popular “Ask the Expert” and “Chemical Biology WIKI”. We produce monthly podcasts for the scientists who prefer to hear about the advances in chemical biology. To enhance accessibility of our content to a broad scientific audience, we help authors who are reporting macromolecular structures to develop 3D interactive figures (web-enhanced objects). On top of all of this, we cover other relevant topics such as science policy, funding, and education issues—issues in which the scientific community we serve express the utmost interest.

As evidenced through coverage by the ISI Web of Knowledge, Medline, and the Chemical Abstracts Service, ACS CB is meeting its mission to provide a much-needed forum for innovation in the chemical and biological sciences.

As we launched the journal, the ACS was also busy implementing changes in its manuscript submission and production systems. We and other journals use the new ACS Paragon Plus Environment to provide a manuscript submission and peer-review system with the high level of quality our authors and reviewers expect from the ACS, while also giving us the flexibility to meet future demands. The ACS Paragon Plus Environment is powered by ScholarOne Manuscript Central, which is used by more than 1400 journals worldwide. On the production front, ACS CB was the first ACS journal to use XML, also known as extensible markup language. XML is a flexible way to create common information formats and share both the format and the data on the Internet, intranets, and elsewhere. We use the flexibility of XML to deliver our content to our web pages and link it to related material within ACS publications and elsewhere.

ACS CB, like all ACS journals, is committed to presenting current, compelling issues and providing the highest-quality venue in which to publish. The quality of information and research that we provide you is a direct result of the commitment and leadership of our Editor-in-Chief, Laura Kiessling, and members of our Board of Editors: Jennifer Doudna, Kai Johnsson, Anna Mapp, Michael Marletta, Peter Seeberger, and James Williamson. We are also indebted to a truly wonderful Editorial Advisory Board that represents the broad chemical and biological communities.

Editor's
LETTER

We are especially indebted to the leadership of the ACS Executive Director and CEO, Madeleine Jacobs; the ACS Publications President, Robert Bovenschulte; and the Senior Vice President of the Journals Publishing Group, Brian Crawford, for their decision to launch a new, web-centric publication in the emerging field of chemical biology for the ACS. *ACS CB* staff, including Managing Editor Sarah Tegen, Art Director Theresa Dubé, Webmaster Scott Bell, Editorial Assistant Elizabeth Betts, Special Projects Liaison Grace Miller, and Journals Editing Manager Diane Needham and her team, have been instrumental in creating each of these first 12 issues. A special thanks goes to Marketing Manager Jonathan Morgan and the ACS Sales and Marketing team for ensuring the widespread distribution and online access in the institutional marketplace worldwide.

As Laura Kiessling wrote in the inaugural issue of *ACS CB*, "Our vision of *ACS Chemical Biology* is that it will be nimble—it will respond to new scientific directions. We are interested in innovative ideas about how we can best serve our readership and the broader scientific community." Just as Laura invited your input in that first issue, we continue to invite your input and involvement as we go forward into a new year of publication.

Thank you for participating in this bold, new venture by the ACS. You have made *ACS CB* the success it already is.



Evelyn Jabri
Executive Editor

Cryptophycins: Cryptic No Longer

Blue-green algae produce a large class of peptolides, or compounds with peptide and macrolide characteristics, called cryptophycins. Some cryptophycins have anticancer activity. However, toxicity issues and the lack of large-scale fermentation methods for their generation necessitate development of other methods for the efficient synthesis of molecules of this class. Now, Magarvey *et al.* (p 766 and Point of View p 747) present a detailed characterization and analysis of the cryptophycin biosynthetic pathway.

Cryptophycin structures suggest that the compounds are assembled by a mixed polyketide synthase/nonribosomal peptide synthetase gene cluster. A comparative secondary metabolomic analysis based on related gene clusters from other blue-green algae led to the identification and subsequent cloning, sequencing, and analysis of an ~40 kb cluster encompassing the cryptophycin biosynthetic genes. In addition, biosynthetic precursors were used to characterize the source of each biosynthetic unit and to assess the versatility of the enzymes involved in cryptophycin assembly. Finally, two key enzymes in the pathway were identified and characterized, and this led to their use in a novel chemoenzymatic assembly of cryptophycin 2. Characterization of the biosynthetic pathway of these intriguing compounds provides important chemical, biological, and bioinformatic tools for the generation of known and novel cryptophycins.

Anti-Diabetics by Design

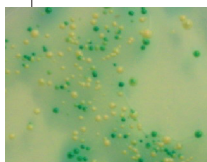
The phosphoinositide 3-phosphatase PTEN is involved in several important cellular processes, including regulation of insulin signaling. Inhibition of PTEN activity leads to increased glucose uptake, an indication that PTEN could be an attractive drug target for diabetes. Rosivatz *et al.* (p 780) report the generation and biological evaluation of specific small-molecule inhibitors of PTEN.

On the basis of the landscape of the PTEN active site and the general phosphatase inhibitory activity of vanadate-based compounds, the authors designed and synthesized a range of vanadates and bis-peroxovanadates. Of eight compounds tested for their ability to inhibit recombinant PTEN activity, the 3-hydroxypicolinate vanadium(IV) complex VO-OHpic was the most potent and specific PTEN inhibitor. With VO-OHpic, they explored the effects of PTEN inhibition in cells. Cells exposed to VO-OHpic exhibited increased phosphorylation of the kinase Akt, translocation of phospho-Akt to the plasma membrane, increased PtdIns(3,4,5)P3 levels, reduced activity of the Akt-dependent transcription factor FoxO3a, and increased glucose uptake. Specific small-molecule PTEN inhibitors such as VO-OHpic will be useful probes for investigating the biological roles and potential therapeutic relevance of PTEN.

Editing Messenger RNA Messages

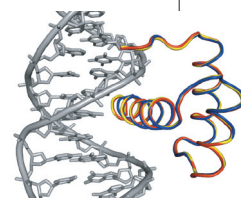
Enzymes called adenosine deaminases that act on RNA (ADARs) catalyze the deamination of adenosine to generate inosine. Because inosine is decoded as guanosine, this RNA editing process can change the message in the RNA strand and lead to structural diversity in the resulting protein. To probe the biochemical and structural basis for ADAR activity, Pokharel and Beal (p 761) have developed a screen that enables the discovery of mutant ADAR and RNA substrate combinations capable and incapable of this type of RNA editing.

The screen was designed so that mutant ADAR or substrate sequences could be rapidly identified *via* a colorimetric yeast assay. It is known that ADAR2 can deaminate within a stop codon, generating a tryptophan codon. Therefore, the ADAR2 substrate was placed upstream of an α -galactosidase gene, such that if RNA editing takes place, green colonies result; otherwise, white colonies grow. This approach identified several functional ADAR2 mutants and RNA sequences that affect RNA editing efficiency. This strategy enables elucidation of the structural and biochemical characteristics of ADAR activity and could facilitate the identification of RNA editing regulators and the engineering of new ADARs with altered properties.



Engineering the Interface

Re-engineering protein interaction interfaces is a powerful method for exploring protein function. Although the interaction interfaces between proteins and various small molecules, DNA, or other proteins have been successfully redesigned, the engineering of highly conserved protein–DNA interaction interfaces has not been explored in depth.



Now, Simon *et al.* (p 755) use novel nucleosides and phage display technology to re-engineer the recognition surface between the homeodomain (HD), a highly conserved DNA-binding domain, and its DNA recognition motif.

The re-engineering of the HD–DNA interface was accomplished by exploiting the hydrophobic contact between an isoleucine in the HD and the C5-methyl group of a thymidine in the DNA. Thymidine analogues containing alkynyl appendages at the C5 position were synthesized and incorporated into the HD–DNA recognition sequence. Phage display was then used to identify mutant HDs that bound to the modified DNA. X-ray crystallization and biochemical studies of one mutant, HD₆, revealed that its fold and stability profile are similar to those of the wild-type HD. That such a highly conserved interaction interface can be so substantially redesigned and yet retain its structure and function is striking. This work demonstrates that even highly conserved binding interfaces are adaptable, paving the way for re-engineering various other domains in the proteome.

Introducing our AUTHORS



Nathan A. Magarvey

Current position: Harvard Medical School, Department of Biological Chemistry and Molecular Pharmacology. Postdoctoral Fellow with Prof. Christopher T. Walsh

Education: Dalhousie University, B.S. in biochemistry, 1998; University of Minnesota Ph.D. in microbiology, immunology and cancer biology with Prof. David H. Sherman, 2005

Nonscientific interests: Camping, hiking, hockey, tennis

Of all the small molecules created, those from nature act with the highest level of specificity, a testament to evolutionary pressures leading to their selection. Equally striking are the assembly strategies to create pools of these molecules. In this paper, we clone the genes for synthesizing the tubulin-binding agent cryptophycin, a mixed polyketide/nonribosomal peptide product from a symbiotic cyanobacterium, and delineate the steps in its assembly. The cryptophycin assembly system is unique in that it produces ~25 analogues. Using an artificial system, we explore the flexibility of the cryptophycin biosynthesis and succeed in using the assembly system to create molecules that have not been made previously and/or those that are difficult to synthesize in the lab. Our hope is that these studies serve as a mechanism to assist in bringing the cryptophycins and other natural products to the war on cancer. (Read Magarvey's article on p 766 and Point of View on p 747.)

Microbes biosynthesize some of the most complex and medically useful compounds. My research is directed toward understanding how they do it. In this article, we describe how the cyanobacteria *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 biosynthesize cryptophycins, potent anticancer agents. Specifically, we have learned the identity of the precursors used to assemble the cryptophycins; cloned, sequenced, and annotated the cryptophycin biosynthetic genes; explored the flexibility of the biosynthetic system through extensive feeding experiments; and heterologously expressed the CYP450 from the cryptophycin biosynthetic pathway, which enabled us to stereospecifically install an epoxide critical to the activity of the cryptophycins. (Read Beck's article on p 766 and Point of View on p 747.)

Current position: University of Michigan, postdoctoral scholar with Prof. David H. Sherman, University of Michigan

Education: Stanford University, B.S. in chemistry, 1995; The Scripps Research Institute, Ph.D. in cellular and macromolecular structure and chemistry with Prof. John H. Elder, 2001

Postdoctoral work: Uppsala University, postdoctoral scholar with Prof. Bengt Mannervik, 2001–2003

Nonscientific interests: Traveling, jazz trumpet, scuba diving



Zachary Q. Beck



Matthew D. Simon

Current position: Massachusetts General Hospital, Department of Molecular Biology, postdoctoral fellow with Prof. Robert Kingston

Education: Tufts University, B.A. in biochemistry, 1999; University of California, Berkeley, Ph.D. in chemistry with Prof. Kevan M. Shokat, 2006

Nonscientific interests: Sound art, radio production

I am interested in the biochemical mechanisms underlying transcriptional regulation and epigenetic inheritance. To facilitate these investigations, we have been working on the synthetic introduction of prosthetic groups capable of functioning as chemical handles. As one example, we have incorporated non-natural nucleosides into DNA and used phage display to select for mutant transcription factors that bind with specificity for the non-natural strands. This project has allowed us to probe the adaptability of a homeodomain–DNA interaction and then examine one of the reengineered interfaces. (Read Simon's article on p 755.)

Introducing our AUTHORS

ACS
chemical
biology



Morris E. Feldman

Current position: University of California, San Francisco, graduate group in biophysics, Ph.D. candidate with Prof. Kevan M. Shokat

Education: University of Chicago, B.A. in biological sciences, 1999

Nonscientific interests: Rock climbing

This work grew out of my interest in the structural basis of molecular recognition. Transcription factors bind to DNA in a sequence-dependent manner to control the expression levels of genes. We chemically modified a transcription factor binding site and selected transcription factor mutants that could tolerate the modification in order to investigate the binding determinants of the protein–DNA interface. Structural and biochemical analysis revealed a remarkable resiliency of the interface, which maintained sequence-specific binding to the DNA despite perturbation of hydrophobic contacts between the transcription factor and its binding site. (Read Feldman's article on p 755.)

My research focuses on the localization and function of various phosphatidylinositols (PIs) in mammalian cells. I am interested in PI recognizing enzymes and their role in signaling and lipid metabolism. Close collaboration between chemists and biologists enabled us to design and characterize a powerful and specific vanadium-based small-molecule inhibitor of the enzyme PTEN. The vanadium complex was useful for dissecting PI 3,4,5-trisphosphate-dependent signaling, because the effects of other growth-factor-induced pathways could be excluded by raising PI 3,4,5-trisphosphate levels solely through the inhibition of PTEN. This inhibitor will help to resolve unknown pathways and will be essential in exploring the role of PTEN in PI dynamics and signaling. (Read Rosivatz's article on p 780.)

Current position: Imperial College London, Division of Cell and Molecular Biology, postdoctoral research associate with Dr. Rudiger Woscholski

Education: University of Innsbruck, Austria, Institute of Medical Chemistry and Biochemistry, M.S. in biology with Prof. Johann Hofmann, 2000; Technical University of Munich, Germany, Institute of Molecular Pathology and University of Innsbruck, Austria, Institute of Medical Chemistry and Biochemistry, Ph.D. in molecular genetics and biochemistry with Dr. Karl-Friedrich Becker and Prof. Florian Ueberall, 2004

Nonscientific interests: Mountaineering, jogging, cycling, nature, London



Erika Rosivatz



Subhash Pokharel

Current position: University of Utah, Department of Chemistry, Ph.D. candidate with Prof. Peter A. Beal

Education: Tribhuvan University, Kathmandu, Nepal, M.Sc. in chemistry, 1994, and bachelor's degree in law, 1999

Nonscientific interests: Hiking, biking, swimming, spending time with my wife and daughter

As a chemical biologist, I am always interested in developing versatile and simple tools to answer basic biological questions at the molecular level. The main focus of my research is to develop new approaches to control editing reactions catalyzed by RNA editing enzymes. In this paper, we report a high-throughput screen for ADAR/substrate combinations capable of RNA editing that can be carried out in yeast growing on agar plates. In the future, we intend to extend this screen to discover regulators of RNA editing reactions and to evolve ADARs that edit at new sites. (Read Pokharel's article on p 761.)

Spotlight

The Bee with Methyl-C's

Recently, the genomic sequence of the honeybee was completed. Though this insect may seem like a cousin of the research workhorse, the fruit fly, these two organisms probably diverged from their common ancestor 300 million years ago. For perspective, this is the same time period in which humans and chickens



Georgette Douma, Getty Images.

shared a common ancestor. Among the sequence was an interesting secret that makes the bee more like a human than a fruit fly. In a recent report, Wang *et al.* (*Science* 2006, 314, 645–647) report that the bee genome encodes enzymes resembling the human DNA deoxycytosine methyltransferases (DNMTs).

In humans, these enzymes methylate CpG dinucleotides in promoter regions, thereby regulating the activity of a particular gene's transcription. Methyl-CpG modifications in mammals can be linked to processes ranging from genome maintenance to tumorigenesis to sex-specific gene regulation. For the first time, it is clear that this type of methylation reaches into an invertebrate. The study steps beyond genomic predictions and tests several of the methylase enzymes for expression and activity. As in mammals, some of this social insect's enzymes are expressed in a tissue-specific or developmental stage-specific manner. Biochemical tests with an *in vitro* translated batch of each methylase showed that the predicted bee DNMTs are *bona fide* methylases. Additionally, a complete nuclease digest of bee genomic DNA to single nucleotides revealed methyl-C as a naturally occurring monomer. The methylation is not as widespread as in mammals, so the underlying function for this modification may be different in bees. This does appear likely, because the authors identified several methylated regions of the bee genome and they lie within coding genes rather than in promoters. This difference might make bees a honey-sweet new organism for studying gene regulation and the effects of the DNMTs. **Jason G. Underwood, Ph.D.**

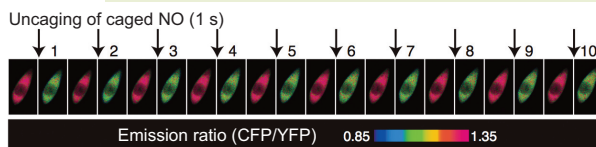
Published online December 15, 2006 • 10.1021/cb600481e CCC: \$33.50
© 2006 by American Chemical Society

Piccell Picks NO Releasing Cells

Nitric oxide (NO) is a signaling molecule that plays a role in several important biological processes, including blood vessel dilatation and neurotransmission. Many fluorescent NO indicators enable the detection of intracellular NO levels, but existing methods for monitoring the release of NO from cells suffer from lack of sensitivity and reversibility. Now, Sato *et al.* (*Anal. Chem.*, published online Nov 9, 2006, DOI: 10.1021/ac061791b) present Piccell, a novel, cellular, FRET-based indicator for NO that enables exploration of the spatiotemporal dynamics of NO release from cells.

Piccell is a cell line containing a genetically encoded molecular indicator that undergoes a change in FRET in the presence of guanosine 3',5'-cyclic monophosphate (cGMP). The protein soluble guanylate cyclase, endogenously present in the cell line from which Piccell is derived, generates cGMP upon binding to NO, resulting in a subsequent FRET signal. The authors initially demonstrated that addition of NO to Piccell generated an exquisitely sensitive, selective, and reversible FRET response. To determine whether the indicator could detect NO released from cells, they cocultured Piccell with vascular endothelial cells or hippocampal neurons. Endothelial cells transiently release NO upon stimulation with ATP. Addition of a specific concentration of ATP to the Piccell–endothelial cell coculture resulted in a change in FRET that could be correlated with a specific picomolar concentration of NO. Interestingly, repeated cycles of decreasing and increasing

FRET signals were observed in the Piccell–neuron coculture in the absence of any external



Reprinted with permission from Sato, M., *et al.*, *Anal. Chem.*, DOI: 10.1021/ac061791b. Copyright 2006 American Chemical Society.

stimuli. Investigation into the mechanism of this intriguing finding revealed that the spontaneous and periodic NO release was dependent on the *N*-methyl-D-aspartic acid receptor and that it synchronized with Ca²⁺ spikes in the neurons. This work establishes Piccell as a powerful tool for investigating the regulation and function of NO release from cells. **Eva Gordon**

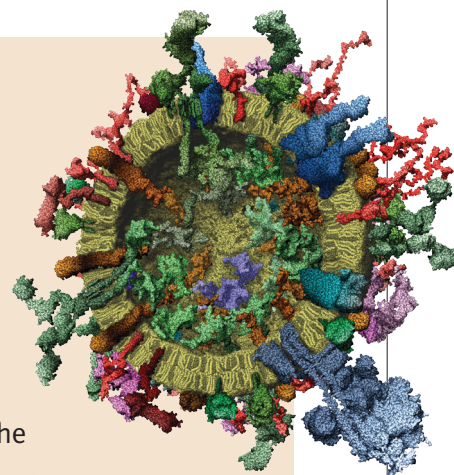
Sizing Up the Synaptic Vesicle

Trafficking vesicles are the vehicles by which the various compartments within a cell transport both membrane constituents and soluble material to each other and to the extracellular space. Although many of the vesicle constituents that act as traffic cops and crossing guards to regulate this highly complex process have been extensively characterized, a comprehensive description of the global structure of trafficking vesicles has not been attempted. Using the synaptic vesicle (SV) as a model trafficking organelle, Takamori *et al.* (*Cell* 2006, 127, 831–846) now quantitatively describe the SV in terms of its protein and lipid composition.

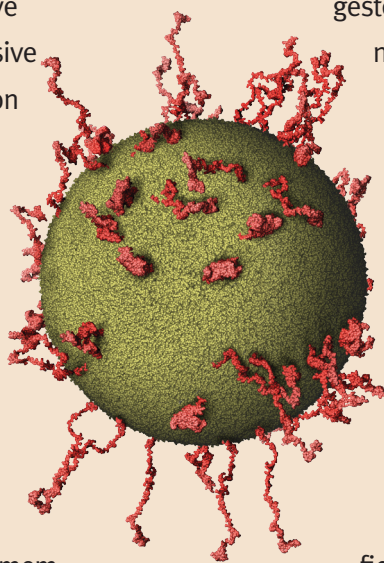
SVs perform the critical function of transporting neurotransmitters to nerve terminals of neurons. An impressive array of molecular characterization technologies, including mass spectrometry, fluorescence correlation spectroscopy, electron microscopy, molecular modeling, and molecular dynamics simulations, was employed to paint a 3D picture of the SV structure. Hundreds of proteins from diverse structural and functional classes were identified as mem-

bers of the SV or at least as visiting constituents, and the SV was found to contain high levels of cholesterol but low amounts of phosphatidylinositol. In addition, the amount in grams and the copy number of the major protein and lipid constituents on the average SV were calculated, along with the vesicle size, density, and mass. On the basis of this quantitative composition data and known structural information of the vesicle constituents, a molecular model of an average SV was generated. Analysis of the SV structure sug-

gested that it is strongly dominated by its protein constituents and that proteins essential for exocytosis and neurotransmitter loading are present in high copy number. This remarkable model provides insight into how SVs function and lays the groundwork for the portrayal of other trafficking vesicles. **Eva Gordon**



Both images reprinted from *Cell*, 127, Takamori, S., *et al.*, Molecular anatomy of a trafficking organelle, 831–846. Copyright 2006, with permission from Elsevier.



Silence of the Plants

Methylation of DNA plays an important role in genome stability and gene regulation in many eukaryotes. In plants, a special RNA interference (RNAi) mechanism guides the methylation of particular DNA segments to trigger gene silencing. Small interfering RNAs (siRNAs) of ~24 nucleotides are first processed out of longer double-stranded precursors. These short RNAs then go on to program the target recognition of an RNA–protein complex. It still remains a mystery whether these complexes use their siRNAs to target a complementary RNA or DNA at the proper methylation locus. Among the most critical protein components in all known RNAi mechanisms are the Argonaute proteins (AGOs). In the mustard plant, *Arabidopsis*, AGO4 is the specialized Argonaute family member that participates in RNA-guided DNA methylation. Now, a recent study indicates that this one AGO protein may have two separable functions in the RNAi pathway. Qi *et al.*

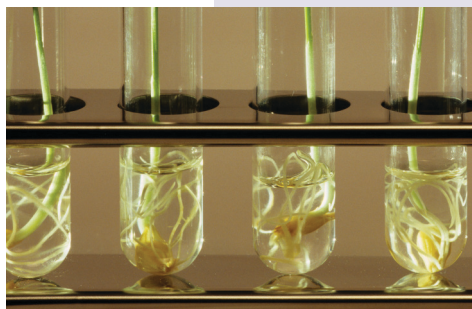
(*Nature* 2006, 443, 1008–1012) began by purifying AGO4 from the plant and sequencing the associated 24 nucleotide RNAs. Among these RNAs, almost half were

derived from abundant repeat sequences in the genome, and others were derived from DNA transposons, consistent with AGO4's role in genome maintenance. The authors then probed whether RNA

Parasites Roll the Dicer

In the last several years, scientists have investigated the RNA interference (RNAi) phenomenon in a wide array of model organisms. In general, the RNAi pathway is fed by sequences of double-stranded RNAs (dsRNAs) that are chopped up into small interfering RNAs (siRNAs) and selectively incorporated into gene-silencing complexes. In many organisms, these silencing complexes go on to bind to and trigger the degradation of complementary messenger RNAs. Among the first protein components linked to the RNAi pathway was Dicer, the enzyme that makes siRNAs from long double-stranded precursors. Since this discovery, the enzyme has been found in organisms ranging from single-celled *Giardia* to humans. Now, the Dicer from another unicellular eukaryote, *Trypanosoma brucei*, has been uncovered. This parasitic organism causes the deadly African sleeping sickness in humans and is spread by its other host, the biting tsetse fly. Shi *et al.* (*RNA* 2006, 12, 2063–2072) used a clever candidate-gene approach to identify Dicer in this organism. Because

gene comparison methods had not been successful in finding this parasite's equivalent enzyme, the authors took advantage of the hallmark of all Dicers, the double-strand cutting RNase III domain. Comparative genomics of *T. brucei* and another species, *T. cruzi*, which does not display RNAi, yielded two possible candidates. One of these two localized to the cell's cytoplasm and could dice long radioactive dsRNA into fragments 24–26 nucleotides in length. Both of these features were consistent with the roles of the candidate genes in RNAi. The authors went on to further prove the function by disrupting and enhancing the Dicer's expression in the organism. Interestingly, they also found that a single point mutation in a yet uncharacterized domain of Dicer can disrupt RNAi function. Altogether, this study demonstrates that an RNAi enzyme that has eluded researchers to date is alive and well in this tiny parasite. The dramatic sequence differences between *T. brucei* and human Dicer enzymes make it a compelling target for drug design. Jason G. Underwood, Ph.D.



Andy Sotiriou, Getty Images.

or DNA could be the target of AGO4. They found that a purified AGO4 complex from plants could cleave a complementary RNA sequence derived from one of the transposons. Mutation of key catalytic residues abolished this cleavage, but interestingly, methylation and silencing could still take place at many loci. Thus, although AGO4 in the plant can make siRNAs by using its

catalytic center, this activity is not necessary for stimulating and guiding methylation. The protein must also allow programming by RNA products from other enzymes. How the RNAi mechanism is tied to DNA methylation is still not well understood, but these results place AGO4 at the center of attention. Jason G. Underwood, Ph.D.

Microbes Go Digital

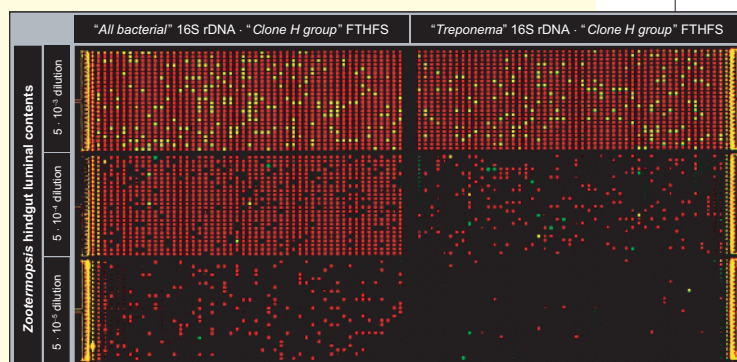
Microbes perform incredibly diverse functions within the environments in which they reside. But whether they are munching on a hamburger in the human stomach or a nice juicy piece of wood in the termite gut, the organisms responsible for specific enzymatic activities within a given ecosystem are often difficult to identify among the plethora of other microbes with whom they coexist. Now, Ottesen *et al.* (*Science* 2006, 314, 1464–1467) combine microfluidics and digital polymerase chain reaction (PCR) to separate, amplify, and analyze multiple genes from individual bacteria within these complex ecosystems, enabling the systematic identification of bacteria carrying a particular gene.

Both termites and bacteria within the termite hindgut rely on bacterial formyl-tetrahydrofolate synthetase (FTHFS) for key metabolic activities, but the

identity of the bacterial species responsible for this function had not been elucidated. To address this problem, the authors used a microfluidic device to separate and partition out single cells from the community of organisms

present in the hindgut of the *Zootermopsis nevadensis* dampwood termite. The individual cells were then used as genomic templates for PCR amplification of a specific genotype (the clone H group) of FTHFS, along with 16S ribosomal RNA (rRNA) to enable identification of the bacterial species containing the gene. Using this method, they determined that ~1% of the bacteria residing in the termite hindgut carried clone H group FTHFS and that of this

group, a cluster of the genus *Treponema* represented the majority of the clone H group FTHFS-encoding cells. With



From Ottesen, E. A., *et al.*, *Science*, Dec 1, 2006, DOI: 10.1126/science.1131370. Reprinted with permission from AAAS.

this information in hand, the authors designed more specific PCR reactions to further characterize the *Treponema* population in *Z. nevadensis*, and they determined that treponemal cells constituted 10–12% of the bacterial community. This report presents an innovative method for identifying uncultivated bacterial species within complex ecosystems by exploiting their rRNA fingerprint and their ability to perform specific metabolic activities. **Eva Gordon**

Fountain of . . . Resveratrol?

Resveratrol, a polyphenolic compound that is available as a dietary supplement, is known to extend the lifespan of diverse species, from yeast to worms to fish. The compound is produced by various plants and is notably present in the skin of red grapes and in red wine. Though this information alone may be ample evidence for some to rationalize the consumption of large amounts of red wine on a regular basis, it would be wise to first fully understand the biological consequences of resveratrol intake. Two groups (Lagouge *et al.*, *Cell* 2006, DOI: 10.1016/j.cell.2006.11.013

and Baur *et al.*, *Nature* 2006, 444, 337–342) report their investigations into the molecular mechanisms behind the fountain-of-youth-like effects of resveratrol

and the impact of resveratrol treatment on the physiology and survival of mice.

Baur *et al.* systematically evaluated the physiological and health effects of resveratrol on mice.

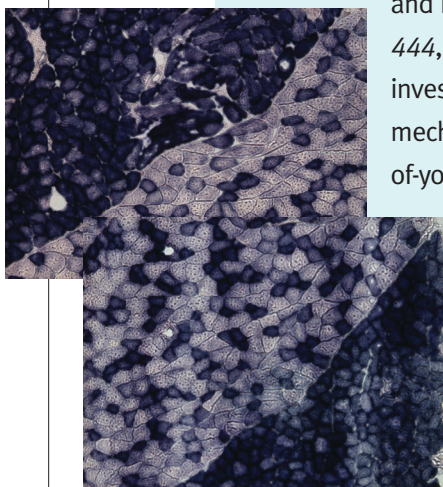
Middle-aged mice (all of 1 year old!) were fed either a standard diet (SD), a high-calorie diet (HC), or a high-calorie diet with resveratrol (HCR). Compared with HC mice, HCR mice exhibited prolonged survival and improved quality of life as assessed through monitor-

ing of balance and motor coordination. In addition, the HC group had increased levels of several molecular indicators for diabetes and shorter lifespan compared with the HCR and SD groups. The HC group also had larger livers and displayed greater overall pathology in the liver and the heart than the HCR or SD groups.

Both research groups uncovered information that helps explain the molecular basis for these remarkable effects. Mice treated with resveratrol exhibited improved sensitivity to insulin, indicative of decreased risks for cardiovascular disease and diabetes. In addition, resveratrol-treated mice had more mitochondria and increased mitochondrial activity than untreated mice. Microarray experiments were also performed to determine which gene expression levels were affected by resveratrol intake. Not surprisingly, many of the most dramatic changes in expression patterns were observed within groups of genes functionally involved in mitochondrial biogenesis and function and in energy homeostasis. Strikingly, resveratrol opposed the effects of high calorie intake in 144 out of 153 significantly altered pathways.

Resveratrol is known to interact with SIRT1, a member of the sirtuin family of nicotinamide adenine dinucleotide-dependent deacetylases. SIRT1 deacetylates peroxisome proliferator-activated receptor γ coactivator (PGC-1 α), a protein involved in mitochondrial biogenesis and function. Delving even further into the molecular

(continued on page 735)



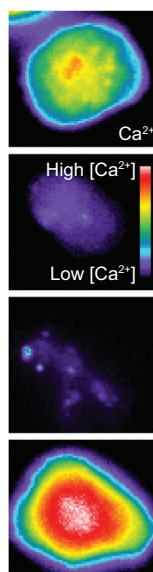
Both images reprinted from Cell, Lagouge, M., *et al.*, Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC-1 α , DOI: 10.1016/j.cell.2006.11.013, Copyright 2006, with permission from Elsevier.

Caged Antigens

T cells use their T cell receptor (TCR) to scan the cellular environment for foreign antigenic peptides presented by major histocompatibility complex (MHC) molecules on antigen-presenting cells. Recent evidence suggests that the spatial pattern of signaling and adhesion molecules activated by the TCR–MHC interaction, called the immunological synapse, plays a direct role in regulation of the immune response. To probe the spatial regulation of T cell signaling, DeMond *et al.* (*J. Am. Chem. Soc.*, 2006, 128, 15,354–15,355) have developed a method for selectively activating T cells in a spatially and temporally defined manner.

Using the well-studied TCR/MHC pair AND/MCC-IE^k, the authors

replaced a critical lysine in the antigenic peptide with a lysine containing the caging group 6-nitroveratryloxycarbonyl (NVOC) appended to its side chain nitrogen. They anticipated that the presence of the caging group would prevent TCR activation but that exposure to UV light, which removes the NVOC group, would restore the interaction and induce T cell activation. Indeed, T cells stimulated with mobile intercellular adhesion molecule-1 (ICAM)-containing lipid bilayers displaying either uncaged MCC-IE^k or NVOC-MCC-IE^k that had already been exposed to UV light adopted a rounded morphology and fluxed calcium, an indication of T cell activation.



Reprinted with permission from DeMond, A. L., *et al.*, *J. Am. Chem. Soc.* 128, 15,354–15,355. Copyright 2006 American Chemical Society.

In contrast, when T cells were exposed to NVOC-MCC-IE^k lipid bilayers that had not been exposed to UV light, the cells crawled and exhibited a low level of intracellular calcium, signifying that the TCR was not activated. Remarkably, however, upon exposure to UV light, these cells stopped migrating, adopted a rounded morphology, and formed a ring of ICAM characteristic of the immune synapse. Caged peptides

such as NVOC-MCC-IE^k provide unique molecular tools with which to probe the spatial and temporal regulation of the molecules involved in T cell activation. **Eva Gordon**

Fountain of . . . Resveratrol?, *continued from page 734*

pathways affected by resveratrol treatment *in vivo*, Lagouge *et al.* discovered that activation of SIRT1 by resveratrol affects PGC-1 α target gene expression in muscle and fat, specifically genes involved in oxidative phosphorylation and mitochondrial biogenesis. These effects ultimately enabled resistance to diet-induced obesity in mice. Interestingly, the authors also discovered that three single-nucleotide polymorphisms in the human *Sirt1* gene correlate with energy homeostasis, further substantiating the role of SIRT1 in energy regulation.

These studies elegantly reveal the intriguing physiological consequences of supplementing a high-calorie diet with resveratrol. A variety of physical and metabolic indicators suggest that this compound essentially reverses the negative effects of overeating and favorably modulates known longevity pathways. In addition, the delineation of resveratrol's mechanism of action presented in these investigations demonstrates that small molecules can effectively manipulate the key pathways involved in regulating metabolic- and age-related disorders. Perhaps someday, as we continue to decipher the workings behind metabolism and aging, we will figure out exactly what compound to put in that fountain. **Eva Gordon**

Does History Repeat Itself? The Emergence of a New Discipline

Laura M. C. Barter, David R. Klug, and Rüdiger Woscholski*

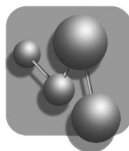
Chemical Biology Centre, Imperial College London, Exhibition Road, London SW72AZ, United Kingdom

The influence of chemistry in the development of the life sciences led to the emergence of biochemistry departments in many institutions and organizations. This subject focused on the “study of the chemistry of life processes” (1) and over time became independent of its original disciplines and established specializations in its own right. Biochemistry continues to be an area of expansion, with many degree courses offered to students around the world. Biochemistry was initially a truly interdisciplinary area, but it evolved quickly and in due process distanced itself increasingly from its roots within chemistry to become a “biological” entity.

Chemical biology is another, more recent, offshoot from chemistry that aims to combine the physical and the life sciences (2). By our definition, however, many important distinctions exist between the emerging discipline of chemical biology and the now-established disciplines within biochemistry. In particular, we at the Chemical Biology Centre (CBC) in London characterize our approach as multidisciplinary rather than interdisciplinary. Such progress requires a mastery of the cutting-edge physical and biological sciences; the best and most relevant parts of each must be combined to solve well-identified problems. If history repeats itself, this new approach of applying chemistry to solve biological problems will become a substantial and evolving research area; but will it, and should it, lead to the establishment of new departments? We argue that the creation of new departments

and undergraduate courses is not necessarily the way forward. We believe that the modern discipline of chemical biology will grow *via* new and flexible research structures that capture the quantitative mature approach prevalent in physics, chemistry, and engineering and apply it to the rapidly evolving fields of biology, biochemistry, and medicine.

The demand is increasing for a cohort of scientists for whom multidisciplinary thinking is the norm in both academic and industrial research. Imperial College London, the Institute of Cancer Research, and the Cancer Research UK London Research Institute, all highly ranked research institutions in Europe, joined forces to create the CBC five years ago. This was driven by the desire to bring together those scientists whose research activities would be particularly synergistic. The key objectives of the CBC were both to develop multidisciplinary research programs across the life science-physical science interface and to provide training for that cohort of researchers seen as so vital to postgenomic science and technology. The CBC offers lecture courses, examinations, seminars, workshops, and research training for physical scientists who wish to work in the chemical biology area with a particular emphasis on new approaches in molecular medicine. The CBC also supports the devel-



Chemical Biology Centre

*Corresponding author,
r.woscholski@imperial.ac.uk.

Published online December 15, 2006

10.1021/cb600468u CCC: \$33.50

© 2006 by American Chemical Society

Box 1. CBC Training Courses

M.Res./Ph.D. in Protein Membrane Chemical Biology. This course starts with a full-time, one-year M.Res. qualification in membrane and protein chemical biology.

Training includes a multidisciplinary research project, taught courses in advanced biochemistry and biomolecular techniques, specialist lectures in transferable skills, and group discussion sessions. CBC DTC continues to oversee a three-year Ph.D. assignment, which leads to a degree awarded by Imperial College London.

Ph.D. Studentships in Single-Cell Proteomics. Six Ph.D. studentships affiliated with the DTC are assigned to support this EPSRC-funded CBC project, which aims to create an unrivaled suite of technologies for the study of single cells using proteomic approaches.

M.Res. in Bioimaging Sciences. This full-time, one-year M.Res. course features a nine-month multidisciplinary research project and taught lecture courses. The fundamentals of modern imaging methodologies are covered, including techniques and applications in medicine and the pharmaceutical industry, and the chemistry behind imaging agents and biomarkers.

M.Res. in Biomedical Physical Chemistry. This full-time, one-year course enables students to bridge the gaps of language, perspective, and methodology between the life and physical sciences. A nine-month multidisciplinary research project, taught courses in physical and chemical technologies, and training in biomedical research, advanced biochemistry, and practical biomolecular techniques are included.

opment of new tools and techniques and facilitates their translation among traditional academic institutions and between academia and industry. Both tool development for the life sciences and the quantitative investigation of biological engineering rules rely on a close interaction and collaboration between physical scientists and life scientists. As a virtual center, the CBC looks for new ways for these researchers to interact. It has a policy of zero barrier to entry, zero barrier to exit with more than 50 research groups currently involved in CBC research projects. The CBC is also home to various targeted research programs such as the Non-Specific Drug Binding Initiative funded by GlaxoSmithKline and the Single Cell Proteomics Project supported by the Engineering and Physical Sciences Research Council (EPSRC) and the Biotechnology and Biological Sciences Research Council.

In the context of graduate education, a major priority for the CBC was to create graduate programs for students who have

degrees in the physical sciences (chemistry, physics, engineering, *etc.*) to equip them with the necessary knowledge, skills, and language to confidently apply their physical science backgrounds to problems in the life sciences. In this respect, we looked more to the waves of physical scientists who emerged to establish fields such as modern structural biology than to the growth of biochemistry departments. The CBC is particularly aware that many students educated in the physical sciences are keen to be trained in projects that are focused on problems in the life sciences. The CBC's main strategy is to attract this pool of physical scientists, train them to use their skills and knowledge to solve biological problems with advanced physical sciences approaches, and modify and develop those approaches accordingly.

To achieve these objectives, the CBC Doctoral Training Centre (DTC) was established with funding from the Life Sciences Interface Programme of the EPSRC. The focus of the DTC is on chemical biology at, within, or

near cell membranes. This Ph.D. program is complemented by three master courses, which either precede the Ph.D. education, or, in some cases, function as a stand-alone training for those students wishing to broaden their knowledge and understanding in the art of multidisciplinary research without the commitment of undertaking a Ph.D. (See Box 1 for further information on the Ph.D. and master's in research (M.Res.) courses offered by the CBC DTC.)

The courses under the CBC umbrella provide physical science graduate students with (i) cross-departmental, cross-institutional, multidisciplinary training toward careers at the physical science/life science interface, (ii) the necessary guidance and experience to apply their physical science skills to problems in the life sciences, and (iii) the ability to move with confidence into biological, biotechnological, and biomedical research, bringing them relevant disciplines, skills, and approaches of their undergraduate training.

For the CBC to achieve its aims, input from a broad range of disciplines is essential. It is clear that advances will depend on the collaborative efforts of biochemists, chemists, medics, physicists, and engineers. Thus the M.Res. programs are designed to bridge the "cultural" divide that can exist between these disciplines because of differences in language, perspective, and methodology. The courses foster development within the type of multi-



CBC chemistry laboratory

Box 2. Format of CBC M.Res. Courses

Taught Component. A fixed lecture program of core courses in the first term is followed by optional courses to complement the research project undertaken by the student in the second term.

Research Project. In October, the nine-month multidisciplinary research project is chosen. Students present a literature report on the topic of their research project by the end of the first term, a talk in July or August, and a final thesis in September.

M.Res. Conference. A meeting for all M.Res. students to present their research projects and to learn about the work carried out by their fellow cohorts. Keynote speakers from other institutions are also invited.

Research Seminars and Colloquia. The chemistry department organizes monthly research seminars by leaders in particular fields. Students interact and discuss scientific issues with these visitors to the department, as well as with staff members. The CBC organizes two or three afternoon colloquia annually, which focus on a particular research area.

Visiting Professors—Advanced Courses. The CBC invites two visiting professors every year for a two-month appointment. The professors deliver eight hours of lectures on their areas of research. Interactions between the students and the professors are encouraged.

disciplinary setting that students will expect to enter on completion of the degree. The courses have also been designed to meet the demands for both breadth of knowledge and the physical techniques employed in this area of work. Moreover, they provide a depth of experience in experimental practice that is gained from undertaking a single nine-month research project. Students concentrate on one research project, which is jointly supervised by at least one physical and one biological scientist. The projects are expected to involve the research laboratories and teams of both collaborating supervisors. This allows the students to benefit from interaction with supervisors as well as postdoctoral and postgraduate researchers from both disciplines. This multidisciplinary approach provides a hands-on experience of biological techniques and systems that is both challenging and rewarding for physical science graduates. The format of a single project running through the entire year of the M.Res. course has been adopted primarily because it takes time to become accustomed to working in a multidisciplinary environment. A single nine-month project is

challenging but offers the satisfaction of a real scientific achievement at the end. A depth of knowledge and confidence is gained that cannot readily be found in structured practical work or shorter projects. This direct experience of multidisciplinary research is something that we regard as key to the development of modern chemical biology.

The training that the M.Res. students receive is complemented by a structured lecture program and seminars. Training is not limited to the M.Res. cohort of students. The CBC encourages all DTC students to meet, interact, and discuss scientific issues with visitors to the chemistry department and staff members by organizing colloquia, workshops, and student conferences. The CBC also runs a visiting professor program where a professor is invited to reside in the department for a period of up to two months, and interaction between the students and professor is encouraged. Students benefit from the stimulation of hearing about the most recent scientific advances from esteemed scientists, and lectures delivered by the professor widen the students' scientific horizons. Moreover, contact with scientists from other institutions exposes students to a broader range of scientific approaches and attitudes. (See Box 2 on the format of the training and courses.)

Specialist transferable-skills lectures also form part of the training. Students are trained in a wide range of transferable skills following exposure to the various teaching and learning aspects of the M.Res. courses, such as safety awareness, effective commu-

Box 3. CBC Transferable-Skills Courses

The Joint DTCs' Teamwork Course. An intensive course in teamwork, personal development, information gathering, and presentation, with a focus on networking and relationship building among Ph.D. students within the CBC DTC as well as with other DTCs. Elements include a business game about a biotechnology start-up company and a presentation exercise based on a research grant pitch.

Science Communication Course. An intensive two-and-a-half day course directed by Gareth Mitchell, a lecturer in radio broadcast communication in the Science Communication Group at Imperial College London and a practicing scientific broadcaster for the BBC. Students develop skills on both sides of the camera and the microphone, as well as effective scientific writing under real-time pressure.

Decision-Making Course. This course on the methodology and analysis of decision making follows models in the teamwork course. Specific business content, such as financial analysis, prepares students for working in pharmaceutical companies and other industries.

Box 4. Contacts

Chemical Biology Centre

www.chemicalbiology.ac.uk, +44 (0)20 7594 5880, cbcadmin@imperial.ac.uk

M.Res. in Biomedical Physical Chemistry

Laura Barter, +44 (0)20 7594 1885, l.barter@imperial.ac.uk

M.Res./Ph.D. in Protein Membrane Chemical Biology

Rudiger Woscholski, +44 (0)20 7594 5305, r.woscholski@imperial.ac.uk

M.Res. Bioimaging Sciences

Nick Long and Ramon Vilar, +44 (0)20 7594 5781/7594 1967,
n.long@imperial.ac.uk or r.vilar@imperial.ac.uk

Department of Chemistry

Doris Pappoe, +44 (0)20 7594 5864, d.pappoe@imperial.ac.uk; online application: www.imperial.ac.uk/P1397.html

nication, time and project management, and management skills. The Graduate School of Engineering and Physical Sciences at Imperial College London also provides transferable skills training. Examples of course topics include intellectual property management, personal organization and effectiveness, research ethics, and technical presentation skills. The CBC has also devised a suite of three transferable-skills courses, specifically for the DTC students, on science communication, team work, and decision making. These courses form the core of the students' transferable-skills training. (See Box 3 for further information about these courses.)

A number of multidisciplinary DTCs are currently running in the U.K. (Examples of DTCs funded by the EPSRC can be found at www.epsrc.ac.uk/PostgraduateTraining/LSIDoctoralTrainingCentres/default.htm.) The CBC has already developed links with some of them, and we hope these interactions will continue to expand. For example, the CBC transferable-skills courses are given jointly with the DTCs in Warwick, Leeds/Sheffield, Edinburgh, Glasgow, and Dundee. The courses not only bring together our cohort of DTC students but also facilitate networking and interdisciplinary communication with the other multidisciplinary life/physical science DTCs.

Our students have responded positively to this training program and are keen to embrace the opportunities provided by the CBC courses. An important indicator of success for the current courses in the CBC is the scientific output. We are delighted that research undertaken during this short nine-month period of the M.Res. research project has resulted in several publications (3–9), one of which is in this issue (9). Almost all of our students in the DTC have yet to finish their Ph.D. work, so we expect that this publication list will grow over time. This will corroborate the multidisciplinary nature of the training offered by the CBC and the DTC.

Chemical biology education is still in its infancy, and time will tell whether it will evolve into its own discipline. Postgraduate training will certainly be the quickest route to alleviating the lack of suitably trained physical scientists and deliver future leaders of chemical biology research. Although undergraduate training of the future will no doubt include more chemical biology, either through the introduction of corresponding modules within existing degree courses or through the creation of dedicated degree courses, it is unlikely to be the sole provider for the next generation of chemical biologists. We ourselves are focused on chemical biology as a postgraduate discipline for those with first-class expertise in disciplines such as chemistry and physics but who

want to use this knowledge in a meaningful and useful way in a biological context. (See Box 4 for contact information.) The multidisciplinary environment of the CBC is designed to take these core strengths and train research cohorts who are each multidisciplinary in their own right. In this sense, every Ph.D. student is expected to be an expert in both an area of life science and an area of physical science. Chemical biology is therefore not for the faint of heart.

REFERENCES

- Berg, J. M., Tymoczko, J. L., and Stryer, L. (2006) Biochemistry, 6th ed., W. H. Freeman and Co., New York.
- Larijani, B., Rosser, C. A., and Woscholski, R. (2006) Introduction, in *Chemical Biology: Techniques and Applications* (Larijani, B., Rosser, C. A., and Woscholski, R., Eds.) pp 1–10, John Wiley & Sons, Ltd., Chichester, U.K.
- Treanor, B., Lanigan, P. M. P., Kumar, S., Dunsby, C., Munro, I., Auksoy, E., Culley, F. J., Purbhoo, M. A., Phillips, D., Neil, M. A. A., Burshtyn, D. N., French, P. M. W., and Davis, D. M. (2006) Microclusters of inhibitory killer immunoglobulin like receptor signaling at natural killer cell immunological synapses, *J. Cell Biol.* **174**, 153–161.
- Seddon, J. M., Squires, A. M., Conn, C. E., Ces, O., Heron, A. J., Mulet, X., Shearman, G. C., and Templer, R. H. (2006) Pressure-jump X-ray studies of liquid crystal transitions in lipids, *Philos. Trans. R. Soc. London, Ser. A* **364**, 2635–2655.
- Conn, C. E., Ces, O., Mulet, X., Finet, S., Winter, R., Seddon, J. M., and Templer, R. H. (2006) Dynamics of structural transformations between lamellar and inverse bicontinuous cubic lyotropic phases, *Phys. Rev. Lett.* **96**, 108012-1–108012-4.
- Ces, O., and Mulet, X. (2006) Physical coupling between lipids and proteins: a paradigm for cellular control, *Signal Transduction* **6**, 112–132.
- Baciu, M., Sebai, S. C., Ces, O., Mulet, X., Clarke, J. A., Shearman, G. C., Law, R. V., Templer, R. H., Plisson, C., Parker, C. A., and Gee, A. (2006) Degradative transport of cationic amphiphilic drugs across phospholipid bilayers, *Philos. Trans. R. Soc. London, Ser. A* **364**, 2597–2614.
- Burke, M. G., Woscholski, R., and Yaliraki, S. N. (2003) Differential hydrophobicity drives self-assembly in Huntington's disease, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13928–13933.
- Rosivatz, E., Matthews, J. G., McDonald, N. Q., Mulet, X., Ho, K. K., Lossi, N., Schmid, A. C., Mirabelli, M., Pomeranz, K. M., Erneux, C., Lam, E. W.-F., Vilar, R., and Woscholski, R. (2006) A small-molecule inhibitor for phosphatase and tensin homologue deleted on chromosome 10 (PTEN), *ACS Chem. Biol.* **1**, 780–790.

Ready for Their Close-Ups: Investigating Single Molecules

Over the past several decades, biologists and chemists have amassed reams of data about the natural world by investigating its properties at the molecular level. However, a majority of such studies have focused on collections of molecules that typically number in the billions. Because molecules can change their behavior on the basis of interactions with their nearby neighbors, such studies can be misleading—they sum up a population's average properties but give little insight on how molecules behave as individuals. Biophysicist Taekjip Ha of the University of Illinois at Urbana–Champaign is working to fill this knowledge gap by employing a variety of techniques, including pioneering use of single-molecule FRET. By characterizing single biomolecules, he is magnifying scientific understanding of some pivotal molecular components of life.

From Physics to Biochemistry

Ha was born in 1968 in Seoul, South Korea. With both parents as teachers, he quickly learned the importance of a good education. He strived to excel as a student, regardless of the subject matter he studied. Though not particularly focused on science as a child, Ha decided in high school to concentrate his efforts on gaining admission to the physics department of Seoul National University (SNU).

“People who got good scores on South Korea’s national college entrance exam said that’s where they were going,” recalls Ha. “I said, wow, that’s what I have to do. I have to go to the physics department to be with these really smart people.”

Against his parents’ initial wishes—they’d encouraged him early on to become a doctor or lawyer—Ha began his physics

studies at SNU in 1986. He filled his curriculum with physics and mathematics courses. He remembers a class called mathematical analysis as his favorite. The course involved using fundamental theorems to “construct your own mathematical universe,” says Ha, a line of study that taught him logical thinking. He continues to use this skill to set up and analyze his current experiments.

Early in his undergraduate studies, Ha set a goal for himself to attend a university in the United States to pursue his doctoral degree. By the time he graduated in 1990, he had gained admission to a Ph.D. program in physics at the University of California, Berkeley. He chose the school for its highly ranked physics program and its large department. “I thought that perhaps I could have more choices with research,” he says.

Ha started his studies with a plan to do theoretical research in condensed-matter physics, a discipline that studies the materials that make up semiconductors and superconductors. However, after attending a course by experimentalist Eugene Cummins, who studies atomic physics, Ha switched his own focus to experimental study. Soon afterward, he joined the lab of Raymond Jeanloz in Berkeley’s geophysics department. There, he worked on a project to place nitrogen and carbon under very high pressures, with the goal to create a material harder than diamonds.

After several months, Ha and his colleagues had little success with the project. Taking a temporary leave of absence from Berkeley, he returned to South Korea for a year to fulfill South Korea’s military service requirements. Upon his return, Ha joined the lab of Daniel Chemla, a prominent

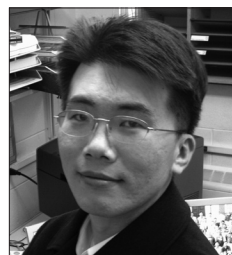


Image courtesy of Taekjip Ha

Profiles provide insights into the lives, backgrounds, career paths, and futures of scientists who serve as Experts on *ACS Chemical Biology's* online Ask the Expert feature. Dr. Ha will begin answering your questions in mid-January, 2007. Readers are encouraged to submit questions to the Experts at www.acschemicalbiology.org. The editors will post the most interesting exchanges on the web site.

Published online December 15, 2006

10.1021/cb6004799 CCC: \$33.50

© 2006 by American Chemical Society

“There were so many things I wanted to study that I couldn’t do it on my own. I figured it was time to start my own group.”

physicist known for his studies of quantum optics of semiconductors. Soon after joining Chemla’s group, Ha began working closely with then-staff scientist Shimon Weiss. Recognizing that Ha had a talent for assembling machines, ranging from switch-boxes to more complicated instruments, Weiss suggested that he take advantage of this skill to build a near-field scanning optical microscope, a machine equipped with a small aperture and a short-pulse laser able to measure a material’s properties with high time and spatial resolution.

Once Ha finished building the machine, he originally planned to use it to study semiconductors. His initial experiments were not successful. However, Ha notes that his seeming failure was actually a turning point for his career. Researching other uses for his new device, he found a few papers that discussed utilizing the machine to measure the properties of single fluorescent molecules. Ha immediately grasped the significance of such research.

“Most of the knowledge we have about chemistry, biology, and physics, we get by studying molecules in the billions and measuring their average properties,” he says. “But looking at single molecules, you have the ability to gain a much deeper understanding of what’s going on.”

Because other groups were already collecting measurements of single fluorescent molecules, Ha decided to add a more complicated twist to his own work by collecting information with FRET. In this technique, an excited donor dye molecule transfers energy to an acceptor molecule, providing information about their interaction. To bring the two molecules into proximity, Ha, Weiss, and their colleagues worked with Berkeley scientist Paul Selvin, who had successfully attached dye molecules to segments of single-stranded DNA (ssDNA). When complementary strands hybridized, the two dye molecules were in proximity to each other.

In a proof-of-concept experiment, Ha and his colleagues used the technique to measure FRET between a single pair of donor and acceptor molecules, tetramethylrhodamine and Texas red. The researchers published their work in 1996 (1).

Four Focuses

Ha published several related papers before completing his doctoral degree in 1996. “I ended up doing pretty well in terms of publications,” he recalls. “You could publish fairly simple experiments because everything in this area was so new.”

Ha remained in Weiss and Chemla’s lab an additional year doing postdoctoral work and training other graduate students in the methods he used. In 1997, he accepted a second postdoctoral fellowship in the lab of Steven Chu at Stanford University. Days after Ha joined the new group, Chu won the Nobel Prize in Physics. Chu’s award and additional responsibilities left Ha and other lab members with unexpected free reign over their projects. Though Ha had initially been charged with using an atomic force microscope to measure conformational changes in enzymes, he decided to add his own twist based on his previous research: he combined atomic force microscopy with FRET, using quantum dots applied to the microscope’s tip as donor molecules and attaching acceptor molecules to his samples. The technique had only limited success, but the fluorescent part of the instrument was useful for observing the folding of single RNA molecules, which he studied in collaboration with James Williamson at the Scripps Research Institute. Ha published his first paper in Chu’s lab on RNA folding in 1999 (2).

After this publication, Ha searched for a new system on which to build his own independent research. He found an ideal candidate in helicase, an enzyme that unwinds double-stranded DNA (dsDNA). After reading a special review issue of *Cell* on motor proteins, Ha realized that helicase

was then the only remaining motor protein mentioned in the reviews that had not been studied at the single-molecule level.

Ha began attempting to measure helicase’s unwinding activity using FRET. However, the enzyme stuck to the glass surface of his lab dishes, hindering its study. Collaborators in Stanford’s chemistry department, including Chris Chidsey, suggested using a polymer brush to prevent adhesion. The problem was solved, giving Ha the impetus to move his work to an independent lab.

“There were so many things I wanted to study that basically I couldn’t do it on my own. I figured it was time to start my own group and hire people to work on all these different ideas,” he says.

In 2000, Ha accepted an offer for an assistant professorship at the University of Illinois at Urbana–Champaign, where he developed four broad areas on which to focus his research: the enzyme helicase, a DNA recombination intermediate called the Holliday junction, a recombination-mediating protein called RecA, and a class of proteins called SNARE, an acronym for soluble NSF attachment receptor, that mediate membrane fusion.

Single Molecules in Action

Ha realized that researchers knew little about how helicase unwinds dsDNA. He and his colleagues, including collaborator Tim Lohman of Washington University, were curious about whether the molecule worked as a monomer to unwind DNA or whether only a dimer could complete the task. Using FRET, the researchers investigated this problem. They published compelling evidence in 2002 suggesting that helicase was only active as a dimer. The team’s research showed that if one monomer of the pair dissociates from DNA, unwinding stalls (3).

Recently, Ha’s team published another paper outlining a quirk in helicase’s mechanism (4). Using FRET again, the researchers

found that if helicase is obstructed in its path along duplex DNA, then the molecule will quickly snap back to its original location to repeat its sweep over the same stretch of DNA. Ha says that his team is not sure of the biological significance of this activity. However, their hypothesis is that helicase could be performing an essential housekeeping function for DNA.

“Helicase moving on a single strand of DNA many times could be used to prevent the accumulation of unwanted proteins on a DNA segment. Like a snowplow, it continues to sweep over and over again,” explains Ha.

About a third of his lab members continue to study helicase. Another system that Ha and his colleagues are investigating is the Holliday junction, a four-stranded DNA molecule that forms as an intermediate during recombination. Researchers had known that this structure can have two alternatively folded conformations and that it could convert between these two states. However, notes Ha, scientists had not been able to observe this conversion in real time. In collaboration with David Lilley of the University of Dundee in Scotland, Ha and his colleagues followed the molecule’s changes between isomers, research they published in 2003 (5). More recently, the team has used optical tweezers to further investigate the junction’s conformational changes under gentle tension.

A third project on which Ha and his team are focused involves RecA, a protein that catalyzes pairing of ssDNA with complementary regions of dsDNA during recombination. This protein filament is highly dynamic in cells, growing and shrinking by adding or subtracting RecA monomers as it accomplishes its task. Using FRET, Ha’s group investigated this process. The researchers found that about five monomers are necessary to start the growth of a RecA filament. Additionally, their work suggested that this filament grows and shrinks one monomer at a time (6).

“Our measurements for the first time gave concrete numbers for this highly dynamic process,” says Ha.

He and his team recently began studying the SNARE family of proteins, a group thought to play pivotal roles in the membrane fusion necessary for exocytosis and vesicle trafficking in the neural junction and elsewhere. The researchers were interested in looking at the proteins’ action at the single-vesicle level. Ha’s team developed an assay in which they created vesicles with donor or acceptor fluorescent dyes incorporated into their membranes. When these donor and acceptor vesicles mix, they emit a FRET signal.

Using this assay, Ha’s team, in collaboration with Yeon-Kyun Shin of Iowa State University, has found that vesicles do not fuse in a single step. Rather, the SNARE proteins guide vesicle membranes through several intermediates (7). Their initial findings used yeast SNARE proteins as a model. However, Ha notes that his team’s current work focuses on the proteins’ action in neurons.

The Future of FRET

Though he and his colleagues have made important contributions to the four systems he initially set out to study, Ha notes that his work is far from complete. He is cultivating several ideas for future experiments using FRET and other techniques.

“So far, we’ve been doing relatively simple experiments for single-molecule measurements, such as the interaction between one protein and one DNA. In terms of the future, we’ll have to make the system more relevant biologically, because in the cell, proteins don’t function in isolation,” he says. “There are interactions between multiple components.”

He adds that his group is planning to make FRET even more useful by developing techniques that involve multiple colors of fluorescent dyes. He also plans on continuing to push the boundaries

of his research through incorporating optical tweezers and other technologies. Additionally, he would like to study other biological systems, such as ribosomes and chromatin remodeling.

Ha notes that this combination of unique techniques and a novel focus will probably continue to turn up interesting aspects of biological phenomena. “I’m interested in studying things that will make a difference—problems or approaches that are not being pursued elsewhere,” he says. “This way, hopefully whatever we discover will be quite unique in terms of contribution to the field.”

—Christen Brownlee, Science Writer

REFERENCES

1. Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996) Probing the interaction between two single molecules—fluorescence resonance energy transfer between a single donor and a single acceptor, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6264–6268.
2. Ha, T., Zhuang, X., Kim, H. D., Orr, J. W., Williamson, J. R., and Chu, S. (1999) Ligand-induced conformational changes observed in single RNA molecules, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9077–9082.
3. Ha, T., Rasnik, I., Cheng, W., Babcock, H. P., Gauss, G. H., Lohman, T. M., and Chu, S. (2002) Initiation and reinitiation of DNA unwinding by the *Escherichia coli* Rep helicase, *Nature* **419**, 638–641.
4. Myong, S., Rasnik, I., Joo, C., Lohman, T. M., and Ha, T. (2005) Repetitive shuttling of a motor protein on DNA, *Nature* **437**, 1321–1325.
5. McKinney, S. A., Declais, A. C., Lilley, D. M. J., and Ha, T. (2003) Structural dynamics of individual Holliday junctions, *Nat. Struct. Biol.* **10**, 93–97.
6. Joo, C., McKinney, S. A., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006) Real-time observation of RecA filament dynamics with single monomer resolution, *Cell* **126**, 515–527.
7. Yoon, T. Y., Okumus, B., Zhang, F., Shin, Y. K., and Ha, T. (2006) Multiple intermediates in SNARE-induced membrane fusion, *Proc. Natl. Acad. Sci. U.S.A.*, in press.

Seeing Is Believing

Valérie Jacquier and Stephen W. Michnick*

Département de Biochimie, Université de Montréal, C.P. 6128, succursale centre-ville, Montréal, Québec H3C 3J7, Canada

ABSTRACT Modern visualization techniques are affording a peek into complex cellular processes. A recent paper describes an automated fluorescence microscopy method to map the subcellular localization of up to 100 different proteins in the same sample.

In their thoughtful commentary, Kim Morrison and Greg Weiss (1) describe how the origins of modern chemical biology can be traced, in part, to the pioneering applications of dyes and chemical processes to resolve the form and function of tissues and cells. Visualizing the structures of cells and the components that bind to chemicals has been the source of inspiration for so many discoveries: Anna Atkins' use of the cyanotype process developed by John Herschel to image botanical specimens, the synthesis of the first example of aniline dyes by William Perkins and their subsequent use by Rudolf Virchow to uncover details of cellular structure and lineage, and Paul Ehrlich's crucial insights that led to the receptor concept, among many others. We've come a long way since these pioneering works. Now, we can visualize the physical locations and associations of individual proteins or complexes of proteins with fluorescent dyes conjugated to artificially created antibodies, thus achieving sensitivities, spatial resolutions, and specific information that Ehrlich and his contemporaries could only dream of.

In a world in which we are bombarded with huge protein data sets, we must remember that where proteins are expressed, in what forms, and what other proteins they interact with can provide the deepest insights into their functions. However, can we bridge the gap between the massive data analysis afforded by contemporary analytical techniques such as mass spectrometry and the rich but usually limited data provided by cellular imaging? In a recent article in *Nature Biotechnology*,

Schubert and colleagues (2) describe a significant step in this direction, with an automated fluorescence microscopy method to map the subcellular localization of up to 100 different proteins in the same sample, an approach that could be used to unravel complete protein networks. This technology enables one to monitor when and where proteins are found together or excluded from a complex and how protein organization is altered in diseased cells. Further, they demonstrate the potential of this approach for facilitating the development of diagnostics and targeted drugs.

Currently available imaging technologies enable the detection of a limited number (~10) of different fluorophores in a sample (3, 4). Schubert and colleagues (2) expand the number of proteins that can be simultaneously investigated by using one fluorophore attached to many different monoclonal antibodies (Figure 1). They visualize one protein after another in a fixed tissue or cell sample by sequentially adding, imaging, and photobleaching each antibody-conjugated fluorophore. The images are then superimposed to obtain a protein colocalization map. This technology, which the authors call multi-epitope ligand cartography (MELC) technology, has the advantage of not requiring a detailed analysis for separating the contributions from individual, simultaneously present fluorophores.

To estimate how many particular proteins are present at each position in the sample, the authors use appropriate threshold values to determine whether each pixel contains each protein epitope. This leads to the creation, for each pixel, of a vector they call

*Corresponding author,
stephen.michnick@umontreal.ca.

Published online December 15, 2006

10.1021/cb600461v CCC: \$33.50

© 2006 by American Chemical Society

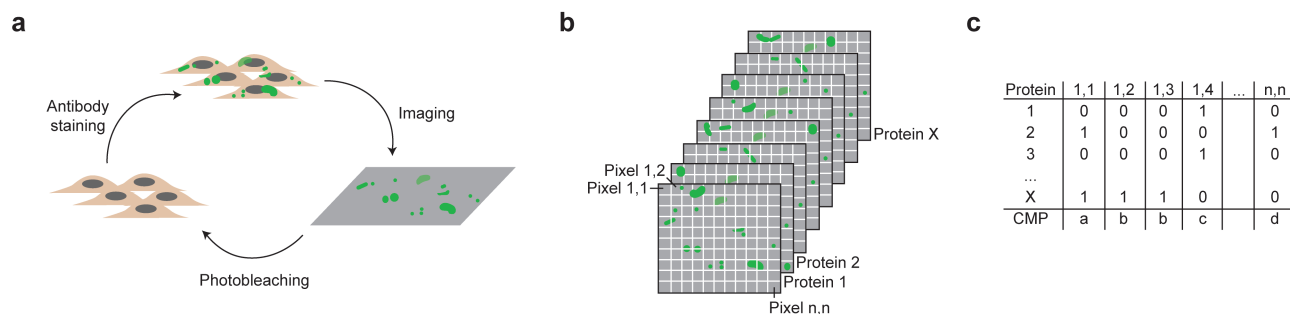


Figure 1. MELC technology. a) A fixed cell or tissue sample is subjected to sequential cycles of staining with a fluorescent monoclonal antibody, imaging, and photobleaching. b) The proteins are imaged one after another, and the images are superimposed to obtain a protein colocalization map. c) Each pixel can be represented as a vector of 0s and 1s to indicate whether it contains a particular protein. Each unique protein combination corresponds to a particular CMP. Adapted with permission from Macmillan, copyright 2006 (12).

combinatorial molecular phenotype (CMP), which represents the combination of proteins found at that pixel. This vector has a length corresponding to the number of proteins and is composed of 0s and 1s (1 if the signal for that protein is higher than the threshold value, 0 if it is lower). The authors show that although the protein distributions vary from cell to cell within a tissue, it is possible to distinguish between healthy and diseased tissues by comparing the distributions of CMPs in these tissues. For example, the abundance of a particular CMP motif in skin tissues can discriminate between tissues with psoriasis or atopic dermatitis and healthy tissues. The MELC technology thus provides a tool for the identification of candidate target proteins in disease.

Furthermore, to show that this methodology can also be used to decipher functional protein networks in a single cell, the authors monitored the colocalization and antilocalization of 23 different cell surface proteins in a spontaneously migrating rhabdomyosarcoma cell line called TE671. They found that four cell surface CMPs were present throughout the migration process and that the alanine-specific protease aminopeptidase N (APN) belonged to these four CMPs, an indication that this protein plays an important role in cellular migration. Indeed, after selective inhibition of APN activity, the cells did

not enter the migratory state and the distribution of the CMPs was affected. Therefore, MELC can be used to identify key proteins in functionally relevant CMP motifs and to selectively inhibit them in order to analyze all functional protein networks.

With this new technology, the number of proteins that can be simultaneously investigated in a tissue or cell sample is limited only by the amount of available fluorescently tagged antibodies (or other fluorescent ligands). However, as usual, caveats exist: antibody labeling is generally highly specific, but parallel methods should be used to validate the accuracy of protein recognition. This is especially true in the case of diseased tissues, because antibodies have the tendency to non-specifically bind to dead cell components. Similarly, the large size of an antibody (typically 150 kDa) might prevent the recognition of neighboring proteins inside protein complexes because of steric hindrance. According to the authors, this limitation could be overcome by carefully determining the proper sequence of labeling cycles.

The MELC technology is a versatile method for unraveling the complexity of protein networks in fixed samples. Live cell imaging of the spatial and temporal dynamics of individual proteins and protein–protein interactions can be achieved with

genetic fusion of reporter fluorescent proteins or other tagging technologies. For example, GFP variants or the use of small fluorescent molecules that can be covalently attached to genetically modified proteins can allow for the real-time analysis of protein localization and trafficking in living cells, as well as the monitoring of protein–protein interactions by FRET (5–7). Similarly, protein fragment complementation assays can monitor the changes in protein complex levels and localization in living cells in response to chemical, genetic, or environmental perturbations (8, 9). Obviously, these techniques all require modifications of the proteins under investigation and cannot be done with living tissues, certainly not on the scale achieved with MELC in single cells, nor can they directly capture subtle changes in protein states, such as post-translational modifications and their causal links to cellular perturbations (10). Thus, despite great efforts to develop genetically encoded reporters of protein function, the best mileage can still be found with good old antibodies.

Schubert and colleagues (2) describe efforts to develop MELC into an impressive industrial-level process. However, it is reasonable to imagine doing the same thing in your basic mom-and-pop operation. The central limitation of their approach is the

availability of specific antibodies. Universities and funding agencies, in collaboration with industry, should focus resources on generating these crucial reagents and coordinate efforts to make them available to the scientific community at a reasonable cost. Further, efforts to devise alternative affinity reagents, including single-chain antibodies and protein alternatives to antibodies, should continue to be supported (11). Imagine a world in which researchers could request a set of affinity reagents for proteins and their different states and have them synthesized on demand. In this world, visualizing the organization of cellular networks and how they evolve in time and space could become a standard tool for building models of cellular machinery. This might seem far-fetched at the moment, but as Schubert and colleagues (2) have shown us, seeing is believing.

Acknowledgment: V.J. acknowledges a Swiss National Science Foundation postdoctoral fellowship. S.W.M. is the Canada Research Chair in Integrative Genomics.

REFERENCES

- Morrison, K. L., and Weiss, G. A. (2006) The origins of chemical biology, *Nat. Chem. Biol.* 2, 3–6.
- Schubert, W., Bonnekoh, B., Pommer, A. J., Philipsen, L., Bockelmann, R., Malykh, Y., Gollnick, H., Friedenberger, M., Bode, M., and Dress, A. W. (2006) Analyzing proteome topology and function by automated multidimensional fluorescence microscopy, *Nat. Biotechnol.* 24, 1270–1278.
- De Rosa, S. C., Herzenberg, L. A., Herzenberg, L. A., and Roederer, M. (2001) 11-Color, 13-parameter flow cytometry: identification of human naive T cells by phenotype, function, and T-cell receptor diversity, *Nat. Med.* 7, 245–248.
- Perfetto, S. P., Chattopadhyay, P. K., and Roederer, M. (2004) Seventeen-colour flow cytometry: unravelling the immune system, *Nat. Rev. Immunol.* 4, 648–655.
- Chen, I., and Ting, A. Y. (2005) Site-specific labeling of proteins with small molecules in live cells, *Curr. Opin. Biotechnol.* 16, 35–40.
- Giepmans, B. N., Adams, S. R., Ellisman, M. H., and Tsien, R. Y. (2006) The fluorescent toolbox for assessing protein location and function, *Science* 312, 217–224.
- Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat. Chem. Biol.* 1, 13–21.
- MacDonald, M. L., Lamerdin, J., Owens, S., Keon, B. H., Bilter, G. K., Shang, Z., Huang, Z., Yu, H., Dias, J., Minami, T., Michnick, S. W., and Westwick, J. K. (2006) Identifying off-target effects and hidden phenotypes of drugs in human cells, *Nat. Chem. Biol.* 2, 329–337.
- Remy, I., and Michnick, S. W. (2001) Visualization of biochemical networks in living cells, *Proc. Natl. Acad. Sci. U.S.A.* 98, 7678–7683.
- Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A., and Nolan, G. P. (2005) Causal protein-signaling networks derived from multiparameter single-cell data, *Science* 308, 523–529.
- Sidhu, S. S., and Fellouse, F. A. (2006) Synthetic therapeutic antibodies, *Nat. Chem. Biol.* 2, 682–688.
- Murphy, R. F. (2006) Putting proteins on the map, *Nat. Biotechnol.* 24, 1223–1224.

Cryptophycin Anticancer Drugs Revisited

Jürgen Rohr*

Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, Kentucky 40536-0082

Imagine a new biotechnologically producible anticancer drug of marine origin with a tubulin-destabilizing mechanism of action that is not negatively affected by P-glycoprotein, a commonly used efflux system of resistant cancer cell lines. This would add another unique natural product into the armamentarium of cancer chemotherapeutics and justify the tremendous efforts that have been invested into marine drug discovery efforts over the past decade or so. The cryptophycins (*e.g.*, cryptophycin 1; Figure 1, panel a) are among the most promising candidates for such a new drug. Like many other natural products, several of marine origin, the cryptophycins interfere with the dynamics of tubulin polymerization and depolymerization. This activity has a long history. Microtubules are dynamic cell structures that play an important role in cell division. Various natural products, such as colchicine, combretastatins, vinca alkaloids, taxanes, epothilones, discodermolides, and dolastatins, bind to tubulin. They eventually cause metaphasic mitotic arrest and consequently apoptotic cell death (1–9). The oldest drug in this context is colchicine. Its damaging effect on tumor vasculature was known in the 1930s, but it was too toxic as an anticancer agent and is now used mostly to treat severe inflammation from gout (10). However, derivatives of the structurally closely related combretastatins A-1 and A-4P were developed and are currently in phase 2 clinical trials as anticancer agents (10). Other tubulin binders, especially several natural (*e.g.*, vincristine and vinblastine) and semisynthetic (*e.g.*, vin-

desine and vinorelbine) vinca alkaloids and some taxanes (*e.g.*, paclitaxel and docetaxel), have a solid place within regimens of anticancer chemotherapeutics. Various epothilones and their derivatives are in clinical trials as well. Epothilones are easily produced by fermentation and are not affected by P-glycoprotein, so they are favorable over the well-established drug Taxol (paclitaxel) (1, 2). Colchicine, combretastatins, vinca alkaloids, and cryptophycins inhibit the polymerization of tubulin, destabilize the microtubules, and eventually cause metaphasic arrest, whereas taxanes, epothilones, and discodermolides follow a different mechanism of action: they stabilize the microtubules and thus cause misarrangements. This eventually leads to the same overall effect: mitotic arrest and cell death. Interestingly, some recently developed colchicine glycosides obtained through a glycorandomization study (*e.g.*, Col19) seemed to act as microtubule stabilizers; this shows that chemical derivatization can change the principal mechanism of action of a tubulin-binding drug (11). Figure 1 shows examples of tubulin-binding natural products and semisynthetic derivatives.

When the cryptophycins, a group of >25 cyanobacterial metabolites with strong tubulin-destabilizing activities (12–14), were discovered, hopes were great that one of these natural products could be developed into a useful anticancer drug. In fact, the prototype cryptophycin 1, the major representative of this class of natural anticancer drugs from the cyanobacterial symbiont

ABSTRACT A recent publication reveals the biosynthetic building blocks, genetic code, and broad substrate tolerance of the enzymes of the cryptophycin biosynthetic pathway. This work lays the foundation for the production of poorly accessible yet very promising members of this family of anticancer compounds from lichen cyanobacterial symbionts. Chemoenzymatic production or precursor-directed biosynthesis might bring candidates from this family of natural products back to clinical trials.

*Corresponding author,
jrohr2@email.uky.edu.

Published online December 15, 2006

10.1021/cb6004678 CCC: \$33.50

© 2006 by American Chemical Society

The scientific results might provide the basis for the biotechnological development of promising cryptophycin analogues.

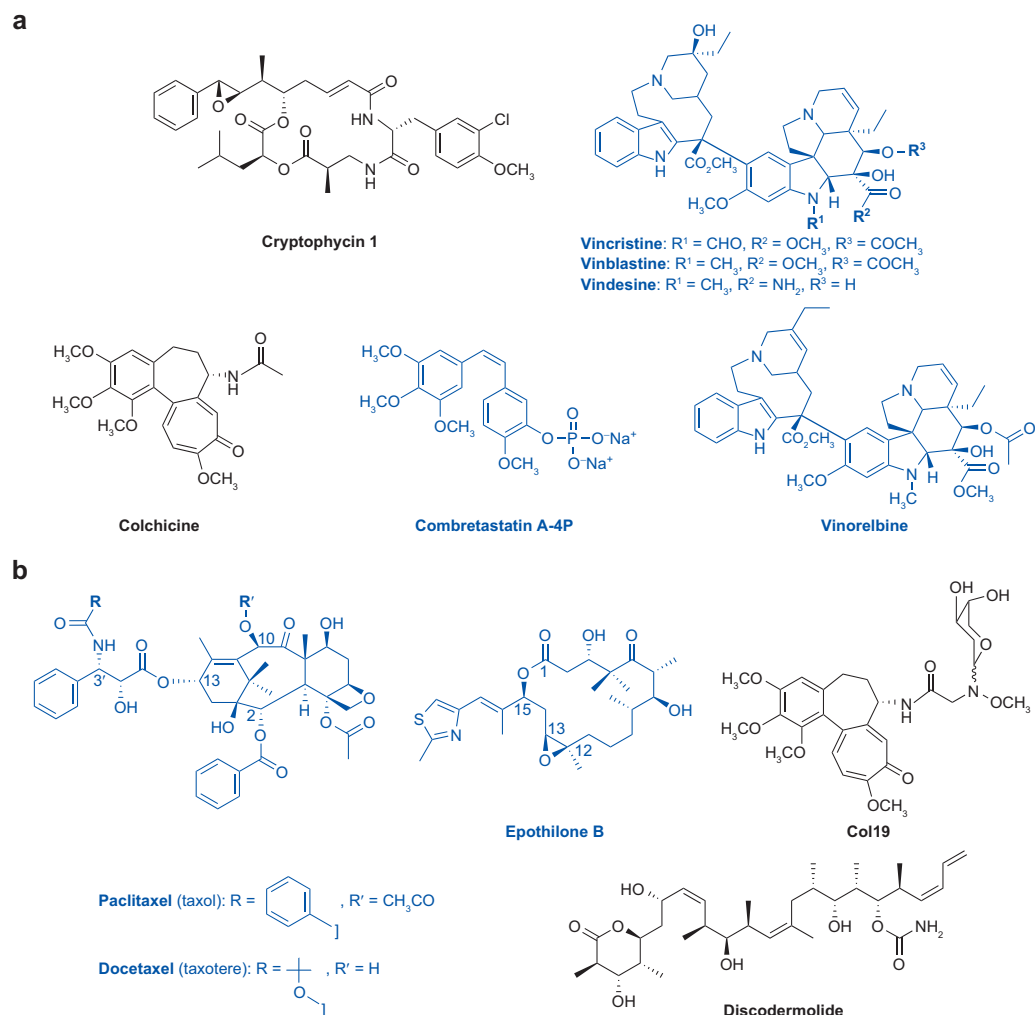


Figure 1. Chemical structures of selected examples of natural product drugs or their semisynthetic derivatives that interfere with the dynamics of tubulin polymerization and depolymerization. a) Tubulin-destabilizing drugs. b) Tubulin stabilizers. Drugs that are currently used as clinical anticancer drugs or that are in clinical trials are depicted in blue.

Nostoc sp. ATCC 53789, is one of the most potent tubulin-destabilizing agents ever found (12). In addition, the cryptophycins, like the epothilones, were not substrates of P-glycoprotein, an efflux pump that makes multidrug-resistant cancer cell lines immune against a multitude of anticancer drugs (4, 5, 12). Consequently, cryptophycin 52 (Figure 2, panel b), a synthetic analogue, was developed and reached phase 2 clinical trials (15–17). The synthetic analogue 52

was chosen because no large-scale biotechnological production method existed for the cryptophycins. Eventually, the high production costs and toxic side effects of cryptophycin 52 stopped its development and that of any other analogue of the cryptophycin family. Nobody wanted to restart all of the trials with a different analogue, although preclinical studies showed that other analogues would have been a better choice (18). Nonetheless, studies to find new cryp-

clinical development. Magarvey *et al.* (22) used a comparative secondary metabolomic analysis to identify the cryptophycin biosynthetic genes (*crp*). This study is unique in that it marks the first time a gene cluster of a natural product has been identified in this way. The cryptophycin gene cluster (Figure 3) covers >40 kb of the *Nostoc* genome. The core of the *crp* gene cluster consists of two modular type-1 polyketide synthase (PKS) genes (*crpA* and *crpB*) and two nonriboso-

tophycin analogues or to develop semisynthetic/biotechnological methods for the generation of promising cryptophycin analogues continued (18–21).

Now, the research group of David H. Sherman of the University of Michigan in collaboration with Richard E. Moore of the University of Hawaii (22) has studied in a very comprehensive way the biosynthesis of the cryptophycins (see p 766 in this issue). The studies cover a wide range of topics, including incorporation experiments, cloning of the gene cluster, important biosynthetic enzymes, and the production of new cryptophycin analogues by precursor-directed biosynthesis. The scientific results might provide the basis for the biotechnological development of promising cryptophycin analogues and thus might reopen the door to their

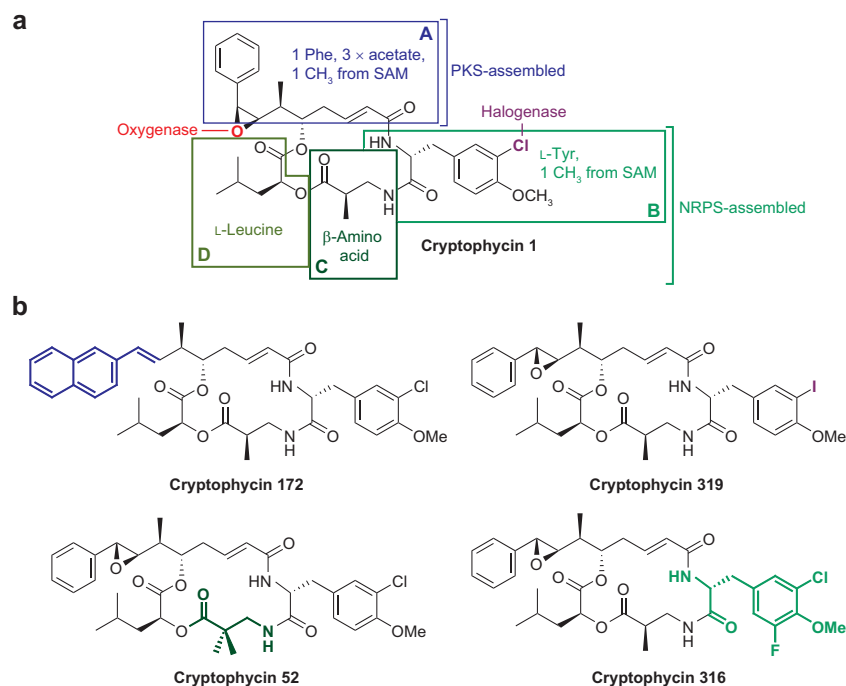


Figure 2. Building blocks of cryptophycins. **a)** Biosynthetic building blocks of subunits A–D of the cryptophycins. **b)** Examples of new cryptophycin analogues obtained by precursor-directed biosynthesis *via* synthetic or natural analogues of the typical building blocks. SAM = *S*-adenosyl-methionine

mal peptide synthetase (NRPS) genes (*crpC* and *crpD*), as one could expect from the chemical structure of the molecule. These four large genes encode multienzyme complexes, which cover six modules of chain elongation and are followed downstream by four small genes encoding post-PKS/NRPS tailoring enzymes (*crpE* to *H*). Interestingly, all genes are assembled with the same reading direction and mirror exactly the biosynthetic assembly-line sequence of events. The most important tailoring enzyme is CrpE, a P450 oxygenase responsible for the epoxide formation found in several cryptophycins, such as cryptophycins 1, 2, and 52. Establishing this epoxide in the desired β -stereochemistry was a challenge and required several synthetic steps (23, 24). The characterization of this enzyme allowed the first stereospecific synthesis of cryptophycin 2 through a chemoenzymatic synthe-

sis *via seco*-cryptophycin 4 involving two enzymes, the thioesterase from CrpD and CrpE.

Furthermore, through incorporation of stable-isotope-labeled precursors, the biosynthetic origin of all four subunits (A–D) of the cryptophycins could be determined (Figure 2, panel a). The substrate flexibility of the biosynthetic enzymes within the assembly line of the PKS/NRPS multienzyme complex of the cryptophycin pathway was proven by precursor-directed biosynthesis that used unnatural starter units, such as various phenylalanine derivatives and analogues. Surprisingly, these artificial amino acids were incorporated not only into the starter unit but also sometimes into unit B. In addition, halogenase CrpH was tested with bromide and iodide ions, and interestingly, it could also incorporate bromide and iodide atoms into unit B. This result is

typical for marine halogenases, because brominated and iodinated natural products are common in the ocean. Finally, unit C could also be altered through the use of suitable β -amino acid precursors. This way, even cryptophycin 52 could be produced, a drug that was previously accessible only through total synthesis and that was chosen for the above-mentioned clinical trials. Overall, the efforts yielded 30 new cryptophycin analogues (for four typical examples, see Figure 2, panel b).

In summary, the work of Magarvey *et al.* (22) applied both established and novel methods and yielded a plethora of very interesting and relevant scientific results for cryptophycin biosynthesis. One could imagine that they have presented a solid base from which cryptophycins can be revisited for clinical use as anticancer agents. First, any of the new analogues produced by precursor-directed biosynthesis may have advantageous properties as clinical drugs. Second, epoxide-containing cryptophycins may be more easily accessible now through a chemoenzymatic approach. Finally, cryptophycin 52 is now biotechnologically producible, which may make this drug more easily accessible and less costly to produce if further pursued.

REFERENCES

- Bergstralh, D. T., and Ting, J. P. (2006) Microtubule stabilizing agents: their molecular signaling consequences and the potential for enhancement by drug combination, *Cancer Treat. Rev.* 32, 166–179.
- Attard, G., Greystoke, A., Kaye, S., and De Bono, J. (2006) Update on tubulin-binding agents, *Pathol. Biol. (Paris)* 54, 72–84.
- Geney, R., Chen, J., and Ojima, I. (2005) Recent advances in the new generation taxane anticancer agents, *Med. Chem.* 1, 125–39.
- Fojo, A. T., and Menefee, M. (2005) Microtubule targeting agents: basic mechanisms of multidrug resistance (MDR), *Semin. Oncol.* 32, S3–S8.
- Breier, A., Barancik, M., Sulova, Z., and Uhrík, B. (2005) P-Glycoprotein—implications of metabolism of neoplastic cells and cancer therapy, *Curr. Cancer Drug Targets* 5, 457–468.
- Pellegrini, F., and Budman, D. R. (2005) Review: tubulin function, action of antitubulin drugs, and new drug development, *Cancer Invest.* 23, 264–273.

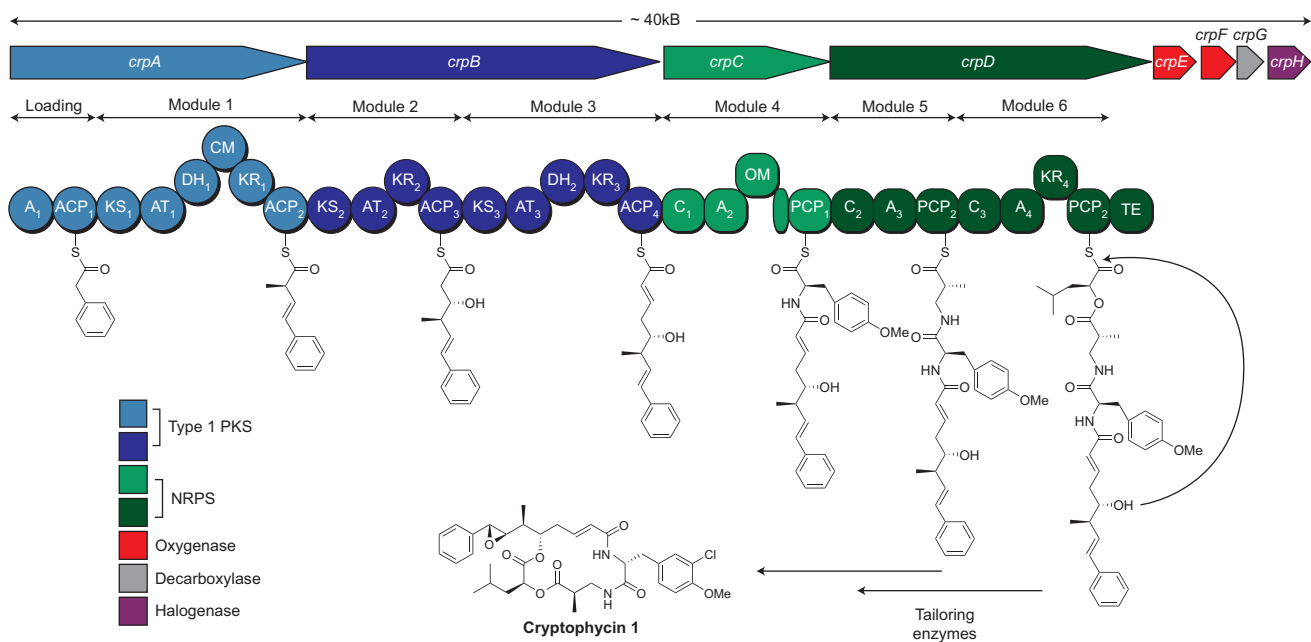


Figure 3. The cryptophycin gene cluster and the biosynthetic assembly line process. Domains of the PKS and NRPS multienzyme complexes: **A** = adenylation domain, **ACP** = acyl carrier protein, **KS** = β -keto synthase, **AT** = acyl transferase, **DH** = dehydrase, **CM** = C-methyl transferase, **KR** = ketoreductase, **C** = condensation domain, **OM** = O-methyl transferase, **PCP** = peptidyl carrier protein, and **TE** = thioesterase.

- Verrills, N. M., and Kavallaris, M. (2005) Improving the targeting of tubulin-binding agents: lessons from drug resistance studies, *Curr. Pharm. Des.* **11**, 1719–1733.
- Zhou, J., and Giannakakou, P. (2005) Targeting microtubules for cancer chemotherapy, *Curr. Med. Chem. Anticancer Agents* **5**, 65–71.
- Kavallaris, M., Verrills, N. M., and Hill, B. T. (2001) Anticancer therapy with novel tubulin-interacting drugs, *Drug Resist. Updates* **4**, 392–401.
- Tozer, G. M., Kanthou, C., and Baguley, B. C. (2005) Disrupting tumour blood vessels, *Nat. Rev. Cancer* **5**, 423–435.
- Ahmed, A., Peters, N. R., Fitzgerald, M. K., Watson, J. A., Jr., Hoffmann, F. M., and Thorson, J. S. (2006) Colchicine glycorandomization influences cytotoxicity and mechanism of action, *J. Am. Chem. Soc.* **128**, 14224–14225.
- Smith, C. D., Zhang, X., Mooberry, S. L., Patterson, G. M., and Moore, R. E. (1994) Cryptophycin: a new antimicrotubule agent active against drug-resistant cells, *Cancer Res.* **54**, 3779–3784.
- Panda, D., Himes, R. H., Moore, R. E., Wilson, L., and Jordan, M. A. (1997) Mechanism of action of the unusually potent microtubule inhibitor cryptophycin 1, *Biochemistry* **36**, 12948–12953.
- Corbett, T. H., Valeriote, F. A., Demchik, L., Lowichik, N., Polin, L., Panchapor, C., Pugh, S., White, K., Kushner, J., Rake, J., Wentland, M., Golakoti, T., Heltzel, C., Ogino, J., Patterson, G., and Moore, R. (1997) Discovery of cryptophycin-1 and BCN-183577: examples of strategies and problems in the detection of antitumor activity in mice, *Invest. New Drugs* **15**, 207–218.
- Sessa, C., Weigang-Kohler, K., Pagani, O., Greim, G., Mora, O., De Pas, T., Burgess, M., Weimer, I., and Johnson, R. (2002) Phase I and pharmacological studies of the cryptophycin analogue LY355703 administered on a single intermittent or weekly schedule, *Eur. J. Cancer* **38**, 2388–2396.
- Edelman, M. J., Gandara, D. R., Hausner, P., Israel, V., Thornton, D., DeSanto, J., and Doyle, L. A. (2003) Phase 2 study of cryptophycin 52 (LY355703) in patients previously treated with platinum based chemotherapy for advanced non-small cell lung cancer, *Lung Cancer* **39**, 197–199.
- D'Agostino, G., del Campo, J., Mellado, B., Izquierdo, M. A., Minarik, T., Cirri, L., Marini, L., Perez-Gracia, J. L., and Scambia, G. (2006) A multicenter phase II study of the cryptophycin analog LY355703 in patients with platinum-resistant ovarian cancer, *Int. J. Gynecol. Cancer* **16**, 71–76.
- Liang, J., Moore, R. E., Moher, E. D., Munroe, J. E., Al-awar, R. S., Hay, D. A., Varie, D. L., Zhang, T. Y., Aikins, J. A., Martinelli, M. J., Shih, C., Ray, J. E., Gibson, L. L., Vasudevan, V., Polin, L., White, K., Kushner, J., Simpson, C., Pugh, S., and Corbett, T. H. (2005) Cryptophycins-309, 249 and other cryptophycin analogs: preclinical efficacy studies with mouse and human tumors, *Invest. New Drugs* **23**, 213–224.
- Beck, Z. Q., Aldrich, C. C., Magarvey, N. A., Georg, G. I., and Sherman, D. H. (2005) Chemoenzymatic synthesis of cryptophycin/arenastatin natural products, *Biochemistry* **44**, 13457–13466.
- Chaganty, S., Golakoti, T., Heltzel, C., Moore, R. E., and Yoshida, W. Y. (2004) Isolation and structure determination of cryptophycins 38, 326, and 327 from the terrestrial cyanobacterium *Nostoc* sp. GSV 224, *J. Nat. Prod.* **67**, 1403–1406.
- Shih, C., and Teicher, B. A. (2001) Cryptophycins: a novel class of potent antimitotic antitumor depsipeptides, *Curr. Pharm. Des.* **7**, 1259–1276.
- Magarvey, N. A., Beck, Z. Q., Golakoti, T., Ding, Y., Huber, U., Hemscheidt, T. K., Abelson, D., Moore, R. E., and Sherman, D. H. (2006) Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins, potent anti-cancer agents from *Nostoc* cyanobionts, *ACS Chem. Biol.* **1**, 766–779.
- Liang, J., Moher, E. D., Moore, R. E., and Hoard, D. W. (2000) Synthesis of cryptophycin 52 using the Sharpless asymmetric dihydroxylation: diol to epoxide transformation optimized for a base-sensitive substrate, *J. Org. Chem.* **65**, 3143–3147.
- EGgen, M., and Georg, G. I. The cryptophycins: their synthesis and anticancer activity, *Med. Res. Rev.* **22**, 85–101.

Mix-and-Match Riboswitches

Colby D Stoddard and Robert T. Batey*

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Campus Box 215, Boulder, Colorado 80309-0215

Regulation of gene expression is a tightly controlled process that allows organisms to adapt to a continuously fluctuating physical and chemical environment. Activation or repression of genes promoting specific responses to changing conditions must occur in a precise and efficient manner. Until recently, protein factors have dominated the known gene expression regulatory strategies; however, RNA-based control (also called riboregulation) is now shown to play a significant role in all domains of life (1–3). One form of riboregulation, called riboswitches, is broadly distributed within the 5'-untranslated region of bacterial messenger RNAs (mRNAs) and controls transcription or translation *via* direct binding of small-molecule metabolites or metal ions (reviewed in ref 3). This control element consists of two functional components, a highly structured ligand-binding aptamer domain that directly binds a specific ligand followed by an expression platform that directly controls the expression of downstream genes (Figure 1, panel a).

In transcriptional control, transduction of intracellular ligand concentration into changes in gene expression is achieved through the interplay of mutually exclusive RNA structures (Figure 1, panel a, orange boxes) that direct RNA polymerase to continue or abort mRNA synthesis. Structural analysis of the purine (4, 5), thiamine pyrophosphate (TPP) (6, 7), and S-adenosylmethionine (SAM) riboswitch (8) aptamer domains revealed a common mechanism by which this occurs. Communication between

the aptamer domain and the expression platform is mediated by a sequence at the 3'-end of a P1 helix (Figure 1, panel a, red) that joins the 5'- and 3'-ends of the aptamer domain, a feature common to most riboswitches. Ligand binding establishes an intricate network of hydrogen bonds that stabilize formation of the P1 helix at the expense of one of two mutually exclusive secondary structures in the expression platform. In the structure of the SAM riboswitch (8), this is achieved by encapsulation of the ligand to form a series of interactions between the 3'-end of the P1 helix and other elements of the RNA (Figure 1, panel b). Direct sensing of metabolites and the absence of protein involvement allow riboswitch-mediated regulation of gene expression to be an economical and fast-reacting means of regulation (9).

In their simplest form, riboswitches appear to function through feedback inhibition with control effected by a single ligand. However, recent studies have uncovered different arrangements of the basic riboswitch, presumably to create more sophisticated responses to ligands. The first example of this was found in an mRNA containing two complete TPP riboswitches positioned tandemly upstream of the *thiSGHFE* operon in several *Desulfovibrio* species (10) (Figure 2, panel a). The result of this arrangement is likely that gene expression is influenced by each riboswitch in an independent but additive fashion. Cooperative ligand binding, on the other hand, is observed in the glycine riboswitch (11). This regulatory element consists of two aptamer domains influencing a single

ABSTRACT Riboswitches are noncoding RNA elements found in the 5'-untranslated region of messenger RNA (mRNA) that mediate gene expression in a *cis* fashion in the absence of protein. This common regulatory strategy in bacteria is achieved through the interplay of two distinct domains: an aptamer domain responsible for sensing intracellular concentrations of a specific metabolite and a domain containing a secondary structural switch directly controlling expression. In a recent study, riboswitches have been discovered that are capable of regulating transcription by using an RNA architecture mimicking a Boolean NOR logic gate. Tandem arrangement of elements that recognize S-adenosylmethionine and coenzyme B₁₂ yields an mRNA that is only expressed when both metabolites are in low concentration in the cell.

*Corresponding author,
robert.batey@colorado.edu.

Published online December 15, 2006

10.1021/cb600458w CCC: \$33.50

© 2006 by American Chemical Society

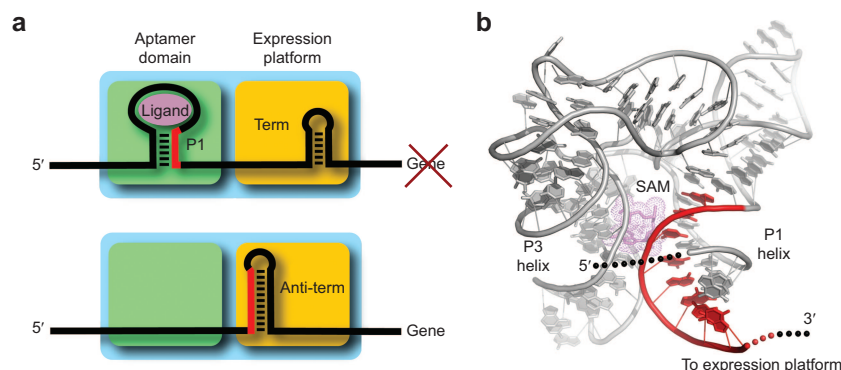


Figure 1. The basic form of the riboswitch. **a)** The basic riboswitch design consisting of an aptamer domain that binds ligand and an expression platform that directs gene expression. The “switching sequence” (shown in red) primarily directs secondary structural switching between the terminator (Term) and the anti-terminator (Anti-term) elements. **b)** The tertiary structure of the SAM riboswitch (PDB code 2GIS) emphasizing the relationship between ligand binding and stabilization of the switching sequence in the P1 helix.

expression platform downstream of the second aptamer (Figure 2, panel b). Glycine binding is a cooperative event in which binding of glycine to the first domain promotes binding to the second, serving to create a more digital regulation in which the on/off switch occurs over a narrower change in ligand concentration. From a structural perspective, it remains unclear how these domains communicate, whether by a limited secondary structure rearrangement or through interacting tertiary structures similar to protein subunit–subunit interactions.

A further advance in our understanding of the complexity of riboswitches emerges from a recent study from the Breaker laboratory in which two heterogeneous tandem

riboswitches were identified (12). Using bioinformatics methods, the authors discovered that the 5′-untranslated region of the *metE* gene in six different isolates of *Bacillus clausii* was controlled by both the SAM (13, 14) and adenosylcobalamin (AdoCbl) riboswitches (15). In this gram-positive bacterium, conversion of homocysteine to methionine by the MetE protein requires the methyl-group donor methyltetrahydrofolate, whereas MetH utilizes methylcobalamin (MeCbl) to perform the same function in a more energy-efficient manner. It was proposed that because the MetE and MetH isozymes have redundant function it may be advantageous to repress the less-efficient MetE enzyme when sufficient concentrations of MetH coenzyme are present

(16–18). MeCbl is derived from AdoCbl, so it is clear why AdoCbl can exert control over the *metE* locus. Control of *metE* by SAM is a form of feedback inhibition because methionine produced by MetE and MetH can be subsequently converted to SAM. This is similar to the switching between MetE and MetH in other bacteria, which is controlled by the MetR activator that indirectly senses AdoCbl and the MetJ repressor that directly senses SAM (19). Thus, the same strategy is achieved with either RNA- or protein-based regulatory elements.

Independent yet coupled regulation of *metE* by SAM and AdoCbl riboswitches is achieved through a linear arrangement of two complete riboswitches (Figure 3). *In vitro* in-line probing experiments demonstrate that each ligand influences only the structure associated with its respective riboswitch and that they do not influence each other’s affinity for the mRNA (12). Further experiments using *in vivo* approaches strongly support this mechanism. The tandem riboswitch architecture allows *metE* expression to be controlled as a natural Boolean NOR logic gate in which distinct ligand inputs can independently control gene repression (Figure 3), as long as the ligand is present in saturating amounts. In general, an output that is negatively controlled (*i.e.*, turned off) by the presence of one OR another of two signals is a Boolean NOR logic gate, also termed a joint denial gate. Thus, the SAM–AdoCbl tandem riboswitch generates an output of turning off gene

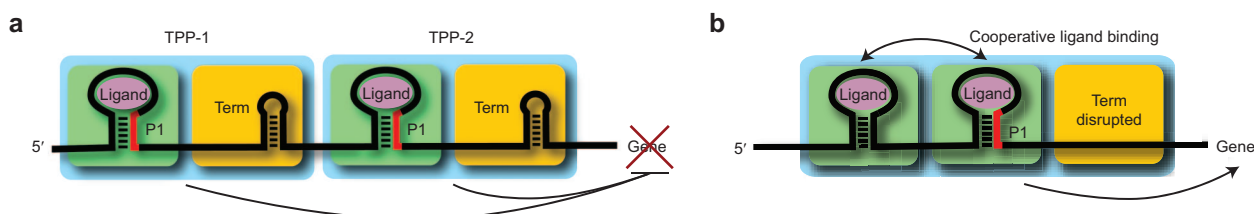


Figure 2. Arrangement of tandem riboswitches. **a)** Two independent TPP riboswitches are tandemly arranged to yield additive control. **b)** Cooperative ligand binding as observed in the glycine riboswitch mediated by two nearly identical aptamer domains that communicate with each other as well as the expression platform. This is a rare example of ligand binding activating gene expression.

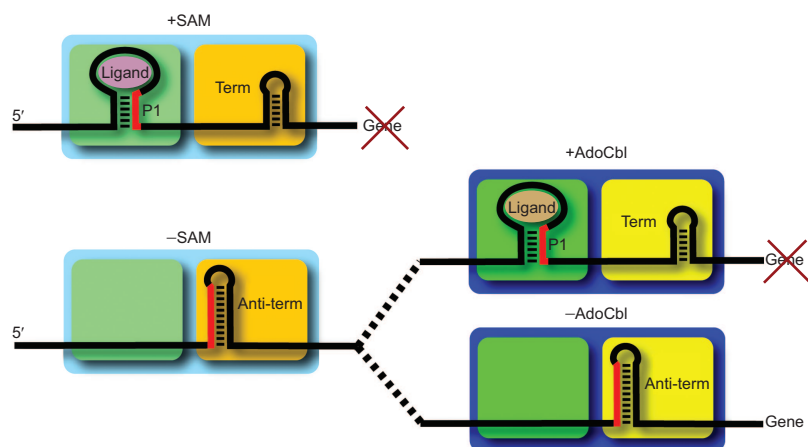


Figure 3. The riboswitch NOR gate. The presence of SAM and/or AdoCbl always results in the off state, whereas only the absence of both results in gene expression being turned on.

expression if either SAM or AdoCbl binds its respective aptamer domain. This condition of independent control by each riboswitch is achieved only at either extreme, high or low, in ligand concentrations. A more physiologically relevant scenario is that each ligand fluctuates around an intermediate concentration range, a reflection of the observation of synergistic control *in vivo* (12).

Natural heterogeneous tandem architecture is not surprising in light of *in vitro* generated RNAs that are capable of multi-input ribozyme activation (20). Tandem arrangement of the *in vitro* selected theophylline and flavin mononucleotide (FMN) aptamers yielded an RNA in which the binding of theophylline significantly promotes FMN binding. Cooperative ligand binding is subsequently detected *via* activation of an attached self-cleaving hammerhead ribozyme. This linear arrangement of aptamers was the first direct demonstration of an RNA-based Boolean logic gate, an AND gate in this case, in which two distinct ligands are required for an output (ribozyme cleavage). These experiments, along with natural tandem riboswitches, suggest that a large repertoire may exist of complex RNA-based

control elements that are yet to be discovered.

As more bacterial genomes are sequenced, we are certain to observe even stranger arrangements of riboswitch elements yielding increasingly sophisticated regulatory responses. In the same paper by Breaker and co-workers, a novel riboswitch is presented that is likely responsive to guanine and another unknown effector molecule in a yet-to-be-defined fashion (12). This mRNA element, found in *Thermoanaerobacter tengcongensis*, controls the *ykkCD* operon that encodes a multi-drug-resistance efflux pump of broad specificity (21). In this RNA, the guanine aptamer domain is predicted to either bind guanine/hypoxanthine to turn off expression or, in the absence of ligand, pair with sequences in the terminator stem to allow transcription (Figure 4, top and middle). Complicating this, however, is the presence of a second conserved stem-loop structure, the *ykkC*

aptamer, which lies between the guanine aptamer and the expression platform. This element is speculated to override the decision-making process of the guanine riboswitch, creating a Boolean “implication gate”. This override gate would activate gene expression if the guanine is high, normally an off situation (Figure 4, middle), and the unknown *ykkC* aptamer ligand is high (Figure 4, bottom). Similar to observations in the glycine riboswitch, the lack of an expression platform separating the guanine and *ykkC* aptamer also may be evidence that the aptamers display some form of cross talk concerning the state of ligand binding. Little direct evidence exists for this hypothesis, so it is still possible that this riboswitch regulates gene expression *via* another mechanism.

The implication of these latest discoveries is that like proteins, RNA can exploit both the 1D and 3D properties of the macromolecule to create new functions out of existing modules. It has long been established that a standard means of readily generating proteins with new biological functions is to mix

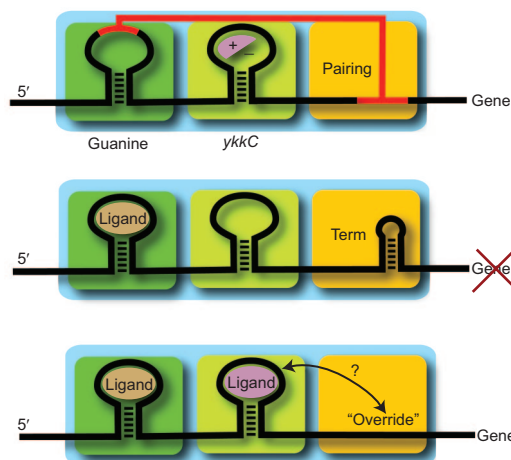


Figure 4. The riboswitch implication gate. In the absence of guanine, regardless of the *ykkC* ligand, the gene is turned on, akin to a standard riboswitch (Figure 1). In the presence of guanine, gene expression is suppressed, also as is standard, except if the *ykkC* ligand is present, which overrides the off state.

and match small, modular domains in novel combinations, akin to “beads on a string”. Similarly, riboswitches can arrange different patterns of either aptamer domains or a completely functional aptamer/expression module to create new regulatory responses that could not be achieved by the simple module alone.

Acknowledgment: This work was supported by a Research Scholar Grant from the American Cancer Society to R.T.B.

REFERENCES

- Gottesman, S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria, *Trends Genet.* **21**, 399–404.
- Mattick, J. S., and Makunin, I. V. (2006) Non-coding RNA, *Hum. Mol. Genet.* **15**, R17–R29.
- Winkler, W. C., and Breaker, R. R. (2005) Regulation of bacterial gene expression by riboswitches, *Annu. Rev. Microbiol.* **59**, 487–517.
- Batey, R. T., Gilbert, S. D., and Montange, R. K. (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine, *Nature* **432**, 411–415.
- Serganov, A., Yuan, Y. R., Pikoyskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R., and Patel, D. J. (2004) Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs, *Chem. Biol.* **11**, 1729–1741.
- Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R., and Patel, D. J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch, *Nature* **441**, 1167–1171.
- Thore, S., Leibundgut, M., and Ban, N. (2006) Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand, *Science* **312**, 1208–1211.
- Montange, R. K., and Batey, R. T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element, *Nature* **441**, 1172–1175.
- Nudler, E., and Mironov, A. S. (2004) The riboswitch control of bacterial metabolism, *Trends Biochem. Sci.* **29**, 11–17.
- Rodionov, D. A., Dubchak, I., Arkin, A., Alm, E., and Gelfand, M. S. (2004) Reconstruction of regulatory and metabolic pathways in metal-reducing delta-proteobacteria, *Genome Biol.* **5**, R90.
- Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W. L., and Breaker, R. R. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression, *Science* **306**, 275–279.
- Sudarsan, N., Hammond, M. C., Block, K. F., Welz, R., Barrick, J. E., Roth, A., and Breaker, R. R. (2006) Tandem riboswitch architectures exhibit complex gene control functions, *Science* **314**, 300–304.
- Grundy, F. J., and Henkin, T. M. (1998) The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria, *Mol. Microbiol.* **30**, 737–749.
- Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) An mRNA structure that controls gene expression by binding S-adenosylmethionine, *Nat. Struct. Biol.* **10**, 701–707.
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA, *Chem. Biol.* **9**, 1043.
- Banerjee, R. V., Frasca, V., Ballou, D. P., and Matthews, R. G. (1990) Participation of cob(I) alamin in the reaction catalyzed by methionine synthase from *Escherichia coli*: a steady-state and rapid reaction kinetic analysis, *Biochemistry* **29**, 11101–11109.
- Gonzalez, J. C., Peariso, K., Penner-Hahn, J. E., and Matthews, R. G. (1996) Cobalamin-independent methionine synthase from *Escherichia coli*: a zinc metalloenzyme, *Biochemistry* **35**, 12228–12234.
- Pejchal, R., and Ludwig, M. L. (2005) Cobalamin-independent methionine synthase (MetE): a face-to-face double barrel that evolved by gene duplication, *PLoS Biol.* **3**, e31.
- Wu, W. F., Urbanowski, M. L., and Stauffer, G. V. (1992) Role of the MetR regulatory system in vitamin B12-mediated repression of the *Salmonella typhimurium metE* gene, *J. Bacteriol.* **174**, 4833–4837.
- Jose, A. M., Soukup, G. A., and Breaker, R. R. (2001) Cooperative binding of effectors by an allosteric ribozyme, *Nucleic Acids Res.* **29**, 1631–1637.
- Jack, D. L., Storms, M. L., Tchieu, J. H., Paulsen, I. T., and Saier, M. H., Jr. (2000) A broad-specificity multidrug efflux pump requiring a pair of homologous SMR-type proteins, *J. Bacteriol.* **182**, 2311–2313.

Structure and Properties of a Re-engineered Homeodomain Protein–DNA Interface

Matthew D. Simon^{†,||}, Morris E. Feldman^{†,||}, Daniel Rauh^{**}, Ann E. Maris[†], David E. Wemmer[†], and Kevan M. Shokat^{†,§,*}

[†]Department of Chemistry, University of California, Berkeley, California 94720, ^{*}Graduate Group in Biophysics, University of California, San Francisco, California 94158, [§]Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, ^{**}Present address: Max Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. ^{||}These authors contributed equally to this work.

The redesign of one member of a highly conserved and large protein family has been useful for the generation of tools to dissect individual protein function in complex biological systems. Protein kinases, for example, contain highly conserved ATP binding pockets and can be studied by introducing a large-to-small mutation at a conserved position in the active site, thereby creating an engineered kinase that is uniquely inhibitable using an enlarged inhibitor that is too bulky to fit into the active site of wild-type kinases (1–3). Similarly, an aspartate-to-asparagine mutation in the active site of a GTPase can switch the nucleotide specificity from GTP to another nucleotide, XTP, allowing the activity of an engineered GTPase to be tracked in the presence of other GTPases (4–7). In these cases, the high sequence conservation of the engineered protein's family allowed extension of an engineering solution from one protein to other members within the family. Although these examples illustrate the utility of engineering protein–small molecule interactions, many proteins exert their regulatory roles in the cell through protein–protein or protein–DNA interactions. Therefore we wondered whether highly conserved, extended interfaces could be re-engineered to incorporate new functionality.

Several extended interfaces have been engineered, including the interacting surfaces between human growth hormone and its receptor (8), zinc fingers and diverse DNA

sequences (9–11), and the protein heterodimerization interface of an endonuclease (12). Perhaps the most dramatic example of an adaptive protein surface is the variable region of an antibody, which can recognize chemically diverse haptens, including those derived from the synthetic elaboration of biomolecules (13). Although these examples highlight the adaptability of protein interfaces, they also demonstrate that protein engineers have focused on interfaces that have a high degree of natural functional variability, and generally the amino acids at these interfaces are poorly conserved. In contrast, to take advantage of the frequent occurrence of highly conserved domains and their interacting surfaces, we are interested in how much adaptability and corresponding potential for re-engineering exists within highly conserved biomolecular interfaces. To address this question, we chose to focus on the homeodomain (HD)–DNA interaction.

The HD is a DNA-binding domain that has been conserved over 500 million years of evolution in both structure and function (14). The 60 amino acid HD is composed of three α helices (Figure 1, panel a), the C-terminal of which is referred to as the recognition helix as a result of base-specific contacts it makes in the major groove of DNA with the sequence TAATXX (Figure 1, panels a and b, in green) (15–18). These conserved contacts include a hydrophobic interaction of either isoleucine or valine at

ABSTRACT The homeodomain (HD)–DNA interface has been conserved over 500 million years of evolution. Despite this conservation, we have successfully re-engineered the engrailed HD to specifically recognize an unnatural nucleotide using a phage display selection. Here we report the synthesis of novel nucleosides and the selection of mutant HDs that bind these nucleotides using phage display. The high-resolution crystal structure of one mutant in complex with modified and unmodified DNA demonstrates that, even with the substantial perturbation to the interface, this selected mutant retains a canonical HD structure. Dissection of the contributions due to each of the selected mutations reveals that the majority of the modification-specific binding is accomplished by a single mutation (I47G) but that the remaining mutations retune the stability of the HD. These results afford a detailed look at a re-engineered protein–DNA interaction and provide insight into the opportunities for re-engineering highly conserved interfaces.

*Corresponding author, shokat@cmp.ucsf.edu.

Received for review August 28, 2006 and accepted November 21, 2006.

Published online December 15, 2006

10.1021/cb6003756 CCC: \$33.50

© 2006 by American Chemical Society

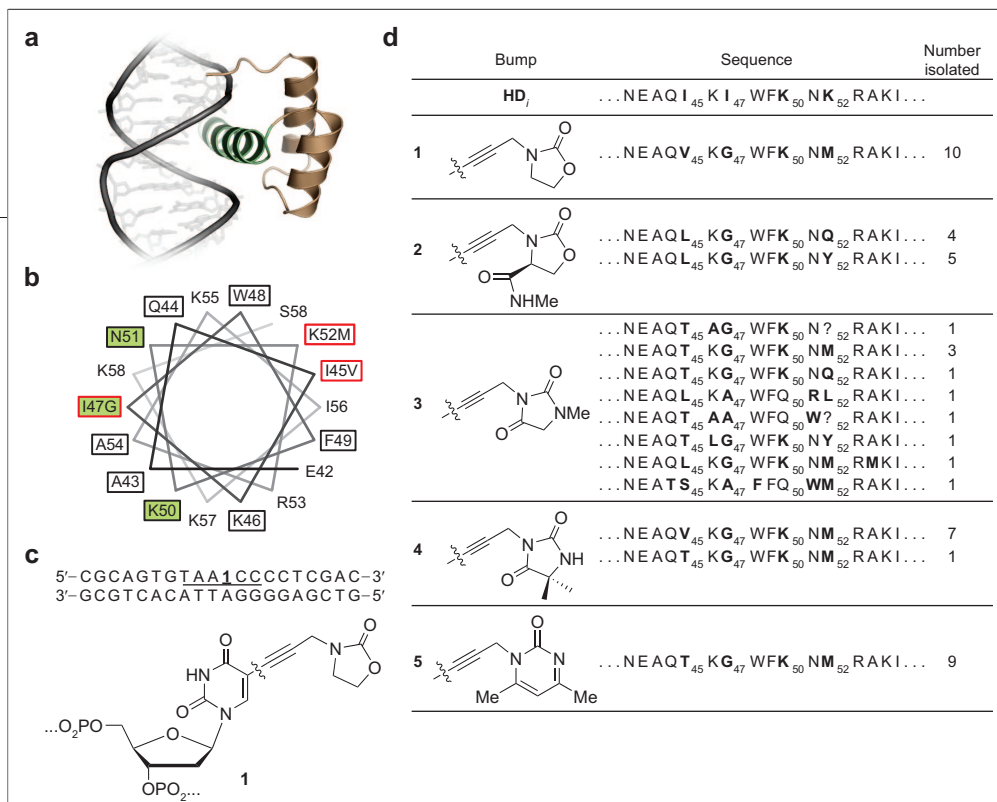


Figure 1. Sequences and schematic representation of the HD mutants recovered from phage selections. a) HD_i is shown bound to TAATCC with its recognition helix (α 3) shown in green. **b)** Helical wheel representation of the HD recognition helix with the three canonical base-specific contacts (Ile47, Lys50, and Asn51) shown in green, residues that were varied in the library presented in boxes, and mutations that comprise HD_q, indicated in the sequence and shown in red. **c)** Example DNA sequence used in the phage selections with oxazolidinone-modified nucleoside 1 incorporated into the HD_i binding site. **d)** Clones recovered from selections for binding to the modified DNA strands TAA1CC (1), TAA2CC (2), TAA3CC (3), TAA4CC (4), and TAA5CC (5). In two cases (indicated with a question mark) the sequencing led to ambiguous results at position 52.

position 47 with the C5 methyl group of thymidine 4, two invariant hydrogen bonds from Asn51 to adenosine 3, and contacts between either a glutamine or a lysine at position 50 that specify the last two base pairs in the recognition sequence (TAATTA or TAATCC, respectively) (19). Although these residues are nearly invariant in naturally occurring HDs, we recently reported the re-engineering of this interface using a phage display selection with a synthetic DNA oligomer bearing an unnatural nucleotide (20). This selected homeodomain (HD_q) binds with specificity for DNA bearing an oxazolidinone appendage projecting into the major groove of the DNA on an alkynyl linker.

Here we report the design and synthesis of other novel nucleosides and the mutations recovered from phage display selections against DNA strands bearing these nucleotides. We also report extensive characterization of one mutant, HD_q, including biochemical dissection of the roles for each

mutation in binding studies, and we analyze the stability of these mutants using CD spectroscopy. Finally, we report the high-resolution crystal structure of HD_q bound to modified and unmodified DNA. Together, these results provide insight into the engineering of highly conserved interfaces.

Design and Synthesis of Derivatized Nucleosides. Our work re-engineering HD–DNA interactions has focused on the well-characterized Q50K mutant of the engrailed homeodomain (HD_i). The Q50K variant is naturally occurring in some HDs (e.g., Bicoid) and alters the binding specificity from the palindromic consensus sequence of the wild-type HD, TAATTA, to the nonpalindromic sequence TAATCC, therefore simplifying biochemical analysis. To engineer HD–DNA interactions, we have focused on regions of the HD_i–DNA interface (21) where the HD makes highly conserved contacts to the DNA (Figure 1, panel a). The hydrophobic contact between Ile47 and the C5 methyl group of thymidine 4 in the recog-

niton motif (TAATTA) was attractive because a large-to-small mutation at Ile47 might create sufficient space to accommodate prosthetic groups appended to the C5 position of T4. Furthermore, the Ile47 contact contributes substantially to DNA binding, but unlike Asn51 (another conserved residue that contacts the DNA), Ile47 is not completely essential ($K_{d, I47A}/K_{d, WT} \sim 20$; $K_{d, N51A}/K_{d, WT} > 1000$ (22)).

With this region of the protein–DNA interface in mind, we have designed synthetic appendages to the DNA that introduce diverse chemical functionality and steric demands into the interface yet are unlikely to compromise the basic structure of

the B-form DNA or clash with the backbone of the helix–turn–helix motif, which is fundamental to the HD fold (20). The use of allyl or propargyl linkers appended to the C5 of thymidine bases allows the incorporation of diverse chemical functionality and the impact of these modifications has been systematically explored elsewhere (23). In these cases, the extended π -system can contribute to the base–base stacking within the DNA helix, thereby stabilizing the desired B-form DNA structure.

Because of the limited space between the major groove of the DNA and the protein backbone of the recognition helix, we chose to focus on flat five- and six-membered heterocycles connected to the base through a propargyl linker at C5 of thymidine (1–5). These alkynyl heterocycles were synthesized as shown in Supplementary Scheme 1. To install these appendages onto the nucleoside base, Sonogashira conditions were used to couple these terminal alkynes to 5'-dimethoxytrityl (DMT)-protected

2'-deoxy-5-iodouridine (see Supplementary Scheme 2) as we have reported previously (20). Installing the DMT-protecting group prior to the Sonogashira coupling proved advantageous both because this route allowed the use of a common starting material and also because this protection scheme allowed the reaction to be performed in THF. Using THF as solvent accelerated the reaction dramatically relative to the rate in DMF and simplified the workup of the desired nucleosides. These modified nucleosides were then incorporated into DNA oligomers (Figure 1, panel c) using solid-phase DNA synthesis as previously described (20).

Phage-Display Selections. Models and biochemical data (20) have demonstrated that a steric clash between Ile47 and alkynyl substituents prevents HD_i from binding to the modified DNA. Therefore, we have sought mutant HDs that bind specifically to the modified DNA strands, presumably including a large-to-small mutation at Ile47. To select for such mutants, we employed a phage selection with a HD displayed on the major coat protein (P8) of M13 phage. The phage display system has been used previously to select for mutants of the engrailed HD that bind to unmodified (24, 25) and oxazolidinone **1** modified DNA (20). The selections using unmodified DNA have validated this approach; the amino acids enriched in the selections largely recapitulated those found in naturally occurring HDs.

To generate the library of mutant HDs, we used a variant of Kunkel mutagenesis to introduce mutations focused around Ile47 (Figure 1, panel b, boxed residues), and to increase the proportion of functional HDs, the library was biased toward approximately four mutations per clone using split-and-pool DNA synthesis to construct the degenerate oligonucleotide used for mutagenesis (20). With this approach, we were able to obtain high coverage of the library using only modest numbers of unique transformants (1.8×10^7 Amp^r colonies).

Using this library and previously established selection conditions, we enriched for DNA-binding phage within 3–4 rounds of selection. To identify mutations enriched by these selections, we sequenced several clones from each enriched pool (Figure 1, panel d). The sequenced clones demonstrated common mutations derived from independent library members, and for TAA1CC a single clone dominated the selection.

As we expected, the sequences largely contained Lys50, consistent with the known role for this residue specifying the last two base pairs of the DNA binding site used in the selection (*i.e.*, TAA \times CC). Although it was encouraging that most of the clones contained a small residue at position 47, this bias was programmed into the original library (50% alanine; 50% glycine). Nonetheless, the bias toward Gly47 (see Figure 1, panel d) was expected on the basis of previ-

ous data demonstrating that I47A has low binding affinity and specificity. From the selected clones, we found a preponderance of mutations at two other positions: Ile45 and Lys52. These mutations are on the back side of the recognition helix from Ile47. Closer examination of the sequences from the selected clones demonstrates that these clones are clustered into two groups: those with consensus mutations (Lys50, Gly47, I45V/T/L, and K52M) and clones that do not have these mutations but are also missing residues identified as particularly important for HD function (*e.g.*, Trp48 and Asn51), suggesting that this second population is composed of residual clones that do not bind the desired DNA.

Having identified consensus mutations in selection against various modified DNA strands, we chose to focus on one mutant in particular, HD_ϕ (I45V, I47G, Q50K, and K52M), because it completely dominated the pool enriched using oxazolidinone **1** and furthermore was the only clone to give a strongly positive result in a phage enzyme-linked immunosorbent assay (20). Replacement of Ile45 by valine is not entirely surprising given that isoleucine and valine are equally likely to occur in natural HDs at position 45 and this substitution is quite conservative. The replacement of Lys52 with methionine, however, is quite surprising; out of 129 human HDs, 103 have arginine at position 52, 10 have lysine, 7 have alanine, but none have methionine (26).

Biochemical Analysis of a Selected Mutant.

When given the choice in a competition electrophoretic mobility shift assay (EMSA), HD_ϕ binds 5-fold more tightly to TAA1CC₂₀ than unmodified DNA, whereas HD_i binds specifically to unmodified DNA (Supplementary

TABLE 1. Affinity for DNA^a and thermal stability of engrailed HD mutants

Mutant	K _d (nM)				T _m (°C)	ΔH (kcal/mol)
	TAATCC ₂₀	TAA1CC ₂₀	TAATGC ₂₀	TAATCG ₂₀		
HD _i	2.1 ± 0.3	17.8 ± 3.5	8.8 ± 0.2	6.0 ± 1.0	52.6 ± 0.7	-30 ± 2
HD _ϕ	5.8 ± 0.6	1.5 ± 0.1	4.5 ± 0.6	18.0 ± 1.5	53.9 ± 0.5	-36 ± 2
HD _i I47G	4.2 ± 0.6	2.3 ± 0.3			47.3 ± 0.5	-29 ± 1
HD _i I47G, I45V	2.3 ± 0.1	3.1 ± 0.3			59.4 ± 0.4	-36 ± 2
HD _i I47G, K52M	4.6 ± 1.0	2.3 ± 0.3			43.1 ± 0.5	-26 ± 1

^a HD binding to 5'-CGCAGTGTAAXXCCTCGAC and its complement was measured by EMSA.

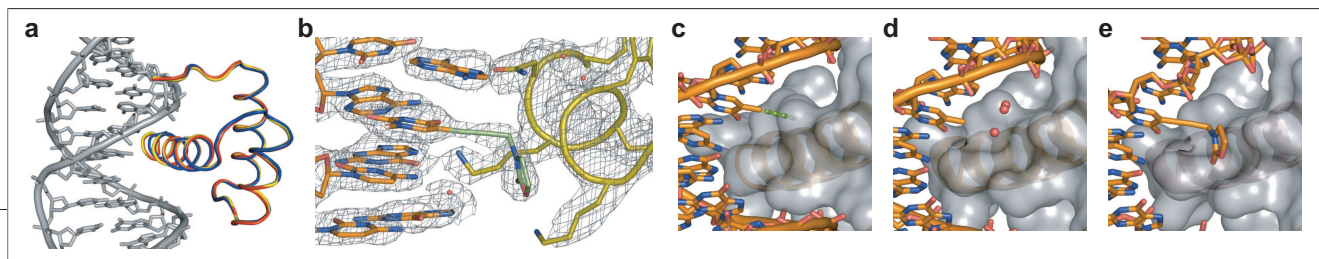


Figure 2. Structure of HD_Φ bound to TAATCC and TAA1CC. **a)** Ribbon overlay of HD–DNA complexes (HD_Φ–TAATCC (15) in blue, HD_Φ–TAA1CC in red, and HD_Φ–TAATCC in yellow). **b)** Electron density in the vicinity of the alkynyl nucleoside. $2F_o - F_c$ electron density from a simulated-annealing composite omit map is contoured at 1σ . **c)** HD_Φ–TAATCC. The hydrophobic contact between Ile47 and the C5 methyl of thymidine in the recognition sequence TAATCC is indicated by green dashes. **d)** HD_Φ–TAATCC. Three waters shown in red occupy the cavity created by the I47G mutation in HD_Φ. For clarity, only these three waters are displayed. **e)** HD_Φ–TAA1CC. The oxazolidinone substituent packs against the cavity created by I47G and displaces the three waters seen in panel d.

Figure 1, panel a). The preference of HD_Φ for TAA1CC₂₀ was seen previously using HD_Φ fused to maltose binding protein (20). Untagged HD_Φ also prefers binding to TAA1CC₂₀ over TAA2CC₂₀, another oxazolidinone-bearing strand (Supplementary Figure 1, panel b).

Next we investigated the role of the two mutations in the hydrophobic core of the HD by mutating these residues back to the wild-type sequence, either individually (HD_i I47G, I45V and HD_i I47G, K52M) or together (HD_i I47G). In EMSAs using either modified or unmodified DNA strands, we were surprised to find that HD_i I47G functioned with specificity and affinity comparable to that of HD_Φ (Table 1 and Supplementary Figure 2). Although the conditions presented also suggest that the selectivity of HDs lacking the hydrophobic mutations is somewhat depressed, the importance of this result is mitigated by modest variability of the selectivities observed under different conditions. Nonetheless, under all conditions tested, HD_Φ is the tightest and most selective HD, demonstrating that HD_Φ binds with the desired specificity for TAA1CC₂₀.

Structural Analysis of the Selected Mutant in Complex with DNA. To understand the structural basis for this re-engineered HD–DNA interaction, we solved X-ray crystal structures of HD_Φ in complex with both modified and unmodified DNA to 2.2 and 1.9 Å, respectively (Supplementary Figure 3 and Supplementary Table 1). Analysis of the structures reveals that, despite the perturbation of the mutations in the HD and the modification to the DNA, HD_Φ still adopts a canonical HD structure as judged by the nearly identical ribbon representations of C α positions for the HD_Φ–TAATCC, HD_Φ–TAATCC, and HD_Φ–TAA1CC structures (Figure 2, panel a). The oxazolidinone modi-

fication is clearly visible in the electron density (Figure 2, panel b) and projects into the cavity created by the I47G mutation (Figure 2, panel e). The average B factor of the oxazolidinone modification (34 Å²) is similar to those of the base to which it is connected (33 Å²) and also of side chains Lys50 (35 Å²) and Asn51 (37 Å²), suggesting that the oxazolidinone is well ordered within the context of the re-engineered interface. In the complex of HD_Φ with unmodified DNA, the cavity created by the I47G mutation (Figure 2, compare panels c and d) is occupied by three ordered water molecules (Figure 2, panel d). In both structures of HD_Φ, with the exception of the waters immediately surrounding the modification and cavity, the waters at the protein–DNA interface are similar to those found in the unengineered HD_Φ–TAATCC interface.

In the structure of HD_i bound to TAATCC, residue Lys50 contacts the last two bases through two alternate conformations allowing it to bind specifically to both base pairs (15). However, in the electron density for HD_Φ bound to either TAA1CC (Supplementary Figure 4) or TAATCC, K50 only occupies one of these conformations, with no apparent structural basis for specifying the final base TAATCC. To test if HD_Φ has decreased specificity for the last C=G base pair (TAATCC) compared to the previous C=G base pair (TAATCC), HD_Φ and HD_i binding to TAATCC and TAATGC were assayed by EMSA (Supplementary Figure 5, panels a and b). We found that HD_Φ retains specificity for the final base pair despite the single orientation of Lys50 observed in the crystal structure. Furthermore, we found that HD_Φ binding to its preferred DNA target is competed off approximately the same concentration of salmon sperm DNA as was

found for HD_i (Supplementary Figure 5, panel c).

We next turned our attention to the two hydrophobic mutations I45V and K52M found in HD_Φ. Examination of the region of HD_Φ around Val45 revealed only minor changes relative to HD_i, such as changes in electron density best explained by Ser35 adopting a second conformation. It is possible that this second conformation requires the slightly reduced steric volume of the I45V mutation, but the importance of this second conformation is unclear. Examination of the environment surrounding the K52M mutation reveals that this mutation could relieve electrostatic repulsion caused by three lysine residues (Lys17, Lys52, and Lys55) in close proximity to one another. The engrailed HD is a member of a small subset of HDs that have basic amino acids at both position 17 and position 52. Most HDs possess a salt bridge between Glu17 and Arg52 in the HD consensus sequence (27). The high density of positive charge caused by the presence of Lys17, Lys52, and Lys55 is destabilizing and can be relieved by K52A and K52E mutations, which stabilize the engrailed HD as previously demonstrated (28). This result suggests that K52M may impact protein stability and led us to wonder more generally about the effects of the HD_Φ mutations on HD stability.

Analysis of the Stability of the Selected Mutant. To examine the stability of the HDs used in this study, we monitored their thermal denaturation by CD spectroscopy (Figure 3). Starting with HD_i, introducing the I47G mutation is destabilizing to the protein ($\Delta T_m = -5.3$ °C). In general, the replacement of an amino acid with glycine is destabilizing because the unfolded state of the protein is entropically stabilized by extra

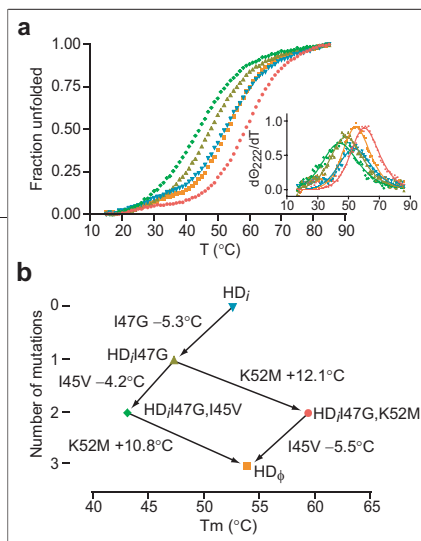


Figure 3. Analysis of the effect of HD_ϕ mutations on HD stability. a) CD spectroscopy was used to monitor the thermal denaturation of HDs. The derivative of the CD curve was used to determine the melting temperature (inset) (31). b) Colors and symbols for each mutant along with the effects of mutations on the thermal stability of the mutant HDs.

conformations available to glycine (29), especially in an α helix (30). As predicted from the relief of repulsive charge–charge interactions, the K52M mutation is stabilizing, not only recovering the stability lost by I47G but further stabilizing HD_ϕ V45I to a net 6.8 °C above HD_ϕ. The dramatic stabilizing effect of the K52M mutation is consistent with the notion that selection favored HDs with stabilizing mutations, but more surprising, the other mutation, I45V, destabilizes the HD by 4–6 °C, depending on the sequence context, bringing the net thermal stability of HD_ϕ close to that of HD_ϕ (Figure 3, panel b, and Table 1). The phage selected mutations returned the stability of the HD, leading to similar stabilities of HD_ϕ with HD_ϕ.

Despite the high conservation of HD–DNA contacts, particularly Gln/Lys50, Asn51, and Ile/Val47, using phage display and nucleoside chemistry, we were able to re-engineer the HD–DNA interface, completely disrupting one of the conserved contacts (Ile47-T4), and replace it with an elaborated nucleoside that packs against the cavity created by an I47G mutation. Although these modifications represent a dramatic perturbation, the high-resolution structure of this interface revealed that the selected mutant functions without affecting the HD fold. Furthermore, we found that the selected mutations tuned the stability of the selected HD to be similar

to that of the starting HD. Analysis of the HD_ϕ–TAA1CC interaction reported here demonstrates that even highly conserved interfaces, such as the HD–DNA interface, contain sufficient adaptability to allow the installation of novel function, in this case specific binding to modified DNA. This latent adaptability is of importance to protein engineers who wish to make tractable the enormity of the proteome by exploiting its conserved motifs and domains, including the HD.

METHODS

Synthesis. Detailed experimental procedures and characterization of the alkynes and nucleosides can be found in Supporting Information.

Phage Library Construction and Selection. The phage library used for these selections and the conditions for the selections have been published elsewhere (20).

Expression and Purification of Mutant HDs. Mutant HDs were expressed as MBP fusions, and the MBP affinity tag was cleaved using Factor Xa. A detailed description of the expression, purification, and characterization can be found in Supporting Information.

Electrophoretic Mobility Shift Assays. Binding of the HDs to DNA was determined essentially as described previously (20, 24), except by using HDs cleaved from the maltose binding protein tag (see above) and conditions described in Supporting Information.

CD. Thermal stability of engrailed HD mutants was measured by CD essentially as described previously (19, 22). Detailed conditions are described in Supporting Information. Melting temperature and enthalpy of denaturation were determined by fitting the derivative of the denaturation curves to the van't Hoff difference equation (31). Fitting the difference data obviates the need to fit baselines to the data, thus reducing the number of free parameters available to the fit. Thermal denaturation curves were numerically differentiated, smoothed over a 3 °C window, and fit to the van't Hoff difference equation using Levenberg–Marquardt least squares minimization (Figure 3, inset of panel a) using scripts written in Matlab (32).

Crystallization, Data Collection, and Structure Refinement. Crystals were grown in hanging drop essentially as described previously (21) except that a higher concentration of poly(ethylene glycol)-400 was used in the well solution. The HD_ϕ–TAATCC structure was solved by molecular replacement using HD_ϕ bound to TAATCC (15) as the initial model. The HD_ϕ–TAATCC structure was then used as the initial model for HD_ϕ bound to TAA1CC. Details of the crystallization, data collection, and structural refinement can be found in Supporting Information.

Accession Codes. Structure factors and final coordinates have been deposited in the RCSB PDB with ID codes 2HOT and 2HOS for HD_ϕ bound to TAA1CC and TAATCC, respectively.

Acknowledgment. We thank L. Rice and C. Wadling for training and assistance with crystallography and structure refinement, the ALS 8.3.1 beamline, J. Chung and I. Gomez Pinto in the James Lab University of California, San Francisco (UCSF) for assistance with large-scale purification of the DNA strands, the Frankel lab (UCSF) for use of their DNA synthesizer, G. Weiss and K. Sato for assistance setting up the phage selections, C. Pabo for a critical discussion of this work, and R. Grant for providing unpublished structure factors for the Q50A HD mutant. M.D.S. was supported by fellowships from the National Science Foundation and the ACS Organic Division. M.E.F. was supported in part by National Institutes of Health (NIH) training grant number GM08284. Mass spectrometry studies were carried out at the UCSF Mass Spectrometry Facility supported by NIH grant number NCCR RR01614. Financial support was provided by the Volkswagen foundation.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Bishop, A. C., Shah, K., Liu, Y., Witucki, L., Kung, C., and Shokat, K. M. (1998) Design of allele-specific inhibitors to probe protein kinase signaling, *Curr. Biol.* 8, 257–266.
- Bishop, A. C., Buzko, O., and Shokat, K. M. (2001) Magic bullets for protein kinases, *Trends Cell Biol.* 11, 167–172.
- Bishop, A. C., Ubersax, J. A., Petsch, D. T., Matheos, D. P., Gray, N. S., Blethrow, J., Shimizu, E., Tsien, J. Z., Schultz, P. G., Rose, M. D., Wood, J. L., Morgan, D. O., and Shokat, K. M. (2000) A chemical switch for inhibitor-sensitive alleles of any protein kinase, *Nature* 407, 395–401.
- Hwang, Y. W., and Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein, *J. Biol. Chem.* 262, 13081–13085.
- Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., Zhang, C., and Shokat, K. M. (2000) Unnatural ligands for engineered proteins: new tools for chemical genetics, *Annu. Rev. Biophys. Biomol. Struct.* 29, 577–606.
- Weijland, A., Parlato, G., and Parmeggiani, A. (1994) Elongation factor Tu D138N, a mutant with modified substrate specificity, as a tool to study energy consumption in protein biosynthesis, *Biochemistry* 33, 10711–10717.
- Weijland, A., and Parmeggiani, A. (1993) Toward a model for the interaction between elongation factor Tu and the ribosome, *Science* 259, 1311–1314.
- Atwell, S., Ultsch, M., De Vos, A. M., and Wells, J. A. (1997) Structural plasticity in a remodeled protein–protein interface, *Science* 278, 1125–1128.
- Greisman, H. A., and Pabo, C. O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites, *Science* 275, 657–661.

10. Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc finger proteins, *Annu. Rev. Biochem.* **70**, 313–340.
11. Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F., 3rd. (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2758–2763.
12. Kortemme, T., and Baker, D. (2004) Computational design of protein-protein interactions, *Curr. Opin. Chem. Biol.* **8**, 91–97.
13. Allen, J. J., Lazerwith, S. E., and Shokat, K. M. (2005) Bio-orthogonal affinity purification of direct kinase substrates, *J. Am. Chem. Soc.* **127**, 5288–5289.
14. Gehring, W. J., Affolter, M., and Burglin, T. (1994) Homeodomain proteins, *Annu. Rev. Biochem.* **63**, 487–526.
15. Tucker-Kellogg, L., Rould, M. A., Chambers, K. A., Ades, S. E., Sauer, R. T., and Pabo, C. O. (1997) Engrailed (Gln50-Lys) homeodomain-DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity, *Structure* **5**, 1047–1054.
16. Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L., and Pabo, C. O. (1994) Structural studies of the engrailed homeodomain, *Protein Sci.* **3**, 1779–1787.
17. Fraenkel, E., Rould, M. A., Chambers, K. A., and Pabo, C. O. (1998) Engrailed homeodomain-DNA complex at 2.2 Å resolution: a detailed view of the interface and comparison with other engrailed structures, *J. Mol. Biol.* **284**, 351–361.
18. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Komberg, T. B., and Pabo, C. O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions, *Cell* **63**, 579–590.
19. Ades, S. E., and Sauer, R. T. (1994) Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50, *Biochemistry* **33**, 9187–9194.
20. Simon, M. D., and Shokat, K. M. (2004) Adaptability at a protein-DNA interface: re-engineering the engrailed homeodomain to recognize an unnatural nucleotide, *J. Am. Chem. Soc.* **126**, 8078–8079.
21. Grant, R. A., Rould, M. A., Klemm, J. D., and Pabo, C. O. (2000) Exploring the role of glutamine 50 in the homeodomain-DNA interface: crystal structure of engrailed (Gln50-Ala) complex at 2.0 Å, *Biochemistry* **39**, 8187–8192.
22. Ades, S. E., and Sauer, R. T. (1995) Specificity of minor-groove and major-groove interactions in a homeodomain-DNA complex, *Biochemistry* **34**, 14601–14608.
23. He, J., and Seela, F. (2002) Propynyl groups in duplex DNA: stability of base pairs incorporating 7-substituted 8-aza-7-deazapurines or 5-substituted pyrimidines, *Nucleic Acids Res.* **30**, 5485–5496.
24. Simon, M. D., Sato, K., Weiss, G. A., and Shokat, K. M. (2004) A phage display selection of engrailed homeodomain mutants and the importance of residue Q50, *Nucleic Acids Res.* **32**, 3623–3631.
25. Sato, K., Simon, M. D., Levin, A. M., Shokat, K. M., and Weiss, G. A. (2004) Dissecting the Engrailed homeodomain-DNA interaction by phage-displayed shotgun scanning, *Chem. Biol.* **11**, 1017–1023.
26. Banerjee-Basu, S., and Baxevanis, A. D. (2001) Molecular evolution of the homeodomain family of transcription factors, *Nucleic Acids Res.* **29**, 3258–3269.
27. Clarke, N. D. (1995) Covariation of residues in the homeodomain sequence family, *Protein Sci.* **4**, 2269–2278.
28. Stollar, E. J., Mayor, U., Lovell, S. C., Federici, L., Freund, S. M., Fersht, A. R., and Luisi, B. F. (2003) Crystal structures of engrailed homeodomain mutants: implications for stability and dynamics, *J. Biol. Chem.* **278**, 43699–43708.
29. Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6663–6667.
30. Serrano, L., Neira, J. L., Sancho, J., and Fersht, A. R. (1992) Effect of alanine versus glycine in alpha-helices on protein stability, *Nature* **356**, 453–455.
31. John, D. M., and Weeks, K. M. (2000) van't Hoff enthalpies without baselines, *Protein Sci.* **9**, 1416–1419.
32. (2004) *Statistics Toolbox User's Guide*; The Mathworks: Natick, MA.

High-Throughput Screening for Functional Adenosine to Inosine RNA Editing Systems

Subhash Pokharel and Peter A. Beal*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112-0850

Deamination of adenosine (A) in RNA is an example of base-modification RNA editing (1). This transformation generates inosine (I) at the corresponding nucleotide position. Because I is decoded as guanosine (G) during translation, the reaction can lead to codon changes in messenger RNA (mRNA) and the introduction of amino acids into a gene product not encoded in the gene (2, 3). Translation of the different coding strands created by this process leads to protein structural diversity. Indeed, A to I RNA editing appears to be necessary to create the structural diversity required for properly functioning central nervous systems in metazoa (4, 5). Two multidomain human proteins have been shown to carry out A deamination in mRNA and have been given the name A deaminases that act on RNA (ADARs) (6). A related family of A deaminases that act on transfer RNA are referred to as ADATs (7). Each ADAR enzyme is made up of identifiable RNA-binding domains containing double-stranded RNA-binding motifs (dsRBMs) and a deaminase domain. Editing site selectivity arises in part by selective binding to double-helical RNA structures found in substrate transcripts mediated by the dsRBMs (8, 9). However, published results suggest the deaminase domain also plays an important role in controlling site selectivity (10, 11). How this takes place is not well understood at the structural–biochemical level.

Genetic strategies are effective for rapidly defining structure–activity relationships in enzyme reactions and for the discovery of

mutant enzymes with new properties. Application of this approach to the study of ADARs requires coupling the generation of libraries of mutant ADARs or mutant substrates with simple screens for editing activity. ADARs are not naturally found in bacteria or yeast, so simple genetic manipulation of these organisms to study A to I RNA editing is not possible. However, active ADARs can be isolated from yeast overexpression systems (12, 13). Therefore, we developed a screening strategy in the yeast *Saccharomyces cerevisiae* that is useful for the rapid identification of active ADAR mutants and new editing substrates capable of supporting the ADAR reaction. Our screen is based on the known ability of ADAR2 to deaminate within a stop codon, converting the sequence to a tryptophan codon (14). Lazinski *et al.* (14) had previously shown that editing of a stop codon found in the hepatitis D virus (HDV) antigenomic RNA could be used as a reporter of RNA editing efficiency by measuring the ratio of long form vs short form of the HDV antigen expressed. This observation suggested to us that a simple ADAR2 substrate upstream of sequence encoding a reporter enzyme could be translated directly and the conversion of a stop codon to a tryptophan codon could control reporter enzyme expression (Figure 1). Thus, we inserted RNA secondary structure known to support ADAR2 editing (human glutamate receptor-B (GluR-B) R/G site hairpin stem) into an mRNA transcript and made the necessary sequence changes to maintain an open reading frame and create a stop codon

ABSTRACT Deamination of adenosines within messenger RNAs catalyzed by adenosine deaminases that act on RNA (ADAR) enzymes generates inosines at the corresponding nucleotide positions. Because inosine is decoded as guanosine, this reaction can lead to codon changes and the introduction of amino acids into a gene product not encoded in the gene. Translation of the different coding strands created by this process leads to protein structural diversity in the parent organism and is necessary for nervous system function in metazoa. The basis for selective editing of adenosines within certain codons is not well understood at the structural/biochemical level. Here we describe a high-throughput screen for ADAR/substrate combinations capable of RNA editing that can be carried out in the yeast *Saccharomyces cerevisiae* growing on agar plates. Results from the screening of libraries of human ADAR2 mutants and libraries of RNA substrates shed light on structure–activity relationships in the ADAR-catalyzed adenosine to inosine RNA editing reaction.

*Corresponding author,
beal@chem.utah.edu.

Received for review September 1, 2006
and accepted November 15, 2006.

Published online December 8, 2006

10.1021/cb6003838 CCC: \$33.50

© 2006 by American Chemical Society

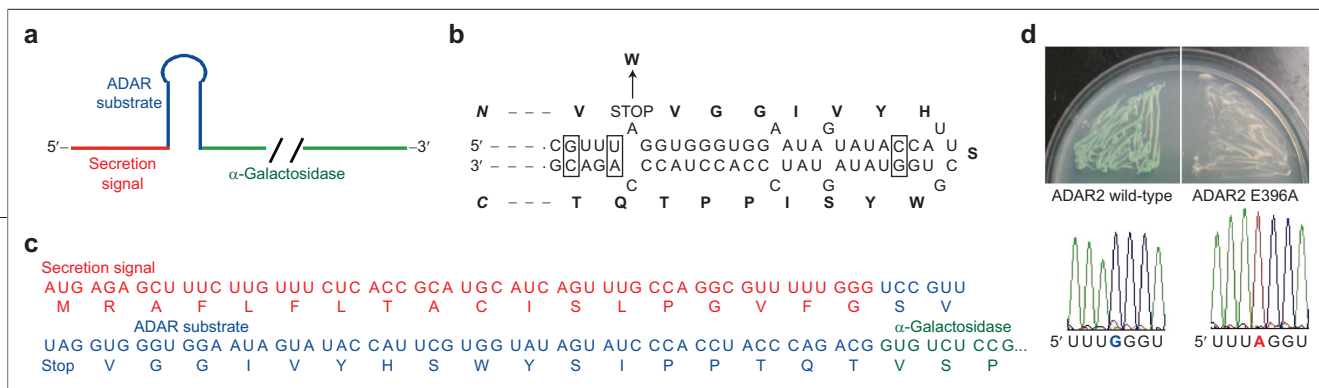


Figure 1. α -Galactosidase reporter for evaluating A to I RNA editing activity in *S. cerevisiae*. **a)** Schematic of reporter mRNA (17, 18). **b)** Sequence of ADAR substrate. The loop and boxed nucleotides have been altered from the original GluR-B pre-mRNA sequence. **c)** Full sequence of reporter mRNA near editing site with the encoded protein sequence shown. **d)** Secreted α -galactosidase activity (top) and reporter mRNA editing (bottom) in yeast expressing wild-type ADAR2 or the inactive mutant E396A.

at the editing site (15). We also reduced the length of the duplex and the size of the loop relative to that naturally found in the GluR-B pre-mRNA. Downstream of this ADAR substrate structure is the sequence encoding α -galactosidase, which is a secreted enzyme readily assayed directly on agar plates containing the colorimetric substrate 5-bromo-4-chloro-3-indolyl- α -galactopyranoside (X- α -Gal) (16). When wild-type human ADAR2 is expressed along with this reporter substrate, yeast grown on X- α -Gal plates appears green, indicating expression and secretion of α -galactosidase (Figure 1) (16). This correlates with editing of the in-frame stop codon as evidenced by sequencing of a reverse transcriptase (RT)-polymerase chain reaction (PCR) product generated from total RNA isolated from this yeast using primers specific for the reporter transcript (Figure 1). No color and no editing are apparent when a known inactive mutant of ADAR2 (E396A) is coexpressed with the reporter mRNA (Figure 1).

To test this system as a method for screening for editing activity, we chose to generate ADAR2 mutants that varied in the identity of two active site residues for which no structure–activity data were previously available. Thr375 is believed to be proximal to the 2'-hydroxyl of the edited nucleotide based on docking AMP into the structure of the deaminase domain (Figure 2) (19). However, this residue is not conserved in the ADAR family (e.g., asparagine in ADAR1) (Figure 2, panel a). In addition, we chose to vary the identity of the amino acid residue at position 376. This residue is a lysine in ADAR2 and a positively charged residue in the known ADARs and ADATs (7).

Its position in the protein suggests a possible interaction with the 3'-phosphodiester of the edited nucleotide (Figure 2, panel b).

Plasmid libraries were created by saturation mutagenesis at the codons for these residues. Yeast expressing the α -galactosidase reporter was transformed with the resulting libraries. Colored colonies growing on X- α -Gal plates were then identified (Figure 2, panel c). Plasmid DNA encoding ADAR2 was isolated from yeast colonies appearing either green or white on X- α -Gal plates. Sequencing revealed the identity of codons leading to active or inactive editing enzyme. In a screen where only the Thr375 codon was varied, plasmid DNA from 21 green and 13 white colonies was isolated and sequenced (Table 1). Common among the clones from green colonies were codons for small hydrophobic residues (8 for alanine, 6 for valine, and 2 for glycine), with wild-type threonine observed once. White colonies typically had position 375 codons encoding large residues (tyrosine, arginine, and lysine), with stop codons and a codon for proline also observed. The green color phenotype was confirmed and shown to correlate with editing of the reporter mRNA by sequencing of the RT-PCR product from yeast expressing either the T375A or T375Y mutants (Supplementary Figure 1). Interestingly, the T375A mutant is less active in this assay than wild-type ADAR2, since a lower level of editing is observed in the sequenced RT-PCR product. The purified ADAR2 T375A mutant also displayed a lower deamination rate than wild-type in *in vitro* deamination assays with a model substrate (data not shown). Thus, although T375A mutant scored as a hit in the screen, this mutation

does reduce ADAR2 deaminase activity. It is clear that under the current screening conditions, it is only possible to distinguish ADAR mutants with large differences in editing activity. Additional studies will be necessary to identify conditions that allow for ranking among ADAR mutants with similar activities.

We also carried out a screen of a library of mutants that varied the identity of residues at both positions 375 and 376 of ADAR2. From the 14 green colonies isolated, we found 10 clones encoding the T375C, K376H mutant (Table 1). In addition, although a positively charged residue is conserved at position 376 in the known ADARs and ADATs, several clones were identified with hydrophobic amino acids encoded for this residue, including the K376I mutant (7). The phenotype and reporter RNA editing were confirmed for the T375C, K376H, and K376I mutants (Supplementary Figure 1).

Important structure–activity relationships for the ADAR2 reaction can be gleaned from analysis of the amino acids accommodated at positions 375 and 376 in the active mutants identified in this study. Small hydrophobic amino acids at position 375 lead to activity and large residues are excluded, consistent with the proposed proximity of Thr375 with the editing site nucleoside (Figure 2, panel b) (19). Interestingly, when the amino acids at positions 375 and 376 were varied simultaneously, only threonine or cysteine was selected at position 375 in active mutants. These residues are capable of hydrogen bonding to the 2'-hydroxyl of the edited A. Threonine was selected in this screen encoded by three different codons, highlighting the

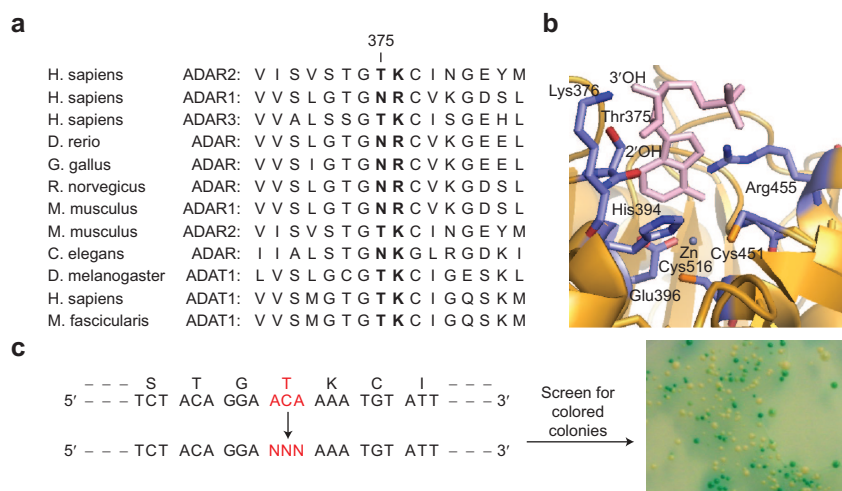


Figure 2. The ADAR active site. a) Sequence alignment for ADARs and ADATs. b) ADAR2 active site residues with docked AMP (19). c) Saturation mutagenesis was carried out at the T375 codon, and the resulting plasmids were screened for the generation of differently colored colonies.

importance of this amino acid at the 375 position of ADAR2.

The selection system described here allows one to vary the sequence of either the editing enzyme or the editing substrate, since both are encoded on plasmids. To evaluate the effect of varying the editing substrate, we prepared two different plasmid libraries with different parts of the substrate sequence randomized (Table 2). In one library, six nucleotides were varied in the sequence complementary to the editing site (ECS library). In a different library, the four-nucleotide loop and two nucleotides of the loop-closing base pair were randomized.

Interestingly, <1% of the colonies observed during selections with the ECS library appeared green, indicating a strong dependence on the local sequence environment of the edited A (Supplementary Figure 2). When plasmids encoding the ADAR substrate from six green colonies observed in this screen were sequenced, the original sequence (CCCAGA) appeared four times along with two other sequences (CGGUAG and CGGUGA) (Table 2). The latter two related sequences are predicted to place the

edited A in a six-nucleotide, purine-rich symmetrical loop. Sequencing of an RT-PCR product from yeast expressing the CGGUGA mutant confirms efficient editing within this structure (~50% conversion to I) (Supplementary Figure 3). Activity with such a structure would not have been predicted given our current understanding of the role of the base pairing partner for the reactive A and the preference for As within duplex secondary structures (11, 15). This particular loop sequence must form a unique structure that properly positions the editing site for reaction with ADAR2. Importantly, a small change to this sequence has a dramatic effect on the screening results. Among the RNAs identified in the white colonies from this screen was the sequence AUCUGA, differing from an active sequence by only three nucleotides. However, folding of this RNA sequence *in silico* predicts an entirely different secondary structure (20). In this case, the edited A is predicted to exist in an A-U pair surrounded by four other stable base pairs. Furthermore, this structure is predicted to have single nucleotide bulges in locations where they do not exist naturally

on the R/G stem. RT-PCR product sequencing confirms little editing within this RNA (Supplementary Figure 3). It is possible that this duplex is too stable to allow for efficient flipping of the reactive adenosine into the deaminase active site (21). However, the placement of bulged nucleotides may also be an important negative determinant for activity with this substrate (22).

A recent report suggested that the five-nucleotide loop of the native R/G hairpin stem found in the GluR-B pre-mRNA is important for controlling editing efficiency at the R/G site (9). For the substrate studied here, most of the sequences in the library support editing since ~80% of transformants with the loop sequence library gave green colonies (Supplementary Figure 2). However, when 10 active clones were sequenced from this screen, the sequence ACUCAC was observed eight times (Table 2 and Supplementary Figure 3). The CUCA sequence contained within is remarkably similar to sequences found naturally in GluR-B R/G hairpin loops (human, CUAA; rat, CUCA; mouse, CUCA; tilapia D, CUAC) (23). Allain *et al.* (9) have presented NMR data indicating that dsRBM I of ADAR2 binds the loop of the native R/G hairpin stem. Our results provide additional evidence for an important role played by this particular loop sequence in facilitating editing at the R/G site. Furthermore, selection of a near wild-type GluR-B R/G editing site loop sequence previously implicated in direct binding to ADAR2 underscores the utility of screening RNA libraries for preferred substrates with this method.

Sequencing of plasmids from white colonies in this screen revealed RNA sequences that could never lead to reporter enzyme expression (*e.g.*, stop codons and frame shift mutations) (Table 2). Indeed, low reporter enzyme expression in these cases was the result of factors other than low editing efficiency. One sequence found in three different white colonies (AAGCAU) maintained an open reading frame and is

TABLE 1. Results of screening of ADAR2 T375X and T375X, K376X libraries

Library	Colony color	Codon(s) (Amino acid)	No. of clones	
T375X	Green (active)	GCA (Ala)	4	
		GCG (Ala)	2	
		GCC (Ala)	1	
		GCT (Ala)	1	
		GTA (Val)	6	
		GGA (Gly)	1	
		GGG (Gly)	1	
		ACA (Thr)	1	
		CAG (Gln)	1	
	ATG (Met)	1		
	ATT (Ile)	1		
	CAT (His)	1		
	White (inactive)	TAA (Stop)	1	
		TAG (Stop)	1	
		TAC (Tyr)	2	
		AGG (Arg)	2	
		CGG (Arg)	2	
		AAG (Lys)	3	
		CCG (Pro)	1	
CCT (Pro)		1		
T375X, K376X		Green (active)	TGT, CAT (Cys, His)	10
			ACA, ATA (Thr, Ile)	1
	ACG, CTT (Thr, Leu)		1	
	TGC, TTG (Cys, Leu)		1	
	ACT, AAC (Thr, Asn)		1	

predicted to form a tetraloop with a loop-closing A-U pair. Sequencing of the RT-PCR product from this sample indicates that RNA editing is supported (data not shown), yet poor expression of α -galactosidase was observed. Therefore, one must be cautious in interpreting reduced α -galactosidase expression solely in terms of editing efficiency without independent confirmation, as other factors can play a role. Indeed, the predicted loop sequence (AGCA) and an adjacent long duplex are features of good substrates for Rnt1p, a *S. cerevisiae* RNase III (24).

In summary, fusing a RNA-editing substrate containing a stop codon in frame with sequence encoding the reporter enzyme α -galactosidase allows for high-throughput screening in *S. cerevisiae* of different combinations of mutant ADARs and RNA substrates to identify combinations that support editing. Simple modifications to this screen should allow for the discovery of regulators of the RNA editing reaction or evolution of new ADARs with altered site selectivity.

METHODS

Construction of Editing Substrate/ α -Galactosidase Reporter Plasmid pR/G α Gal. The α -galactosidase expression plasmid pMEL α was obtained as a gift from EUROSARF, Germany. The α -galactosidase coding sequence present on pMEL α was amplified by PCR and ligated into pYES3/CT (Invitrogen) using standard protocols to generate pR/G α Gal. Sequences for all primers used in this study can be found in Supplementary Table 1.

Co-expression of α -Galactosidase Reporter and Human ADAR2 in Yeast. INVSc1 *S. cerevisiae* strain (Invitrogen) was sequentially transformed with an ADAR2 expression plasmid (YEPTOP2PGAL1 (13) or YEPTOP2PGAL1-E396A) with *URA3* selection followed by transformation with an editing reporter plasmid (pR/G α Gal) with *TRP1* selection using a lithium acetate protocol. Yeast colonies were harvested from these plates and transferred to agar plates containing complete minimal (CM) media-uracil, -tryptophan 2% raffinose, 3% galactose, and 60 μ g mL⁻¹ X- α -Gal. The plates were incubated at 30 °C for 3–5 d until color was apparent.

Sequencing of Reporter mRNA Isolated from Yeast Expressing ADAR2 Mutants. Single yeast colonies transformed with pR/G α Gal and an ADAR2 expression plasmid were harvested and placed into 5 mL of CM-tryptophan, -uracil, 2% raffinose media for 48 h at 30 °C with shaking. Protein expression was induced by adding 3% galactose to the media, and growth was continued for another 48 h at 30 °C with shaking. mRNAs were isolated using the RiboPure-Yeast kit (Ambion) following the manufacturer's protocol. Isolated mRNA

was reverse transcribed and amplified using the Access RT-PCR kit (Promega). The extent of RNA editing was assessed by sequencing the resulting RT-PCR product.

Generation of Plasmid Libraries. The QuikChange XL site-directed mutagenesis kit (Stratagene) was used to generate plasmid libraries following the manufacturer's protocol. Resulting PCR products were used to transform XL 10 gold *Escherichia coli* cells using ampicillin selection. Bacteria colonies were harvested, combined, and grown in Luria-Bertani (LB) ampicillin media overnight at 37 °C with shaking. Plasmid DNA was isolated using a QIAprep Spin Miniprep kit (Qiagen). Sequence randomization at the desired site was confirmed by sequencing.

Screening for Functional Editing Combinations. INVSc1 cells were transformed with either pR/G α Gal or YEPTOP2PGAL1 (13) plasmids using a lithium acetate protocol. A second lithium acetate transformation was performed to introduce plasmid libraries. After the second transformation in screens for ADAR mutants, yeast were plated directly onto agar plates containing CM-tryptophan, -uracil, 2% raffinose, 3% galactose and X- α -Gal. The plates were incubated at 30 °C for 5–10 d until color was apparent. After the second transformation in screens for mutant substrates, cells were plated onto agar plates containing CM-tryptophan, -uracil, 2% glucose. These plates were incubated at 30 °C for 3 days, after which they were replica plated onto agar plates containing CM-tryptophan, -uracil, 2% raffinose, 3% galactose, and X- α -Gal. Within 72 h, green colored colonies were visualized. Yeast colonies appearing either green or white were harvested and lysed. Plasmid DNA was isolated from the lysate by ethanol precipitation and used to transform XL 10 gold *E. coli* cells (Stratagene) with selection on LB-ampicillin plates. Plasmid DNA was isolated from the resulting bacteria colonies and sequenced.

Acknowledgment: P.B. acknowledges support from the National Institutes of Health (GM61115).

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

1. Maydanovich, O., and Beal, P. A. (2006) Breaking the central dogma by RNA editing, *Chem. Rev.* 106, 3397–3411.
2. Burns, C. M., Chu, H., Rueter, S. M., Hutchinson, L. K., Canton, H., Sanders-Bush, E., and Emeson, R. B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing, *Nature* 387, 303–308.
3. Higuchi, M., Single, F. N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P. H. (1993) RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency, *Cell* 75, 1361–1370.
4. Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P. H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2, *Nature* 406, 78–81.

TABLE 2. Results of screening of ADAR2 substrate libraries

Library	Colony color	RNA sequence (5' → 3')	No. of clones	Encoded sequence (N → C)
ECS	Green	CCCAGA	5	TQT
		CGGUGA	2	TVT
		CGGUAG	1	TVA
	White	UAUAAU	2	I-Stop
		AUCUGA	1	NLT
		CCGCGC	1	TAP
Loop	Green	ACUCAC	8	HSR
		CAUCAA	1	PSR
		GCCGUA	1	RRR
	White	AAGCAU	3	QAW
		CCCAAUC	2	Frame shift
		CGUAAA	1	P-Stop

5. Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) A-to-I pre-mRNA editing in *Drosophila* is primarily involved in adult nervous system function and integrity, *Cell* 102, 437–449.
6. Bass, B. L., Nishikura, K., Keller, W., Seeburg, P. H., Emeson, R. B., O'Connell, M. A., Samuel, C. E., and Herbert, A. (1997) A standardized nomenclature for adenosine deaminases that act on RNA, *RNA* 3, 947–949.
7. Gerber, A., Grosjean, H., Melcher, T., and Keller, W. (1998) Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2, *EMBO J.* 17, 4780–4789.
8. Stephens, O. M., Haudenschild, B. L., and Beal, P. A. (2004) The binding selectivity of ADAR2's dsRBMs contributes to RNA-editing selectivity, *Chem. Biol.* 11, 1–20.
9. Stefl, R., Xu, M., Skrisovska, L., Emeson, R. B., and Allain, F. H.-T. (2006) Structure and specific RNA-binding of ADAR2 double-stranded RNA-binding motifs, *Structure* 14, 345–355.
10. Lehmann, K. A., and Bass, B. L. (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities, *Biochemistry* 39, 12875–12884.
11. Kallman, A. M., Sahlin, M., and Ohman, M. (2003) ADAR2 A-I editing: site selectivity and editing efficiency are separate events, *Nucleic Acids Res.* 31, 4874–4881.
12. O'Connell, M. A., Gerber, A., and Keegan, L. P. (1998) Purification of native and recombinant double-stranded RNA-specific adenosine deaminases, *Methods* 15, 51–62.
13. Macbeth, M. R., Lingam, A. T., and Bass, B. L. (2004) Evidence for auto-inhibition by the N-terminus of hADAR2 and activation by dsRNA binding, *RNA* 10, 1563–1571.
14. Sato, S., Wong, S. K., and Lazinski, D. W. (2001) Hepatitis delta virus minimal substrates competent for editing by ADAR1 and ADAR2, *J. Virol.* 75, 8547–8555.
15. Stephens, O. M., Yi-Brunozzi, H.-Y., and Beal, P. A. (2000) Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site, *Biochemistry* 39, 12243–12251.
16. Rupp, S. (2002) LacZ assays in yeast, *Methods Enzymol.* 350, 112–131.
17. Hofmann, K. J., and Schultz, L. D. (1991) Mutations of the alpha-galactosidase signal peptide which greatly enhance secretion of heterologous proteins by yeast, *Gene* 101, 105–111.
18. Aho, S., Arffman, A., Pummi, T., and Uitto, J. (1997) A novel reporter gene MEL1 for the yeast two-hybrid system, *Anal. Biochem.* 253, 270–272.

19. Macbeth, M. R., Schubert, H. L., VanDemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing, *Science* 309, 1534–1539.
20. Mathews, D. H., Sabrina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol.* 288, 911–940.
21. Yi-Brunozzi, H.-Y., Stephens, O. M., and Beal, P. A. (2001) Conformational changes that occur during an RNA-editing adenosine deamination reaction, *J. Biol. Chem.* 276, 37827–37833.
22. Lehmann, K. A., and Bass, B. L. (1999) The importance of internal loops within RNA substrates of ADAR1, *J. Mol. Biol.* 291, 1–13.
23. Aruscavage, P. J., and Bass, B. L. (2000) A phylogenetic analysis reveals an unusual sequence conservation with introns involved in RNA editing, *RNA* 6, 257–269.
24. Wu, H., Yang, P. K., Butcher, S. E., Kang, S., Chanfreau, G., and Feigon, J. (2001) A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III, *EMBO J.* 20, 7240–7249.

Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent Anticancer Agents from *Nostoc* Cyanobionts

Nathan A. Magarvey^{†,¶}, Zachary Q. Beck^{‡,¶}, Trimurtulu Golakoti[§], Yousong Ding[‡], Udo Huber[§], Thomas K. Hemscheidt[§], Dafna Abelson[†], Richard E. Moore[§], and David H. Sherman^{†,*,*}

[†]Department of Microbiology and BioTechnology Institute, University of Minnesota, Minneapolis–St. Paul, Minnesota 55108, [‡]Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, and Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, [§]Department of Chemistry, University of Hawaii, Manoa, Hawaii 96822. [¶]These authors contributed equally to this work.

ABSTRACT The lichen cyanobacterial symbiont *Nostoc* sp. ATCC 53789 and its close relative *Nostoc* sp. GSV 224 are prolific producers of natural products, generating >25 derivatives of the cryptophycin class of secondary metabolites. Cryptophycin 1, the prototypic member of the class, is a potent tubulin-depolymerizing agent, and several semisynthetic derivatives are being developed as anticancer therapeutics. Here we provide a detailed characterization of the cryptophycin metabolic pathway by stable-isotope labeling experiments and through cloning, sequencing, and annotating the cryptophycin biosynthetic gene cluster. A comparative secondary metabolomic analysis based on polyketide (PK)/non-ribosomal peptide gene clusters from the phylogenetically related, non-cryptophycin producing cycad symbiont, *Nostoc punctiforme* ATCC 29133, was used to identify the cryptophycin biosynthetic genes that encompass ~40 kb within the lichen symbiont *Nostoc* sp. ATCC 53789 genome. The pathway encodes a collinear set of enzymes, including three modular PK synthases, two non-ribosomal peptide synthetase modules, and an integrated adenylation/ketoreductase didomain for elaboration of the leucic acid subunit. In addition, genes encoding key tailoring steps, including a FAD-dependent halogenase and CYP450 epoxidase, were identified. The inherent flexibility of the cryptophycin biosynthetic enzymes was harnessed to generate a suite of new analogues by altering the pool of PK starter units and selected amino acid extender groups. Characterization of the cryptophycin CYP450 enabled development of the first stereospecific synthesis of cryptophycin 2, through a tandem chemoenzymatic synthesis from the natural seco-cryptophycin 4 chain elongation intermediate.

Cryptophycins are the largest class of peptolides isolated from cyanobacteria (blue-green algae) to date (1). The lichen cyanobacterial symbiont *Nostoc* sp. ATCC 53789 and its close genetic relative *Nostoc* sp. GSV 224 produce these promising anticancer agents (1–5). Cryptophycin 1 (1), the major representative of >25 naturally occurring analogues, consists of four units, including a phenyl-octenoic acid (unit A) and L-leucic acid (unit D) and two amino acids, 3-chloro-*O*-methyl-*D*-tyrosine (unit B) and methyl β -alanine (unit C), linked in a cyclic ABCD sequence (Table 1). All of the other naturally occurring cryptophycins are analogues that differ structurally from 1 by one or two units in the molecule (Table 1). The *Nostoc*-derived cryptophycins exhibit extensive variation, indicating the flexibility and versatility of the biosynthetic pathway (1–5) (Table 1). Natural cryptophycin variants of unit A differ in their oxygenation state (*e.g.*, alkene vs hydroxyl groups and epoxide vs styrene) and double bond configurations (*trans* vs *cis*). Unit B contains phenylalanine variants, unit C includes methyl β -alanine or β -alanine, and unit D involves α -hydroxy acid subunit diversity (Table 1). Another notable cryptophycin variation, which does not commonly stem from other polyketide (PK) and non-ribosomal peptide biosynthetic pathways, is macrocyclic ring size (16- vs 14-membered peptolide rings) (see Table 1) (1, 3, 4).

1 (Table 1), found as the most abundant product from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224, is one of the most potent tubulin destabilizing agents ever discovered (6). It arrests tumor cells at the G₁-M phase, inducing a block in cellular proliferation, and causes hyper-phosphorylation of Bcl-2, triggering the apoptotic cascade (7). Cryptophycins are also attractive as chemotherapeutic agents because they are not substrates for

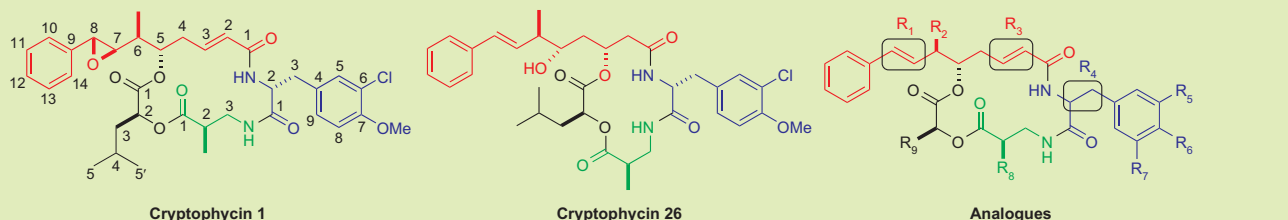
*Corresponding author,
davidhs@umich.edu.

Received for review October 16, 2006
and accepted November 29, 2006.

Published online December 15, 2006

10.1021/cb6004307 CCC: \$33.50

© 2006 by American Chemical Society

TABLE 1. Structural variation of cryptophycins isolated from *Nostoc* strains^a


		Natural cryptophycin analogues																												
		1	2	3	4	16	17	18	19	21	23	24	26	28	29	30	31	38	40	43	45	46	49	50	54	175	176	326	327	
R ₁	β-epoxide	•	•			•		•			•	•		•			•							•	•		•	•	•	
	<i>trans</i> -Styrene			•	•		•		•	•			•		•	•				•	•	•	•			•				
	α-Epoxide																			•										
R ₂	CH ₃	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
	H														•					•										
R ₃	<i>Trans</i> double bond	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
	<i>Cis</i> double bond																												•	
	OH														•		•													
R ₄	L																								•					
	D		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
R ₅	Cl	•		•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	H		•		•							•									•									
R ₆	OCH ₃	•	•	•	•			•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
	OH						•	•													•	•					•			
R ₇	H	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	Cl													•												•		•		
R ₈	CH ₃	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	H														•												•		•	
R ₉	Isobutyl	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	<i>n</i> -Propyl																							•	•			•		
	Isopropyl														•															
	<i>sec</i> -Butyl															•														

^aBiosynthetic subunits are distinguished by color: unit A (red); unit B (blue); unit C (green); unit D (black). These data were assembled according to previous studies (1, 3, 4).

p-glycoprotein pumps and are active against multidrug-resistant tumor cell lines (6). These properties led to the advancement of cryptophycin 52 (LY355703), a synthetic analogue, to phase II clinical trials. In an initial study, dose-limiting toxicities of that analogue restricted its further advancement (8). However, in a subsequent phase II clinical trial performed on patients with

platinum-resistant ovarian cancer, the study concluded that the considerable rate of disease stabilization suggests that LY355703 might warrant further investigation (9). Moreover, another generation of 1 analogues has been synthesized that shield the reactive epoxide as a chlorohydrin or glycinate ester, resulting in improved solubility (10). Preclinical studies with the most promis-

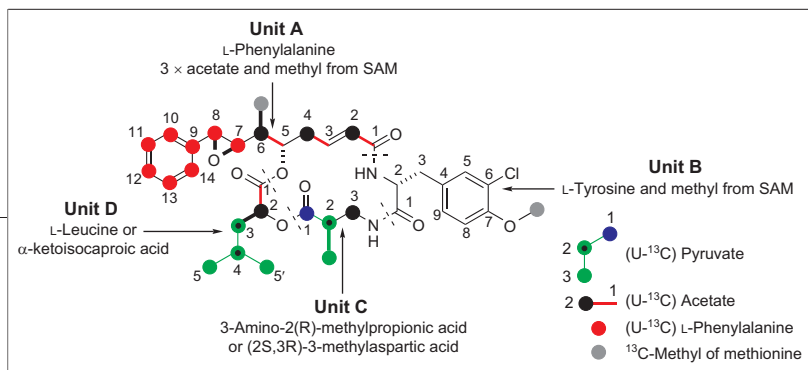


Figure 1. Summary of precursor incorporation experiments and labeling patterns to determine the origin of units A, B, C, and D of cryptophycin 1.

ing of these compounds show a dramatic increase in activity against a variety of tumors (10).

Because of the lack of large scale fermentation methods for isolation of cryptophycins, total synthesis was required to obtain adequate supplies for clinical evaluation (11). Several effective synthetic approaches have been developed for cryptophycin 3, a natural desoxy analogue of **1**. However, the most significant challenge has been the late-stage stereospecific installation of the epoxide moiety across the unit A C7–C8 double bond (Table 1), which is necessary due to the labile nature of this functionality. Initial efforts resulted in no better than a 3:1 mixture of **1** and the α -epoxide diastereomer that was difficult to separate (12). A more effective strategy to control epoxide stereochemistry was accomplished using Evans asymmetric aldol chemistry (13). This problem was addressed in an alternative manner for the synthesis of cryptophycin 52, which employed Sharpless methodology to generate a C7–C8 *syn* diol at the seco-cryptophycin stage and subsequent conversion to the β -epoxide (14, 15). More recently, we reported a convergent, chemoenzymatic synthesis of desoxy-deschloro-cryptophycin 1 by assembling a set of seco-cryptophycin chain elongation intermediates as the *N*-acetyl cysteamine (SNAC) ester and converting them to the corresponding cyclic depsipeptide using a cryptophycin (Crp) thioesterase (TE) mediated approach (16). Completion of the final step remained dependent on successful development of the predicted CYP450 derived from the native Crp biosynthetic system to install the β -epoxide functional group.

Here we provide a complete analysis of the biosynthetic origin and assembly for the A–D subunits comprising the Crp pathway through use of stable isotope precursor-labeling studies and the cloning, sequencing, and biochemical characterization of the Crp metabolic system. A comparative secondary metabolomic analysis was performed to localize and identify the 40 kb *crp* mixed polyketide synthase (PKS)/non-ribosomal peptide synthetase (NRPS) gene cluster. Annotation provided a detailed view of the collinear metabolic system

and revealed a number of unique enzymatic steps involved in subunit assembly, peptolide ring elaboration, and post-PKS/NRPS tailoring reactions. Precursor-directed biosynthesis using unnatural starter units afforded a suite of novel cryptophycins, providing direct evidence for the

flexibility of the biosynthetic enzymes along the assembly line. Finally, we demonstrate that CrpE is the cryptophycin CYP450 that provides an efficient *in vitro* method to generate the β -epoxide with complete regio- and stereochemical control. This enabled a novel chemoenzymatic synthesis of cryptophycin 2 using a tandem macrocyclization/epoxidation reaction sequence with CrpD TE and CrpE.

RESULTS AND DISCUSSION

Identification, Cloning, and Sequencing of the Cryptophycin Biosynthetic Gene Cluster: Comparative Cyanobiont Secondary Metabolome Analysis. The subunit structures (units A–D) comprising **1** (Figure 1) suggest an assembly from carboxylic acid and amino acid precursors by a mixed PKS/NRPS system. Degenerate polymerase chain reaction (PCR) primer sets are widely used to amplify and clone DNA fragments encoding segments of NRPS adenylation (A) (amino acid selecting) domains and ketosynthase (KS) (condensing enzyme) domains of type I PKSs (17, 18). Therefore, our initial strategy was to amplify A and KS domain DNA fragments and use the mixture of amplicons to detect cosmids containing PKS and NRPS genes from a *Nostoc* sp. ATCC 53789 genomic library. In the case of *Nostoc* sp. ATCC 53789, an unusually large number of nonoverlapping clones containing both NRPS and PKS genes were isolated reflecting a highly complex secondary metabolome. Four nonoverlapping PKS and NRPS gene-containing cosmids were partially sequenced to reveal distinct metabolic systems, but none were consistent with a Crp pathway (data not shown). Interestingly, three of the four cosmid DNA sequences revealed high similarity to secondary metabolic gene clusters found within the sequenced genome of the cycad symbiont, *Nostoc punctiforme* ATCC 29133 (data not shown) (19).

Nostopeptolides are the only previously described secondary metabolites produced by *N. punctiforme* ATCC 29133 (20), but there are no reports of its ability to produce cryptophycins. A Crp-sensitive bioassay using *Cryptococcus neoformans* (2) as an indicator strain confirmed the lack of Crp production by *N. punctiforme*

(data not shown). Therefore, we considered that a bioinformatics approach using A and KS domain sequences from *N. punctiforme* compared to A and KS domain sequences from *Nostoc* sp. ATCC 53789 would identify PKS and NRPS pathways unique to this lichen symbiont. Of the *Nostoc* sp. ATCC 53789 A domain sequences cloned and analyzed, only six were absent from *N. punctiforme*. One of the six DNA fragments contained within cosmid pNAM123 had a predicted aromatic amino acid specificity code (Asp235, Ala236, Ser239, Thr278, Ile299, Ala301, Gly322, Ile330) (21) and was selected as the candidate Crp pathway A domain for unit B. Significantly, PCR primers designed from the pNAM123 DNA insert (see Methods) generated an amplicon from *Nostoc* sp. GSV 224 genomic DNA whose sequence was 98% identical to the pNAM123 insert (data not shown). As further evidence that the Crp biosynthetic gene clusters from *Nostoc* sp. GSV 224 and *Nostoc* sp. ATCC 53789 are virtually identical, a fosmid was cloned from *Nostoc* sp. GSV 224 that partially contained *crpD* and encompassed all of the downstream genes associated with Crp production. The overlapping DNA sequences from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 were >99.9% identical and contained *crpE*, *crpF*, *crpG*, and *crpH* in the same order (Figure 2). Interestingly, the DNA sequences from both species were essentially identical up to the transposase region, after which they diverged completely, suggesting that the *crp* gene clusters were integrated at different genomic loci within the two species. These data also provide evidence that *crpH* represents the terminus of the Crp biosynthetic gene cluster. Probing a *Nostoc* sp. ATCC 53789 cosmid library provided pDHS500, whose insert was sequenced and found to contain NRPS and PKS genes consistent with the predicted architecture for the 3'-half of the Crp pathway. Additional library probing using PCR-based screening provided pDHS501 that overlapped with pDHS500 and contained the complete upstream PKS gene portion of the putative *crp* cluster.

The *crp* gene cluster (40,304 bp) is flanked by transposases and inverted repeats (interestingly, the curacin A and jamaicamide cyanobacterial biosynthetic gene clusters (18, 22) are also flanked by transposases) that appear to represent the limits of the metabolic system (Figure 2). The first open reading frame (ORF) within *crp* is a type I PKS gene (8823 bp) that is followed by a second modular PKS gene (10,407 bp) and then two modular NRPS genes (5829 and 10,029 bp) designated

crpA-D. A series of ORFs (*crpE-H*) downstream of the Crp PKS and NRPS genes are predicted to encode enzymes that catalyze functional group modifications (e.g., epoxidation (*crpE*) and chlorination (*crpH*)) present in **1** and many of its analogues. The *crp* biosynthetic gene cluster architecture (Figure 2) and a summary of each deduced protein sequence and corresponding functional role are summarized in Supplementary Table 1.

Proposed Assembly of the Cryptophycins. Biosynthesis of Type I PKS-Unit A: δ -Hydroxy-phenyloctenoic Acid.

As a first step toward unraveling the biosynthetic origin of unit A in the cryptophycins, sodium [1,2- $^{13}\text{C}_2$]acetate was provided in precursor incorporation studies to *Nostoc* sp. GSV 224 (Figure 1 and Supplementary Experiment A). In this experiment the precursor was diluted with unlabeled acetate to minimize the formation of interconnected [1,2- $^{13}\text{C}_2$]acetate units. The ^1H -decoupled ^{13}C NMR spectrum of labeled **1** (Figure 1) appeared as a 1:0.64:1.2 cluster of peaks (1.8%) for the signals of six contiguous carbons in unit A, specifically C1 to C6, which was consistent with the incorporation of three intact acetate units. Each triplet was composed of a singlet for the natural abundance ^{13}C flanked by peaks of a doublet for the incorporated ^{13}C . The level of ^{13}C enrichment (integration of doublet/integration of singlet) averaged 1.1%. The coupling constants associated with the doublets rigorously established that the three intact acetate units had been assimilated into C1–C2, C3–C4, and C5–C6. No ^{13}C enrichment was observed in any of the other carbon signals for unit A. Similar results were obtained when sodium [U- $^{13}\text{C}_3$]pyruvate was provided as a precursor to the cyanobacterium.

To determine the fate of the acetate protons in the formation of the C1–C2–C3–C4–C5–C6 segment of unit A, sodium [2- ^{13}C , $^2\text{H}_3$]acetate was provided as a precursor to the bacterial cells (Supplementary Experiment B). The ^2H -decoupled ^{13}C NMR spectrum of the labeled **1** exhibited enhanced signals (1.8%) for C2 and C6 and isotopically shifted ^{13}C signals for C4. The C4 signals were found in a 1:0.6:1.2 triplet at 36.70, 36.35, and 36.02 ppm assigned to undeuterated, monodeuterated, and dideuterated C4, respectively. About 80% of the deuterium on the ^{13}C incorporated into C4 had been retained, whereas all of the deuterium on the ^{13}C incorporated into C2 and C6 had been lost.

To establish the origin of oxygen atoms attached to C1 and C5 in unit A, sodium [1- ^{13}C , $^{18}\text{O}_2$]acetate was provided to a culture of *Nostoc* sp. GSV 224 (Supple-

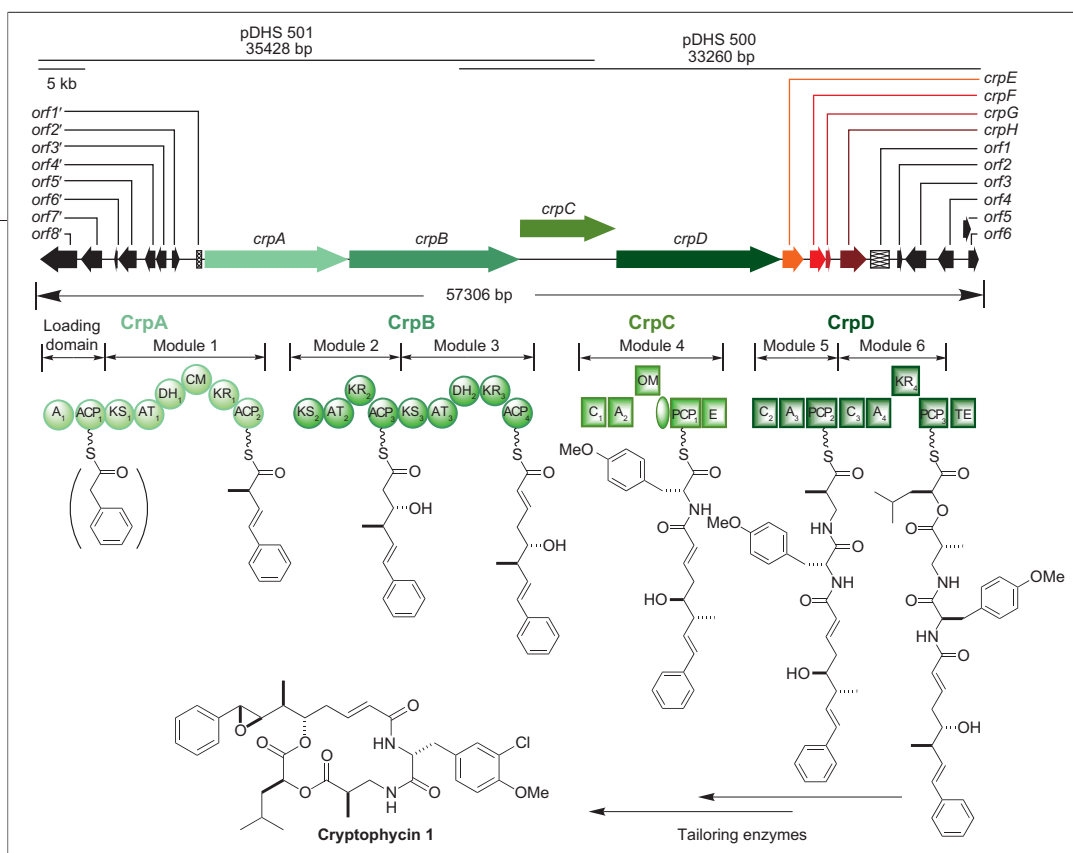


Figure 2. The Crp gene cluster and the deduced PKS/NRPS assembly line. Block arrows indicate ORFs. Genes that are cross-hatched are truncated ORFs or presumed pseudogenes. Genes depicted in green colors indicate that the corresponding gene products are predicted to function as NRPSs or PKSs. Genes shown in orange to red are predicted to encode tailoring enzymes or enzymes involved in precursor biosynthesis. Genes depicted in black are predicted to fall outside of the Crp biosynthetic gene cluster. Domains found within the proposed Crp PKS and NRPS assembly line are shown in circles (PKS) and squares (NRPS) with domain abbreviations (AT, acyltransferase domain; KS, ketosynthase domain; CM, C-methyltransferase domain; DH, dehydratase domain; KR, ketoreductase domain; ACP, acyl carrier protein; A, adenylation domain; C, condensation domain; PCP, peptidyl carrier protein; OM, proposed O-methyltransferase domain; E, epimerase; TE, thioesterase domain) denoting their function.

mentary Experiment C). The ^{13}C NMR spectrum of the labeled **1** showed ^{13}C peaks (1.3% enrichment) for C1 and C5 that were isotopically shifted upfield from the natural abundance lines by 0.03 and 0.04 ppm, respectively. This result revealed that essentially all of the ^{18}O on the ^{13}C incorporated into C1 and C5 from this precursor had been retained.

A precursor incorporation experiment with L-[methyl- ^{13}C]methionine established that the methyl group on C6 of unit A originates from the C₁ pool (Figure 1 and Supplementary Experiment D). The ^{13}C NMR spectrum showed a substantial enhancement (15%) of the carbon signal at δ 13.5.

The origin of the remaining carbons in unit A (i.e., C7, C8, and those in the phenyl group) were shown to be phenylalanine-derived from a precursor incorporation experiment with L-[U- $^{13}\text{C}_9$, ^{15}N]phenylalanine (Supplementary Experiment E). The precursor was provided to *Nostoc* sp. GSV 224, and the bacterial culture was harvested and processed for **1**. The proton-decoupled ^{13}C NMR spectrum of **1** exhibited doublet and doublet resonances (2.3% enrichment from the precursor)

for C7 and C8, respectively, due to ^{13}C enrichment from the precursor, with the natural abundance ^{13}C superimposed at the center of the carbon signal (other carbons appeared as multiplets). The direct ^2H NMR analysis of **1** obtained following supplementation of the cultures with L-[$^2\text{H}_8$]phenylalanine showed poor but positive incorporation of the precursor as indicated by weak deuterium signals for the phenyl group (Supplementary Experiment F). No significant level of deuterium on the benzylic epoxymethine

carbon, however, was detected. On the basis of these precursor incorporation results, phenylacetyl-CoA is a potential starter unit for the biosynthesis of unit A in **1** with chain elongation proceeding by a sequential addition of three malonate units (Figure 1 and Figure 3, panel a). However, feeding experiments with labeled phenylacetic acid and the corresponding N-acetylcysteine thioester failed to show incorporation. Although it might not be possible for these precursors to be transported across the cyanobacterial cell membrane, it is conceivable that another starter unit is involved (see below).

Recently, the origin of the starter unit of the cyanobacterial toxin microcystin (Mcy) was investigated by testing *in vitro* the activation and acylation of a series of candidate aryl and amino acids by McyG loading module (23). Acceptable substrates for the A loading domain were phenylpropanoids with McyG showing the highest selectivity toward *trans*-cinnamic acid. Interestingly, the CrpA A domain amino acid specificity code is identical to the corresponding McyG A domain, and thus, it is possible that *trans*-cinnamic acid functions as the Crp starter unit since it is also derived from phenyl-

a

Methyl Malonyl CoA	V V Q x x x x x M x S L A x x	G H S Q G
CrpA AT ₁	Y T Q V A L F A I E Y A L Y K L	G H S A G
CrpB AT ₁	Y T Q V A L F A I E Y A L Y K L	G H S V G
AT ₂	Y T Q V A L F A I E Y A L Y K L	G H S A G
Malonyl CoA	Y T Q x x x x x E x A L x x x	G H S V G

b

	235	236	239	278	299	301	322	330	331	517
McyG	V	G	V	W	V	A	A	S	G	K
CrpA	V	G	V	W	V	A	A	S	G	K
CrpC	D	A	S	T	I	A	G	I	G	K
CrpD-M1	G	D	A	V	F	S	L	A	D	K
Ebony	V	D	A	V	V	S	F	G	D	K
CrpD-M2	V	A	I	F	L	G	S	S	G	K
CesA-M1	V	G	V	W	V	G	T	S	G	K

c

Type A (consensus)	H			F S S	W
CrpB-KR ₃ (type B)	H A A G I L D D G	---	F S S M A S I L G S P G Q G N Y		
CrpD-KR ₄	H M A G I I Q E T	---	F C S V N G F F G G T N V A A Y		
Type B (consensus)	H	L D D	F S S		P N Y

d

		M1		M2		M3
N-Mt	V L (D E) x G x G x G	N E L S x Y R Y x A V	V E x S x A R Q x G x L D			
CrpC OMet	V V E I G T G K D	G V F E D L D F S K L	Not present			
ApdB-M3Met	V V E I G T G K D	D V L E D L N F N E Y	Not present			

Figure 3. Sequence alignments between key conserved amino acid motifs of previously characterized PKS and NRPS domains, compared to deduced amino sequence of motifs identified from Crp PKS and NRPS domains. **a)** Identification of AT specificity based on consensus sequences motifs from malonyl-CoA and methylmalonyl-CoA AT domains. **b)** Alignment of the proposed active site region of KR domains CrpA-KR₃ and CrpD-KR used previously to classify KR domains as type A and type B (24). Residue numbering is relative to residue 1 of the EryA KR domain (54). **c)** Comparison of conserved motifs from NRPS *N*-methyltransferase domains (N-Mt domains) (27) with the CrpC and anabaenopeptidase (ApdB)-M3 methyltransferase domain motifs. **d)** The deduced codes of the Crp adenylation domains (CrpA-phenylacetate-like starter unit; CrpC, chloro-*L*-tyrosine, CrpD-M1, methyl β -alanine; CrpD-M2, α -ketoisocaproic acid) as well as NRPS codes for the Ebony protein (*Drosophila melanogaster*), β -alanine; McyG (microcystin PKS), proposed phenylacetate specific A domain; CesA-M1 (cereulide NRPS), α -ketoisocaproic acid specific A domain.

alanine (Figure 3, panel b). A remaining mystery is the mechanistic basis for an unprecedented one-carbon truncation within the growing unit A PK chain derived from a phenylpropanoid starter unit.

The precursor incorporation studies described above are consistent with the established three-module domain organization of CrpA (A-ACP-KS-AT-DH-CM-KR-ACP) and CrpB (KS-AT-KR-ACP-KS-AT-DH-KR-ACP) that are collinear with the **1** unit A structure (Figure 2). Several structural variations that deviate from the collinearity rule in the CrpA/B PKS are noteworthy. One is the smaller macrocycle (14-membered ring) found in cryptophycin 26 (Table 1), which is unique compared to all other cryptophycins. Formation of cryptophycin 26 requires that the C3 hydroxyl group serves as the nucleophile for CrpD TE mediated ring closure (16). Bypassing of the CrpB dehydratase (DH) might preserve the C3 hydroxyl group in cryptophycin 26 as well as cryptophycin 30 (Table 1). In another natural analogue (cryptophycin 327) a C-2/C-3 *cis*-double bond occurs instead of the C-2/C-3 *trans*-double bond common to all other cryptophycins. The CrpB KR₂ domain is predicted to be responsible for C3 ketoreduction, and its amino acid sequence is coincident with an *S*-specific KR (or type B) based on bioinformatics analysis (24–26) (Figure 3, panel c), consistent with the C-2/C-3 double bond. Presumably the CrpB KR₂ domain also generates an *R* C-3 stereoisomer at a low level, with the CrpB DH₂ catalyzing dehydration to the C-2/C-3 *cis* olefin.

NRPS: Incorporation of Unit B and Unit C in the Cryptophycins. Unit B: 3-Chloro-*O*-methyl-*D*-tyrosine.

Unit C: Methyl- β -alanine. Unit B of **1** is 3-chloro-*O*-methyl-*D*-tyrosine, but other rare cryptophycins incorporate *L*-amino acids (e.g., *L*-tyrosine), dichlorotyrosines (e.g., 3,5-dichloro-*O*-methyl-*D*-tyrosine), or desmethyl-deschlorotyrosines (e.g., *O*-methyl-*D*-tyrosine, *D*-tyrosine) (Table 1). The origin of unit B in **1** was explored using *L*-[1-¹³C]tyrosine and *D*- and *L*-3-[(4-¹³C₂H₃)-phenyl]-alanine precursor incorporation studies (Supplementary Experiments G and H, respectively). The ¹³C NMR spectrum from **1** produced in the presence of *L*-[1-¹³C]tyrosine showed a 4% enhancement of C-1 of unit B. Feeding experiments with *D*- and *L*-3-[(4-¹³C₂H₃)-phenyl]-alanine resulted in deuterium-enriched **1**, indicating that unit B is derived from *L*-tyrosine, and the Crp NRPS catalyzes the incorporation of *O*-methyl-*D*- and *L*-tyrosine (i.e., *L*-3-[(4-¹³C₂H₃)-phenyl]-alanine) (Figure 1).

CrpC is a monomodular NRPS containing an elongation module with an A domain that has an NRPS amino acid specificity code (Figure 3, panel a) most similar (75% identical, 87% similar) to the ApdB-M4 NRPS A domain (*O*-methyl-*L*-tyrosine, or *N,O*-dimethyl chloro-*L*-tyrosine). The CrpC A domain, bears an *S*-adenosylmethionine (SAM)-dependent methyltransferase between motifs A8 and A9 (27) of the polypeptide (Figure 2). Protein database searching revealed that the most similar peptide sequence (64% identity, 84% similarity) to the CrpC methyltransferase domain (~400 amino acids) is the ApdB-M3 methyltransferase of the

Apd NRPS (28). We propose the CrpC and ApdB-M3 methyltransferases (*O*-methylation of tyrosine) are a new type of NRPS domain that specify *O*-methylation, which is supported by unique amino acid sequence motifs (Figure 3, panel d). Both the ApdB-M3 and CrpC methyltransferase domains lack a core *N*-methyltransferase amino acid sequence motif (M3), and residues within core methyltransferase motifs M1 and M2 are different. Following the presumed OMet domain are A domain motifs (A9 and A10) of unknown function, a PCP domain, and an epimerase (E) domain (27) that is predicted to catalyze the conversion of *L*-tyrosine to *D*-tyrosine. Unit B is structurally modified by mono- or dichlorination of the tyrosine subunit (Table 1). It remains to be determined whether halogenation occurs at the subunit stage to generate a precursor pool of chlorotyrosine or during a later stage of natural product assembly.

CrpH, whose gene (*crpH*) represents the presumed terminus of the *crp* gene cluster is the predicted halogenase responsible for modification of tyrosine at the position ortho to the OMet moiety (Supplementary Table 1). Database comparisons with the deduced CrpH peptide sequence revealed that it resembles a number of previously described non-heme flavin-dependent halogenases, including those from the rebeccamycin (29) and pyoluteorin (30) pathways. We expect that CrpH is responsible for mono- and dichlorination of the aromatic ring of unit B for the various Crp natural products.

Unit C of **1** is methyl- β -alanine, whereas in a small number of analogues (*e.g.*, cryptophycins 18, 24, 29, 176, 326; Table 1) β -alanine is incorporated for the elongation step. In primary metabolism, phosphopantetheine formation requires β -alanine that is in turn derived from aspartate *via* a PanD decarboxylase (31). To determine if methyl- β -alanine is formed from methyl-aspartate (MASP), a MASP synthesis (32) was developed using [^{13}C]iodomethane instead of iodomethane to provide a 1:1 (*SR/SS*) diastereomeric mixture of the compounds (Supplementary Experiment I). This enabled precursor incorporation studies using a ^{13}C -labeled methyl group in cultures of *Nostoc* sp. GSV 224. A 1.4% enrichment of ^{13}C label from [*methyl*- ^{13}C]-(*2S,3R/S*)-3-MASP in the methyl group of unit C in **1** demonstrated that unit C is derived from this biosynthetic subunit (Figure 1). In addition, feeding studies using [^{13}C]pyruvate resulted in intact incorporation of pyruvate into the methyl- β -alanine unit (Figure 1 and Supplementary Experiment J).

The only other natural product that is known to contain methyl- β -alanine is vicienistatin. However, a combination of feeding studies and analysis of the DNA sequence of the entire vicienistatin biosynthetic cluster (33) indicates a unique origin for this subunit. Specifically, the methyl- β -alanine unit is presumably derived from glutamate that is rearranged by a glutamate mutase to form MASP and finally decarboxylated by a pyridoxal-5'-phosphate-dependent decarboxylase (33, 34). Intact incorporation of pyruvate by the Crp biosynthetic pathway is inconsistent with the formation of MASP derived by rearrangement from glutamic acid and suggests that MASP in *Nostoc* may originate from a biosynthetic pathway analogous to branched chain amino acid biosynthesis. The primary sequence of CrpG is similar to various PanD enzymes (Supplementary Table 1). Interestingly, when ^{13}C -labeled *S*-methyl- β -alanine was fed to cultures of *Nostoc* sp. GSV 224, it was incorporated with a percent yield comparable to the incorporation of the *R*-isomer (Supplementary precursor-directed biosynthesis). It appears that the CrpD A2 domain does not discriminate between these diastereomeric forms of the subunit. This result suggests that CrpG displays high selectivity toward *R*-MASP to afford the *2R*-stereoisomer of methyl- β -alanine. The ability of the CrpD-M1 A domain to activate methyl- β -alanine/ β -alanine is predicted based on its sequence similarity to previously characterized (21) β -alanine A domains (Figure 3, panel b).

Mosaic NRPS/PKS Module. Unit D: Variation and Incorporation of Branch Chain α -Hydroxy Acids. On the basis of reasonable biosynthetic principles, it was evident that unit D of **1** is derived from α -hydroxyisocaproic acid. Interestingly, α -hydroxyisovaleric acid, α -hydroxyisobutyric acid, or α -hydroxyvaleric acid is also incorporated based on the structures of several natural Crp analogues (Table 1; cryptophycins 19, 21, 49, 50, 54, 326). We sought to determine if *L*-leucine is the direct precursor for unit D by conducting precursor incorporation experiments with dL -[5- $^2\text{H}_3$]leucine (Supplementary Experiment K) and with L -[1- ^{13}C]leucine (Figure 1, Supplementary Experiment L). In feeding experiments with dL -[5- $^2\text{H}_3$]leucine, **1** was isolated and shown to have a ^2H resonance at 0.827 ppm (1.8% incorporation). Intact and equal incorporation (0.5%) of the methyl side chains derived from leucine into the unit D methyl groups was evident in the ^{13}C NMR spectrum of **1** isolated from these *Nostoc* cultures. In feed-

ing studies with L-[1-¹³C]leucine, the ¹³C NMR spectrum showed incorporation of 1% of the label from C-1 of the precursor into C-1 of unit D. Precursor incorporation studies with α-hydroxyisocaproic acid, however, did not result in labeled cryptophycins (Supplementary Experiment M). Unlike α-hydroxyisocaproic acid, α-ketoisocaproate is a common leucine biosynthesis intermediate (35), formed from leucine *via* transamination. The direct contribution of α-ketoisocaproate to unit D was tested by precursor incorporation studies with α-[1-¹³C]ketoisocaproic acid. **1** extracted from the resulting cultures was subjected to ¹³C NMR analysis that showed enhancement (0.9%) of the ester carbonyl (Figure 1; Supplementary Experiment N). On the basis of the required conversion of leucine to the corresponding α-keto acid, we surmise that CrpF, a putative non-heme oxygenase (Supplementary Table 1), might catalyze oxidative deamination of leucine yielding α-ketoisocaproate. Final KR-mediated reduction (see below) would be required to generate the nascent unit D precursor.

The final module within the Crp PKS/NRPS is the second module of CrpD (CrpD-M6) that bears an NRPS-like elongation module (C₃-A₄-KR₄-PCP₃) and the terminal TE domain, CrpD TE (16) (Figure 2). Following presumed activation and transfer of α-ketoisocaproic acid to the CrpD-M6 PCP₃, the tethered α-keto acid might serve as a substrate for the upstream KR-like domain. In polyketide biosynthesis, KR domains have been demonstrated to reduce β-keto groups with no reported precedent for α-keto reducing KR. The valinomycin (*vlm*) (36) and cereulide (*ces*) (37) NRPSs contain dehydrogenases similar to that found in CrpD. One of the *ces* dehydrogenases was recently demonstrated to catalyze α-keto reduction of a thioester bound 2-keto isocaproyl-PCP substrate (38).

NRPS C domains typically function as amide synthetases, but the domain placement and predicted substrate of the CrpD C₃ domain suggest that it is an ester synthetase. Highly homologous domains from the *vlm* (36) and *ces* (37) NRPSs were also predicted to bear ester synthetase C domains. Recently the fumonisin-A PCP-C didomain was shown to catalyze this type of ester bond formation (39). Finally, the hydroxyl group of unit A and the leucic acid carboxyl group of unit D are linked to form a lactone, catalyzed by CrpD TE (16).

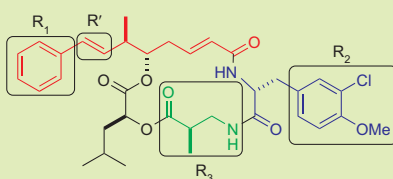
Precursor-Directed Biosynthesis of Novel Cryptophycins.

The unusually large number of natural product analogues generated by the Crp biosynthetic system suggests that it is comprised of flexible enzymes for depsipeptide assembly and tailoring. We sought to assess further the versatility of the Crp assembly line using precursor-directed biosynthesis. A series of amino acids and halogens were used in order to probe the flexibility of enzymes responsible for incorporation of units A, B, and C.

We first targeted the aromatic groups of unit A as this is among the least varied in the natural cryptophycins (Table 2), yet synthetic chemical structural modification to the aromatic group can provide enhanced bioactivity (11). Our demonstration that phenylalanine is converted into a substrate for the CrpA A₁ domain suggested that analogues of this amino acid might provide insight into CrpA initiation domain flexibility. As an initial experiment, we used *p*-methylphenylalanine for precursor incorporation, which led to the new metabolite cryptophycin 111 (Table 2 and Supplementary Information). Further experiments included a diverse set of phenylalanine analogues, including those with larger *p*-alkyl chains (*i.e.*, ethyl and alkyl hydroxyl) and more bulky ring systems (naphthyl and biphenyl), cyano-appended rings, and para-substituted halides (Cl, Br, F, and I) (Table 2). In each case, new cryptophycins were generated with expected starter units, revealing that the Crp loading domain is remarkably flexible. Interestingly, many of these amino acids were also incorporated within unit B (Table 2). Moreover, this work demonstrates that the Crp PKS/NRPS is able to channel and process alternative substrates of varying structural complexity (Table 2). In most cases, the new compounds obtained were modified by the presumed CrpE CYP450, revealing significant tolerance of the epoxidase toward unnatural substrates (Table 2).

The *in vivo* specificity of the CrpH halogenase was also investigated by adding CaBr₂ and CaI₂ to the cyanobacterial culture medium. Surprisingly, both bromine and iodine were incorporated into unit B, indicating an additional opportunity for *in vivo* generation of diverse products (Table 2). This was particularly intriguing in view of recent reports that replacement of chlorine for iodine has not been observed previously in natural product halogenases (40). Interestingly, the epoxidized product of the brominated molecule (Table 2; cryptophycin 104) was extracted from the bacterial culture, but the

TABLE 2. Summary of the precursor-directed biosynthesis to assess substrate tolerance in the Crp pathway^a



Precursors	Cryptophycin products				
	R'		R ₁	R ₂	R ₃
	<i>trans</i> -Styrene	β-Epoxyde			
DL-4-Fluorophenylalanine	110	115	4-Fluorophenyl	*	*
DL-4-Chlorophenylalanine	124	125	4-Chlorophenyl	*	*
	324		4-Chlorophenyl	4-Methoxyphenyl	*
DL-4-Bromophenylalanine	304	305	4-Bromophenyl	*	*
DL-4-Iodophenylalanine	310		4-Iodophenyl	*	*
L-4-Methylphenylalanine	111	117	4-Methylphenyl	*	*
	312		4-Methylphenyl	4-Methoxyphenyl	*
L-4-Ethylphenylalanine	339		4-Ethylphenyl	*	*
L-4-Trifluoromethylphenylalanine	161		4-Trifluoromethylphenyl	*	*
L-4-Cyanophenylalanine	318		4-Cyanophenyl	*	*
DL-3-Fluorophenylalanine	311	303	3-Fluorophenyl	*	*
(2 <i>R</i>)-3-Fluorotyrosine	211	316	*	3-Fluoro-5-chloro-4-methoxy-phenyl	*
(2 <i>R</i>)-3-Fluoro-4-[² H ₃]-methoxytyrosine	210	190	*	3-Fluoro-4-[² H ₃]-methoxyphenyl	*
L-3-Iodotyrosine	320	319	*	3-Iodo-4-methoxyphenyl	*
DL-2-Fluorophenylalanine	181	182	2-Fluorophenyl	*	*
(2 <i>R</i>)-3-(3,4-[² H ₂]-Methylenedioxy-phenyl)alanine	208	209	*	3,4-[² H ₂]-Methylenedioxy-phenyl	*
(2 <i>R</i>)-3-(3-Methyl-4-[² H ₃]-methoxy-phenyl)alanine	189		*	3-Methyl-4-[² H ₃]-methoxy-phenyl	*
L-4-Methylhydroxyphenylalanine	336		*	4-Methylhydroxyphenyl	*
	238		4-Methylhydroxyphenyl	3-Chloro-4-methoxyphenyl	*
L-2-Naphthylalanine	334		*	Naphthyl	*
	172		Naphthyl	*	*
	335		Naphthyl	Naphthyl	*
CaBr ₂	315	104	*	3-Bromo-4-methoxyphenyl	*
		313	*	3-Bromo-4-hydroxyphenyl	β-Alanine
		314	*	3-Bromo-4-hydroxyphenyl	*
CaI ₂	320	319	*	3-Iodo-4-methoxyphenyl	*
β-Alanine	29	21	*	*	β-Alanine
	24		*	4-Methoxyphenyl	β-Alanine
(2 <i>S</i>)-3-Aminoisobutyric acid	342	343	*	*	(2 <i>S</i>)-3-Aminoisobutyric acid
<i>gem</i> -Dimethyl-β-alanine	52		*	*	<i>gem</i> -Dimethyl-β-alanine

^aThe asterisks indicate that the R group is the same as the R group of the illustration.

iodinated form (Table 2; cryptophycin 320) was also isolated as the *trans*-styrene, indicating that unit B might impose structural constraints on CrpE substrate selectivity.

Flexibility of the Crp pathway unit C enzymes was investigated by evaluating precursor incorporation using β-alanine, and (2*S*)-3-aminoisobutyric acid. Both of these amino acids were incorporated into unit C with complete downstream processing to the expected cyclic depsipeptide structures (Table 2; cryptophycins 24, 29,

342, 343). Cryptophycin 52, the synthetic lead molecule that was advanced to clinical trials, is identical to **1** except for a *gem*-dimethyl β-alanine as unit C. On the basis of precursor incorporation studies showing that the CrpD A domain is tolerant of unnatural subunits, we sought to test whether the Crp biosynthetic machinery was capable of forming cryptophycin 52 *in vivo*. Direct supplementation of *gem*-dimethyl β-alanine to *Nostoc* cultures resulted in a product that was shown by NMR to

be cryptophycin 52 (Table 2 and Supplementary Information for NMR characterization of cryptophycins). This is the first report using precursor-directed biosynthesis to produce cryptophycin 52, further advancing the potential of the Crp biosynthetic machinery to access important bioactive Crp analogues. Previous to this work, access to this important analogue was only possible by total synthesis (41).

Chemoenzymatic Synthesis of Cryptophycin 2: Sequential Synthesis from Seco-cryptophycin 4 Using CrpD TE/CrpE P450. Recently, we reported the ability of excised CrpD TE to catalyze *in vitro* macrocyclization of several acyl-peptide SNAC variants of Crp chain-elongation intermediates (16). One of the depsipeptide products obtained from the TE was desoxycryptophycin 2, the presumed substrate for the Crp epoxidase. The desoxycryptophycins have significantly reduced bioactivity, and it is installation of the β -epoxide that represents a significant challenge in Crp total synthesis. The Crp epoxidase was readily identified as the putative product of *crpE* whose deduced amino acid sequence is similar to numerous CYP450s involved in hydroxylation or epoxidation. To explore its specific function, the gene was cloned and overexpressed in *Escherichia coli* to provide a soluble histidine tagged/maltose-binding protein variant (data not shown). Incubation of CrpE with cryptophycin 4 resulted in >75% conversion to a compound with a mass consistent with cryptophycin 2 (Figure 4). Scale-up provided sufficient quantities of the epoxidized material to confirm by ^1H NMR that the compound is a single diastereomer (*i.e.*, the β -epoxide) and identical in all respects to cryptophycin 2 (42). This result confirms that CrpE is the Crp epoxidase that enables efficient, stereospecific installation of a β -epoxide into the cryptophycins. As an additional demonstration of the flexibility of this system, we employed the seco-cryptophycin 4 intermediate in a tandem *in situ* reaction that included CrpD TE, CrpE, and all required cofactors (see Methods). This experimental system provided direct isolation of cryptophycin 2 as a single product without the need to purify the desoxycryptophycin 2 intermediate (Figure 4). Effective *in vitro* reconstitution of these two final biosynthetic steps promises immediate access to new cryptophycins in the search for improved anticancer lead compounds.

In summary, the cyanobacterial symbiont-derived *Nostoc* sp. ATCC 53789 and the highly related *Nostoc* sp. GSV 224 produce the cryptophycin natural products.

These compounds have shown significant potential as anticancer therapeutic agents, and a number of analogues are currently being explored as clinical candidates. In this report, we focused on five allied studies, including (i) cloning, sequencing, and bioinformatics analysis of the Crp gene cluster, (ii) precursor incorporation studies to reveal the specific source of each biosynthetic subunit, (iii) precursor directed biosynthesis to probe the flexibility of Crp metabolic enzymes involved in chain assembly, cyclization, and tailoring, (iv) identification and functional analysis of the CrpE CYP450 that catalyzes β -epoxide formation to yield **1** and related analogues, and (v) tandem *in vitro* assembly of cryptophycin 2 from seco-cryptophycin 4 using CrpD TE and CrpE reconstituted in a single reaction vessel.

Because of the remarkable number of secondary metabolic gene clusters associated with *Nostoc* sp. genomes, a comparative analysis was employed to survey the complete secondary metabolome of *N. punctiforme* and use it as a tool to identify the *crp* gene cluster in both *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224. To our knowledge, this is the first report of natural product gene cluster identification using a comparative metabolomic strategy, made possible by the availability of a full genome sequence in a highly related species. Previous studies have shown the close phylogenetic relationship between lichen cyanobionts (43), but our reported strategy probes further the potential value of investigating cyanobiont phylogeny and secondary metabolism.

A particularly significant characteristic of the Crp biosynthetic system is its ability to generate >25 related natural products that include alterations in subunit structures, including units A, B, C, and D, depsipeptide ring size (14 vs 16), and variations in tailoring reactions, including epoxidation and halogenation. Subunit A structural variants are presumably derived from imprecise PKS processing reactions, including Crp module 3 KR₃ and DH₂ domains. These events result in variant double bond configuration or presence of a second hydroxyl group at C-3 in unit A (Table 1). NRPS enzymes involved in assembly of units B, C, and D show flexibility in amino acid and α -hydroxy acid subunit selectivity, resulting in the incorporation of diverse, yet related, extender groups. Tailoring reactions involving the *crpE*-encoded halogenase can result in one or two chlorine atoms in the unit B aromatic ring of tyrosine. Precursor incorporation studies revealed definitively the source of

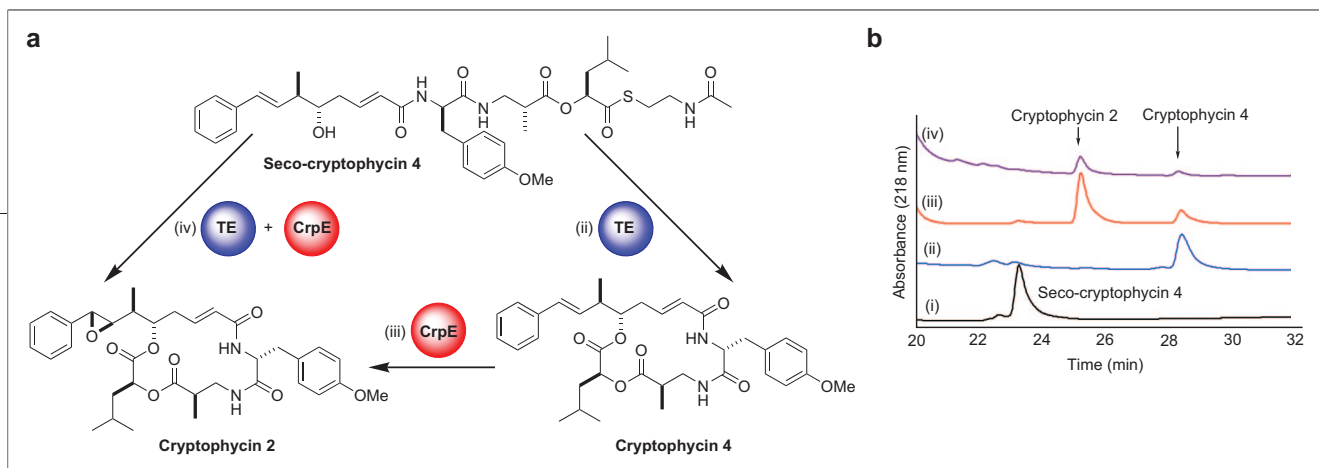


Figure 4. CrpD TE mediated cyclization and CrpE mediated epoxidation. **a)** Illustration of the reactions catalyzed by CrpD TE and CrpE. **b)** HPLC traces of: i, 10.0 μM SNAC-thioester of seco-cryptophycin 4 (black line); ii, CrpD TE catalyzed cyclization of 9.0 μM seco-cryptophycin 4 (blue line); iii, CrpE mediated epoxidation of 9.5 μM cryptophycin 4 (previously cyclized by CrpD TE and HPLC purified) to produce cryptophycin 2 (red line); iv, tandem (one reaction vessel) CrpD TE and CrpE mediated cyclization and epoxidation of 5.2 μM seco-cryptophycin 4 (purple line). The reactions contained 0.1 M Tris pH 7.4, 5% DMSO, 1.8 μM Crp TE and/or 3 μM CrpE with 100 $\mu\text{g mL}^{-1}$ ferredoxin, 0.2 unit mL^{-1} ferredoxin-NADP⁺ reductase, 1.4 mM NADPH, 10 mM glucose-6-phosphate, 8 units mL^{-1} glucose-6-phosphate dehydrogenase at 30 °C from 2 to 4 h.

units B, C, and D. Although unit A was shown to be derived from three acetate groups, the ultimate source of the “phenylacetate” starter group requires further investigation. Phenylalanine is the likely Crp pathway initiation group, but the precise steps involved in conversion to the “phenylacetate” equivalent remains unknown. Precursor-directed incorporation studies revealed additional flexibility for the Crp biosynthetic pathway through its ability to accept a number of struc-

turally diverse starter units that were channeled, processed, cyclized, and released as fully elaborated novel depsipeptide natural products. Finally, we demonstrated for the first time a stereospecific epoxidation reaction resulting in effective conversion of cryptophycin 4 to cryptophycin 2 using the CrpE CYP450. This unique biocatalyst, along with the versatile CrpD TE domain, provides new, useful tools for the chemoenzymatic synthesis of Crp anticancer agents.

METHODS

General Microbiology and Molecular Biology Procedures: Chemicals, Bacterial Strains, Culture Conditions, and DNA Subcloning.

Molecular biology reagents and enzymes were supplied by New England Biolabs except for Pfu (Stratagene), dNTPs (Takara), T4 DNA Ligase (Invitrogen), pSMART-HCKan (Lucigen), and Wizard SV PCR Clean-up kit (Promega). When necessary, other chemicals were purchased from Sigma-Aldrich. *Nostoc* cultures were routinely maintained on BG-11 agar plates prepared as previously described (44). *Nostoc* strains prepared for genomic DNA extraction and production of Crp were grown in BG-11 medium (44) at 28 °C with constant illumination (9 W m^{-2}) and aeration. The *E. coli* strains used for cloning and plasmid harvesting were XL-1 Blue (Stratagene) and TOP-10 (Invitrogen). Protein overexpression was performed in *E. coli* BL21 (DE3) (Invitrogen). All *E. coli* strains were grown in Luria-Bertani (LB) broth and when necessary supplemented with 25 $\mu\text{g mL}^{-1}$ kanamycin. Preparation and manipulation of plasmid DNA from *E. coli* were accomplished using standard methods (45). DNA sequencing was performed at the University of Michigan DNA Sequencing Core or the University of Minnesota Advanced Genetic Analysis Center using ABI Prism sequencers and the dye termination method.

Subtractive Analysis of *Nostoc* A and KS Domains. Amplicons encoding both KS domain (~750 bp) and A domain fragments (~1100 bp) were obtained using conditions and PCR primers described previously (46, 47). The resulting amplicons were cloned into the sequencing vector pCR2.1 (Invitrogen) and 96 clones bearing inserts for both KS and A domains were selected for DNA sequencing. Sequences of KS and A domains were compared by multiple sequence alignment using the Clustal X program (48) (version 1.82) with A and KS domains from the deduced sequence of NRPSs and PKSs encoded by ORFs identi-

fied within the *N. punctiforme* genome sequence (GenBank accession # NZ_AAY0000000).

Cloning, Sequencing, and Annotating the Crp Cluster. High molecular weight genomic DNA was harvested from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 according to established protocols (49) and also using a FASTDNA spin kit (Qbiogene). For cosmid library construction total *Nostoc* DNA was partially digested using *Sau3A*I, dephosphorylated, and size selected (36–40 kb) using a CHEF gel with a CHEF-DR III PFGE system (Biorad), followed by ligation into *Bam*HI-digested SuperCos1 (Stratagene). The ligation mixture was packaged using the Giga-pack III XL Packaging Extract Kit (Stratagene), and the resulting library was titered and amplified according to manufacturer instructions. Fosmid libraries were constructed in an analogous fashion but without prior restriction enzyme digestion, and blunt-ended fragments were cloned directly into pCC1Fos fosmid according to manufacturer instructions (Epicentre). The insert within pNAM123 (encoding a portion of an A domain) was liberated by *Eco*RI digestion, and the resulting DNA fragment was radiolabeled using the RadPrime labeling kit (Pharmacia) with [α -³²P] dCTP (Amersham) according to manufacturer directions. The radiolabeled fragments were used to probe the genomic library using standard colony hybridization protocols (45). The cosmid pDHS500, identified from probing with the pNAM123 insert, was fully sequenced using a shot-gun cloning approach where pDHS500 was fragmented in a nebulizer (IPT Medical Products, Inc.) with the following parameters, 4.4×10^4 Pa of N_2 for 4.0 min. Likewise, the fosmid DNA from pDHS501 was nebulized and the 2–6 kb fragments collected. The resulting pDHS500 and 501 fragments were blunt-ended by Klenow and T7 DNA polymerase and phosphorylated by treatment with T4 kinase and ATP immediately prior to fractionation on a low melting agarose gel. Fragments in the 2–6 kb size range were eluted, concentrated, and used separately for ligation into

the *Sma*I site of pUC18 and pSMART-HCKan (Lucigen). The sequences of pUC18 and pSMART-HCKan inserts were trimmed to remove sequences of SuperCos-1 and pCC1Fos and assembled into contigs using the SeqMan 5.06 program of the DNASTAR software package. The pDHS501 fosmid was identified by a PCR-screening strategy employing primers designed from the *crpD* TE region (16) and a second set (5'-GCATTGTC-ATTCTGGTGAGGC-3' and 5'-CCTGCTGCTAAGGCTATTCCAAG-3') for a portion of *crpB* contained within pDHS500. Two rounds of PCR were used to identify clones that produced amplicons using the PKS primers, but not the TE domain primers. The first round of PCR was used to identify plates that contained amplicons of the desired size. pDHS500 was used as a positive control for both primer sets. Ten-microliter portions of each culture from a given 96-well plate were pooled. The DNA was then used in PCR reactions using the TE and the PKS primer sets. The PCR reactions contained 1 μ L of pooled DNA, 1 μ M of each primer, 200 μ M dNTP, 5 μ L of Taq polymerase buffer, and 0.5 μ L of Taq DNA polymerase, and water was added to a final volume of 50 μ L. The PCR reactions were run at 94 $^{\circ}$ C for 2 min and cycled 30 times with the program: 94 $^{\circ}$ C for 1 min, 50 $^{\circ}$ C for 1 min, 72 $^{\circ}$ C for 1 min. A *crpB* PKS amplicon was generated from one of the clones with the *crpB* primer set but not with the *crpD* TE primer set. The clone (pDHS501) producing this amplicon was identified and shown to extend from the 5' end of pDHS500. ORFs from pDHS500 and pDHS501 were identified using Vector NTI software package version 10 and their deduced amino acid sequences compared to others in the GenBank database. The domain arrangements of the Crp PKS and NRPS proteins were determined using the web-based software program PKS-NRPS (50). The PKS-NRPS program was also used to determine the predicted specificity codes of the Crp NRPS A domains. The *crp* gene cluster DNA sequence has been deposited in GenBank under accession number EF159954.

Overexpression and Purification of CrpE. The CrpE gene was amplified by PCR using cosmid pDHS500 as template. A typical 50 μ L reaction mixture contained 5 ng of pDHS500, 2 μ M forward primer (5'-TGC GGA TCC ATG ATT AAT ACT GCT AAA TCC-3') and 2 μ M reverse primer (5'-ACG CGA ATT CTT ACA ATA CAA CCA TTT TTA ATC C-3') (*Bam*HI and *Eco*RI sites are underlined), 200 μ M dNTP, 5 μ L of 10 \times PCR buffer. Conditions for *crpE* amplifications included an initial 5 min denaturation step (94 $^{\circ}$ C) and cycling conditions of 94 $^{\circ}$ C (30 s), 58.5 $^{\circ}$ C (30 s), and 72 $^{\circ}$ C (1 min 45 s) for 35 cycles followed by a final 72 $^{\circ}$ C extension step (7 min). The *crpE* amplicon was phosphorylated using T4 kinase and ligated into pSMART-HCKan to produce pDing1. The insert within pDing1 was sequenced and shown to be free of PCR errors. pDing1 was digested with *Bam*HI and *Eco*RI and ligated into linear pSj8 (*Eco*RI/*Bam*HI cut). The resulting plasmid, pDing2, and chaperone expressing plasmid pGRO7 were used to cotransform *E. coli* BL21(DE3) to ampicillin (amp) (50 μ g mL⁻¹) and chloramphenicol (chl) (25 μ g mL⁻¹) resistance. A 5 mL overnight culture was diluted in 1 L of LB supplemented with amp (50 μ g mL⁻¹), chl (25 μ g mL⁻¹), 0.25 mM Fe(NH₄)₂(SO₄)₂, 1 mM thiamine, and 0.25 mM of 5-aminolevulinic acid that was added 30 min prior to induction (OD₆₀₀ ~ 0.6). The culture was then cooled to 4 $^{\circ}$ C before inducing with 0.1 mM Isopropyl- β -D-thiogalactopyranoside and 1 g L mL⁻¹ of L-arabinose. The culture was grown at 15 $^{\circ}$ C with constant shaking (200 rpm) for 20 h. The cells were pelleted by centrifugation and resuspended in 80 mL of PBS buffer (140 mM NaCl, 2.7 mM potassium chloride, 10 mM sodium hydrogen phosphate, and 1.8 mM potassium dihydrogen phosphate, pH 7.4, 3 mM β -mercaptoethanol, 10% glycerol). The cell suspension was then sonicated and the lysate collected following centrifugation (40,000g for 45 min). The resulting supernatant was collected and incubated with pre-equilibrated amylase agarose

resin at 4 $^{\circ}$ C for 3 h with agitation. The amylase agarose resin was washed (100 column volumes of lysis buffer) and MBP-His-CrpE eluted with lysis buffer containing 8 mM maltose. Maltose was removed from the protein sample with a PD-10 column, and the MBP-His-CrpE fusion protein was then treated with His-TEV protease at 4 $^{\circ}$ C overnight to remove the MBP-His portion. The CrpE protein was separated from the MBP-His polypeptide by passing the mixture through a Ni-agarose column. The concentration of the purified protein was determined by its predicted extinction coefficient (81,820 M⁻¹ cm⁻¹ at 280 nm). The active CrpE concentration was determined using previously described methods (51).

Generation of Cryptophycin 4 Using the Cryptophycin TE. Synthesis of seco-SNAC-cryptophycin 4 and heterologous expression and purification of the Crp TE (CrpD TE) were conducted following published procedures (16). Seco-SNAC-cryptophycin 4 (2.4 mg) was dissolved to 2 mM in DMSO and then diluted to 100 μ M in 0.1 M Tris pH 7.0. A 10 μ M portion of CrpD TE was added to this mixture and incubated at 30 $^{\circ}$ C for 15 h. The reaction was next extracted with ethyl acetate (3 \times 25 mL), and the ethyl acetate fractions were pooled and concentrated. The total contents of the ethyl acetate extraction were separated by semi-preparative reversed-phase HPLC (C18 Econosil, 10 \times 250 mm, 5 mL min⁻¹, 10–100% acetonitrile/water, 50 min). The peak corresponding to cryptophycin 4 was collected and lyophilized. HPLC $t_{\text{ret}} = 36.3$ min; mass spectrometry (MS) (ESI⁺) m/z 605.2, [M + H]⁺ (C₃₅H₄₅N₂O₇ requires 605.3).

CrpE Reactions and HPLC Analysis. CrpE reactions contained 100 μ g mL⁻¹ ferredoxin, 0.2 units mL⁻¹ of ferredoxin-NADP+ reductase, 1.4 mM of NADPH, 10 mM of glucose-6-phosphate, 8 units mL⁻¹ of glucose-6-phosphate dehydrogenase, 5.2 μ M seco-SNAC-cryptophycin 4, and 1.8 μ M CrpD TE or 9 μ M cryptophycin 4 in 100 μ L of lysis buffer. The reaction mixtures were incubated at 30 $^{\circ}$ C for 2 min, and then 3 μ M CrpE and/or 1.8 μ M CrpD TE was added to initiate the reaction at 30 $^{\circ}$ C for 2–4 h. A control reaction was run in parallel that contained boiled CrpE in place of active CrpE. The reaction mixtures were then separated by analytical reversed-phase HPLC (C18 Econosil, 4.6 \times 250 mm, 1 mL min⁻¹, 30–100% acetonitrile/water + 0.1% TFA, 40 min, 218 nm). Reaction of CrpE with cryptophycin 4 yielded cryptophycin 2 as determined by LC/MS $t_{\text{ret}} = 25.3$ min; MS (ESI⁺) m/z 621.2, [M + H]⁺ (C₃₅H₄₅N₂O₈ requires 621.3), and m/z 643.2, [M + Na]⁺ (C₃₅H₄₄N₂NaO₈ requires 643.3). A large-scale reaction was performed using the same ratio of reagents but 100 \times the volume. For this reaction, spinach ferredoxin reductase was purified from spinach leaves using modified protocols (52). Briefly, 3 kg of crude homogenate obtained from grinding in a Waring blender was fractionated with acetone according to a previously published method (53). The precipitate obtained was dissolved in 58 mM Tris, pH 7.5, and passed through three Hi-Trap Q 5 mL columns in series using an isocratic flow of 5 mL min⁻¹ in the same buffer. The yellow flow-through fractions were concentrated and used directly in the reaction. Purified epoxidized material was analyzed by NMR using a Varian INOVA 600 NMR spectrometer, equipped with a 5 mm HCN probe. The Crp sample was dissolved in 200 μ L of CD₃OD. The NMR spectrum was identical to that reported for cryptophycin 2 with the chemical shifts for the epoxide protons at 3.94 (dd, $J = 7.5, 1.8$ Hz), 3.70 (d, $J = 1.8$ Hz), and 1.80 (m) for the adjacent ring proton (42).

General Chemical Procedures: Labeled Precursors, Stable-Isotope Feedings, and Directed Biosynthesis. Sodium [1,2-¹³C₂] acetate, sodium [2-¹³C,²H₃]acetate, sodium [1-¹³C,¹⁸O]acetate, sodium [U-¹³C₃]pyruvate, L-[methyl-¹³C]methionine (1³C, 96%), L-[U-¹³C₆,¹⁵N]phenylalanine, L-[²H₈]phenylalanine, [1-¹³C]phenylacetic acid, *p*-tolylacetic acid, L-[1-¹³C]tyrosine, DL-[2-¹³C,¹⁵N] aspartic acid, DL-[2,3,3-²H₃]aspartic acid, L-[1-¹³C]leucine, and

[1-¹³C]2-ketocaproic acid were obtained from the Aldrich Chemical Co. The methods for all of the feeding study experiments are contained within the Supporting Information section. In general, *Nostoc* sp. GSV 224 or *Nostoc* sp. ATCC 53789 were cultured in 20-L glass carboys as previously described (3). All feeding studies were performed with both *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 with essentially identical results. The isotopically labeled or biosynthetic precursor was added to each of two 20-L cultures of *Nostoc* sp. GSV 224 or *Nostoc* sp. ATCC 53789 beginning on day 10 after inoculation unless noted otherwise. For precursor-directed biosynthesis, an aqueous solution of the precursor was added all at once or in aliquots (2–8) every other day. Usually an amino acid precursor was dissolved in 0.5 N HCl to a concentration of 10–30 mg mL⁻¹ and fed in 8 aliquots. Typically, a 0.5 mL portion of the solution was added to each of the 2–4 carboys of *Nostoc* sp. GSV 224 culture in 2-d intervals beginning on day 10–12 after inoculation. After 6–10 additions of the amino acid solution, the cultures were allowed to grow for an additional 3–5 d and then harvested. ¹H and ¹³C NMR spectra were obtained at 500 and 125 MHz in CDCl₃. All ¹³C, ¹⁵N, ²H, and ¹⁸O-labeled compounds used in this study were 99, 96–99, 98, and 95 atom %, respectively, unless noted otherwise. Low-resolution MS was performed at the University of Michigan Mass Spectrometry Laboratory on a Waters Ultima magnetic sector mass spectrometer equipped with an electrospray interface.

Isolation of Labeled Cryptophycins. Lyophilized *Nostoc* sp. GSV 224 (10–30 g) was extracted with 4:1 mixture of CH₃CN/CH₂Cl₂ (40 mL g⁻¹) for 48 h and the extract concentrated *in vacuo* to give a dark green solid. This solid (1 g) was applied to an ODS-coated silica column (55 g, 15 × 2.5 cm) and subjected to flash chromatography with 1:3 CH₃CN/H₂O (0.4 L), 1:1 CH₃CN/H₂O (0.4 L), 65:35 CH₃CN/H₂O (0.8 L), CH₃OH (0.4 L), and CH₂Cl₂ (0.4 L). The material was eluted with 65:35 CH₃CN/H₂O (100–400 mg) and further separated by reversed-phase HPLC (Econosil C18, 10 μm, 25 cm × 22 mm, UV detection at 254 nm, 65:35 CH₃CN/H₂O, flow rate 6 mL min⁻¹) to give labeled cryptophycin 1 (*t*_r = 53.0 min) and a number of fractions containing mixtures of other labeled cryptophycins. The natural cryptophycins eluted between 25 and 100 min. Labeled analogues were isolated using previously described procedures (3). The isolation of deuterated cryptophycins was monitored by ²H NMR spectroscopy. The methods and results of the feeding experiments, including all associated structure elucidation information are contained within the Supporting Information section available online.

Acknowledgments: We kindly acknowledge Zhaohui Xu for pSj8 used for CrpE expression, Sabine Gruschow for LC/MS analysis, and Frank Schroeder for NMR analysis. We thank Wu Du, Ghosen Ye, and Jian Liang for expert technical assistance. This research was supported by National Institutes of Health (NIH) Postdoctoral Training Fellowship (NCI CA09676) to Z.Q.B., NIH Grants CA83155 and CA108874 and the Searle Professorship to D.H.S., and NIH Grant CA12623 and National Science Foundation Grant CHE-9530794 to R.E.M.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Chaganty, S., Golakoti, T., Heltzel, C., Moore, R. E., and Yoshida, W. Y. (2004) Isolation and structure determination of cryptophycins 38, 326, and 327 from the terrestrial cyanobacterium *Nostoc* sp. GSV 224, *J. Nat. Prod.* 67, 1403–1406.
- Schwartz, R. E., Hirsch, C. F., Sesin, D. F., Flor, J. E., Chartrain, M., Fromtling, R. E., Harris, G. H., Salvatore, M. J., Liesch, J. M., and Yudin, K. (1990) Pharmaceuticals from cultured algae, *J. Ind. Microbiol. Biotechnol.* 5, 113–123.
- Golakoti, T., Ohtani, I., Patterson, G. M. L., Moore, R. E., Corbett, T. H., Valeriote, F. A., and Demchik, L. (1994) Total structures of cryptophycins, potent antitumor depsipeptides from the blue-green-alga *Nostoc* sp. strain GSV-224, *J. Am. Chem. Soc.* 116, 4729–4737.
- Subbaraju, G. V., Golakoti, T., Patterson, G. M., and Moore, R. E. (1997) Three new cryptophycins from *Nostoc* sp. GSV 224, *J. Nat. Prod.* 60, 302–305.
- Golakoti, T., Yoshida, W. Y., Chaganty, S., and Moore, R. E. (2001) Isolation and structure determination of nostocyclopeptides A1 and A2 from the terrestrial cyanobacterium *Nostoc* sp. ATCC 53789, *J. Nat. Prod.* 64, 54–59.
- Smith, C. D., Zhang, X., Mooberry, S. L., Patterson, G. M., and Moore, R. E. (1994) Cryptophycin: a new antimicrotubule agent active against drug-resistant cells, *Cancer Res.* 54, 3779–3784.
- Lu, K., Dempsey, J., Schultz, R. M., Shih, C., and Teicher, B. A. (2001) Cryptophycin-induced hyperphosphorylation of Bcl-2, cell cycle arrest and growth inhibition in human H460 NSCLC cells, *Cancer Chemother. Pharmacol.* 47, 170–178.
- Edelman, M. J., Gandara, D. R., Hausner, P., Israel, V., Thornton, D., DeSanto, J., and Doyle, L. A. (2003) Phase 2 study of cryptophycin 52 (LY355703) in patients previously treated with platinum based chemotherapy for advanced non-small cell lung cancer, *Lung Cancer* 39, 197–199.
- D'Agostino, G., del Campo, J., Mellado, B., Izquierdo, M. A., Minarik, T., Ciri, L., Marini, L., Perez-Gracia, J. L., and Scambia, G. (2006) A multicenter phase II study of the cryptophycin analog LY355703 in patients with platinum-resistant ovarian cancer, *Int. J. Gynecol. Cancer* 16, 71–76.
- Liang, J., Moore, R. E., Moher, E. D., Munroe, J. E., Al-awar, R. S., Hay, D. A., Varie, D. L., Zhang, T. Y., Aikins, J. A., Martinelli, M. J., Shih, C., Ray, J. E., Gibson, L. L., Vasudevan, V., Polin, L., White, K., Kushner, J., Simpson, C., Pugh, S., and Corbett, T. H. (2005) Cryptophycins-309, 249 and other cryptophycin analogs: preclinical efficacy studies with mouse and human tumors, *Invest. New Drugs* 23, 213–224.
- Eggen, M., and Georg, G. I. (2002) The cryptophycins: their synthesis and anticancer activity, *Med. Res. Rev.* 22, 85–101.
- Barrow, R. A., Hemscheidt, T., Liang, J., Paik, S., Moore, R. E., and Tius, M. A. (1995) Total synthesis of cryptophycins—revision of the structures of cryptophycin-A and cryptophycin-C, *J. Am. Chem. Soc.* 117, 2479–2490.
- Gardinier, K. M., and Leahy, J. W. (1997) Enantiospecific total synthesis of the potent antitumor macrolides cryptophycins 1 and 8, *J. Org. Chem.* 62, 7098–7099.
- Chang, H. T., and Sharpless, K. B. (1996) Molar scale synthesis of enantiopure stilbene oxide, *J. Org. Chem.* 61, 6456–6457.
- Liang, J., Moher, E. D., Moore, R. E., and Hoard, D. W. (2000) Synthesis of cryptophycin 52 using the Sharpless asymmetric dihydroxylation: diol to epoxide transformation optimized for a base-sensitive substrate, *J. Org. Chem.* 65, 3143–3147.
- Beck, Z. Q., Aldrich, C. C., Magarvey, N. A., Georg, G. I., and Sherman, D. H. (2005) Chemoenzymatic synthesis of cryptophycin/arenastatin natural products, *Biochemistry* 44, 13457–13466.
- Chang, Z., Flatt, P., Gerwick, W. H., Nguyen, V. A., Willis, C. L., and Sherman, D. H. (2002) The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit, *Gene* 296, 235–247.
- Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., Sherman, D. H., and Gerwick, W. H. (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*, *J. Nat. Prod.* 67, 1356–1367.
- Meeks, J. C., Elhai, J., Thiel, T., Potts, M., Larimer, F., Lamerdin, J., Predki, P., and Atlas, R. (2001) An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium, *Photosyn. Res.* 70, 85–106.

20. Hunsucker, S. W., Klage, K., Slaughter, S. M., Potts, M., and Helm, R. F. (2004) A preliminary investigation of the *Nostoc punctiforme* proteome, *Biochem. Biophys. Res. Commun.* **317**, 1121–1127.
21. Challis, G. L., Ravel, J., and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, *Chem. Biol.* **7**, 211–224.
22. Edwards, D. J., Marquez, B. L., Nogle, L. M., McPhail, K., Goeger, D. E., Roberts, M. A., and Gerwick, W. H. (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*, *Chem. Biol.* **11**, 817–833.
23. Hicks, L. M., Moffitt, M. C., Beer, L. L., Moore, B. S., and Kelleher, N. L. (2006) Structural characterization of *in vitro* and *in vivo* intermediates on the loading module of microcystin synthetase, *ACS Chem. Biol.* **1**, 93–102.
24. Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases, *ChemBiochem* **4**, 654–657.
25. Reid, R., Piagentini, M., Rodríguez, E., Ashley, G., Viswanathan, N., Camey, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases, *Biochemistry* **42**, 72–79.
26. Siskos, A. P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spitteller, D., Popovic, B., Spencer, J. B., Staunton, J., Weissman, K. J., and Leadlay, P. F. (2005) Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases, *Chem. Biol.* **12**, 1145–1153.
27. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis, *Chem. Rev.* **97**, 2651–2673.
28. Rouhiainen, L., Paulin, L., Suomalainen, S., Hyttiäinen, H., Buikema, W., Haselkom, R., and Sivonen, K. (2000) Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90, *Mol. Microbiol.* **37**, 156–167.
29. Yeh, E., Gameau, S., and Walsh, C. T. (2005) Robust *in vitro* activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebecamycin biosynthesis, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3960–3965.
30. Dorrestein, P. C., Yeh, E., Gameau-Tsodikova, S., Kelleher, N. L., and Walsh, C. T. (2005) Dichlorination of a pyrrolyl-S-carrier protein by FADH₂-dependent halogenase P1tA during pyoluteorin biosynthesis, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13843–13848.
31. Jackowski, S. (1996) Biosynthesis of pantothenic acid and coenzyme A, in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Ed.), pp 687–694, American Society for Microbiology, Washington, DC.
32. Wolf, J. P., and Rapoport, H. (1989) Conformationally constrained peptides. Chiro-specific synthesis of 4-alkyl-substituted gamma-lactam-bridged dipeptides from L-aspartic acid, *J. Org. Chem.* **54**, 3164–3173.
33. Ogasawara, Y., Katayama, K., Minami, A., Otsuka, M., Eguchi, T., and Kakinuma, K. (2004) Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*, *Chem. Biol.* **11**, 79–86.
34. Nishida, H., Eguchi, T., and Kakinuma, K. (2001) Amino acid starter unit in the biosynthesis of macrolactam polyketide antitumor antibiotic vicenistatin, *Tetrahedron* **57**, 8237–8242.
35. Umbarger, H. E. (1996) Biosynthesis of branched-chain amino acids, in *Escherichia coli and Salmonella: Cellular and Molecular Biology 1* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) 2nd ed., pp 442–457, American Society for Microbiology, Washington, DC.
36. Cheng, Y. Q. (2006) Deciphering the biosynthetic codes for the potent anti-SARS-CoV cyclodepsipeptide valinomycin in *Streptomyces tsusimaensis* ATCC 15141, *ChemBiochem* **7**, 471–477.
37. Ehling-Schulz, M., Fricker, M., Gallert, H., Rieck, P., Wagner, M., and Scherer, S. (2006) Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1, *BMC Microbiol.* **6**, 20.
38. Magarvey, N. A., Ehling-Schulz, M., and Walsh, C. T. (2006) Characterization of the cereulide NRPS α -hydroxy acid specifying modules: activation of α -keto acids and chiral reduction on the assembly line, *J. Am. Chem. Soc.* **128**, 10698–10699.
39. Zañeta-Rivera, K., Xu, C., Yu, F., Butchko, R. A., Proctor, R. H., Hidalgo-Lara, M. E., Raza, A., Dussault, P. H., and Du, L. (2006) A bidomain nonribosomal peptide synthetase encoded by FUM14 catalyzes the formation of tricarballic esters in the biosynthesis of fumonisins, *Biochemistry* **45**, 2561–2569.
40. Pee, K.-H. v., and Patallo, E. P. (2006) Flavin-dependent halogenases involved in secondary metabolism in bacteria, *Appl. Microbiol. Biotechnol.* **70**, 631–641.
41. Al-Awar, R. S., Ray, J. E., Schultz, R. M., Andis, S. L., Kennedy, J. H., Moore, R. E., Liang, J., Golakoti, T., Subbaraju, G. V., and Corbett, T. H. (2003) A convergent approach to cryptophycin 52 analogues: synthesis and biological evaluation of a novel series of fragment A epoxides and chlorohydrins, *J. Med. Chem.* **46**, 2985–3007.
42. Ghosh, A. K., and Bischoff, A. (2000) A convergent synthesis of (+)-cryptophycin B, a potent antitumor macrolide from *Nostoc* sp. cyanobacteria, *Org. Lett.* **2**, 1573–1575.
43. Rikkinen, J., Oksanen, I., and Lohtander, K. (2002) Lichen guilds share related cyanobacterial symbionts, *Science* **297**, 357.
44. Golden, S. S., Brusslan, J., and Haselkom, R. (1987) Genetic engineering of the cyanobacterial chromosome, *Methods Enzymol.* **153**, 215–231.
45. Sambrook, J., and Russell, D. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
46. Moffitt, M. C., and Neilan, B. A. (2004) Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins, *Appl. Environ. Microbiol.* **70**, 6353–6362.
47. Neilan, B. A., Dittmann, E., Rouhiainen, L., Bass, R. A., Schaub, V., Sivonen, K., and Borner, T. (1999) Nonribosomal peptide synthesis and toxicity of cyanobacteria, *J. Bacteriol.* **181**, 4089–4097.
48. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* **25**, 4876–4882.
49. Cohen, M. F., Wallis, J. G., Campbell, E. L., and Meeks, J. C. (1994) Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives, *Microbiology* **140**, 3233–3240.
50. Ansari, M. Z., Yadav, G., Gokhale, R. S., and Mohanty, D. (2004) NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases, *Nucleic Acids Res.* **32**, W405–W413.
51. Omura, T., and Sato, R. (1964) The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification, and properties, *J. Biol. Chem.* **239**, 2379–2385.
52. Schurmann, P. (1995) Ferredoxin: thioredoxin system, *Methods Enzymol.* **252**, 274–283.
53. Shin, M. (1971) Ferredoxin-NADP reductase from spinach, *Methods Enzymol.* **23**, 440–447.
54. Donadio, S., and Katz, L. (1992) Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*, *Gene* **111**, 51–60.

A Small-Molecule Inhibitor for Phosphatase and Tensin Homologue Deleted on Chromosome 10 (PTEN)

Erika Rosivatz^{†,*}, Jonathan G. Matthews[†], Neil Q. McDonald^{*,§}, Xavier Mulet^{†,‡}, Ka Kei Ho[†], Nadine Lossi^{†,‡,¶}, Annette C. Schmid^{†,¶}, Marianna Mirabelli^{||,††,‡‡}, Karen M. Pomeranz^{**}, Christophe Erneux^{§§}, Eric W.-F. Lam^{**}, Ramón Vilar^{||,††}, and Rüdiger Woscholski^{†,‡,*,¶}

[†]Division of Cell and Molecular Biology, Imperial College London, Exhibition Road, London SW7 2AZ, U.K., ^{*}Structural Biology Laboratory, The London Research Institute, Cancer Research UK, 44 Lincoln's Inn Fields, London WC2A 3PX, U.K., [§]School of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, U.K., ^{||}The Institute of Chemical Research of Catalonia (ICIQ), Avda. Paisos Catalans 16, 43007 Tarragona, Spain, ^{**}Cancer Research UK Labs, Department of Oncology, Imperial College London, Du Cane Road, London W12 0NN, U.K., ^{††}Department of Chemistry, Imperial College London, Exhibition Road, London SW7 2AZ, U.K., ^{‡‡}Chemical Biology Centre, Imperial College London, Exhibition Road, London SW7 2AZ, U.K., ^{§§}Interdisciplinary Research Institute (IRIBHM), Université libre de Bruxelles, Campus Erasme, 808 Route de Lennik, 1070 Brussels, Belgium, [¶]Current address: Division of Investigative Science, Faculty of Medicine, Imperial College London, Exhibition Road, London SW7 2AZ, U.K., ^{¶¶}Current address: DoNatur GmbH, Am Klopferspitz 19, 82152 Martinsried, Germany

ABSTRACT Phosphatase and tensin homologue deleted on chromosome 10 (PTEN), a phosphoinositide 3-phosphatase, is an important regulator of insulin-dependent signaling. The loss or impairment of PTEN results in an antidiabetic impact, which led to the suggestion that PTEN could be an important target for drugs against type II diabetes. Here we report the design and validation of a small-molecule inhibitor of PTEN. Compared with other cysteine-based phosphatases, PTEN has a much wider active site cleft enabling it to bind the PtdIns(3,4,5)P₃ substrate. We have exploited this feature in the design of vanadate scaffolds complexed to a range of different organic ligands, some of which show potent inhibitory activity. A vanadyl complexed to hydroxypicolinic acid was found to be a highly potent and specific inhibitor of PTEN that increases cellular PtdIns(3,4,5)P₃ levels, phosphorylation of Akt, and glucose uptake in adipocytes at nanomolar concentrations. The findings presented here demonstrate the applicability of a novel and specific chemical inhibitor against PTEN in research and drug development.

*Corresponding authors,
r.woscholski@ic.ac.uk or e.rosivatz@ic.ac.uk.

Received for review August 11, 2006
and accepted November 14, 2006.

Published online December 8, 2006

10.1021/cb600352f CCC: \$33.50

© 2006 by American Chemical Society

Phosphatase and tensin homologue deleted on chromosome 10 (PTEN) shares the CX5R motif, present at the bottom of the active site pocket, with the large family of tyrosine phosphatases (PTPs) and dual specificity protein phosphatases but has little sequence homology to these phosphatase families outside this motif (1). Although there is evidence that PTEN displays protein tyrosine phosphatase activity, it is its lipid phosphatase activity that is associated with PTEN's cellular impacts, such as inhibition of proliferation, survival, and regulation of insulin signaling (2). In particular, the dephosphorylation of the D3 position of the inositol head group of PtdIns(3,4,5)P₃ will counteract phosphoinositide 3-kinase (PI3K) action (3), explaining why the loss of PTEN's function is linked with the development and progression of tumors (4–7). On the other hand, its deregulation and hyperactivation have been implicated in type II diabetes mellitus due to compromised insulin signaling (8–10). Furthermore, in a muscle-specific knock-out of PTEN, a reduction in the development of high-fat-induced insulin resistance and diabetes could be observed (11). PTEN deficiency in mouse adipose tissue leads to decreased production of the hormone resistin, which positively correlates with insulin resistance (12). It has been suggested that impaired control of physiological PTEN inhibition might

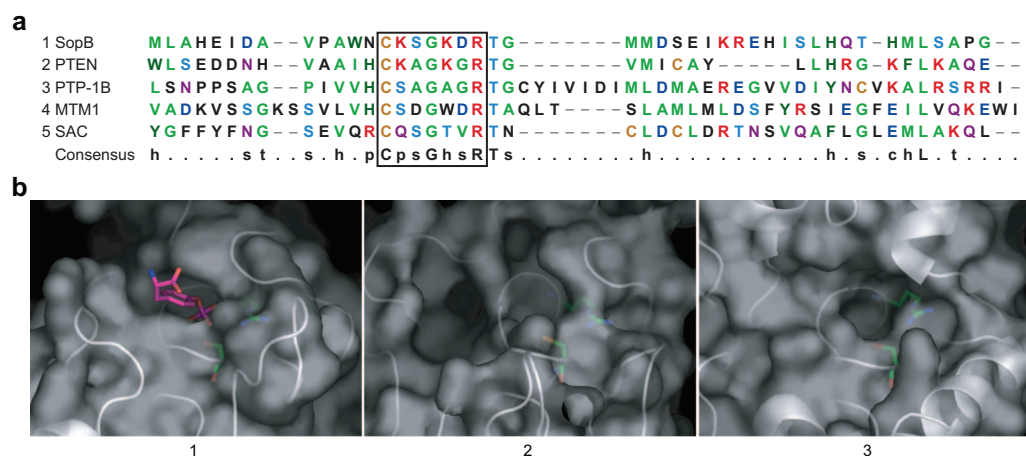


Figure 1. CPB active sites a) Alignment of the CX₅R motif surroundings. b) Molecular surface around the active site of PTP-1B (1), PTEN (2), and MTM (3), drawn using the program Pymol (38). Side chains for the active site cysteine and arginine from the CX₅R motif are also shown with the ribbon representation of the peptide backbone. The PTP-1B structure is of a substrate-trapping mutant active site C215S mutation and the position of a phosphotyrosine is also indicated.

be a factor in conferring leptin resistance and obesity in animals and humans (13). In addition, Nakashima *et al.* (14) report that microinjection of an anti-PTEN antibody increased basal and insulin-stimulated PtdIns(3,4,5)P₃ levels and the translocation of glucose transporter type 4 (GLUT4) to the plasma membrane, where its exocytosis enables the uptake of glucose (15).

While the role of PTEN on PI3K-dependent PtdIns(3,4,5)P₃ generation seems to be well supported, it remains to be seen whether GLUT4 translocation and glucose uptake are controlled by PTEN's PI 3-phosphatase activity. For example, Tang *et al.* (16) show that knock-down of PTEN by small interfering RNA enhances insulin-dependent glucose uptake into adipocytes, whereas Mosser *et al.* (17) propose that PTEN activity does not influence GLUT4 translocation and other metabolic functions of insulin. They found that overexpression of wild-type PTEN abolishes GLUT4 translocation to the same extent as overexpression of PTEN mutants without lipid phosphatase activity. Nevertheless, they do find that overexpression of wild-type PTEN antagonizes the metabolic actions of insulin in a PI3K-dependent fashion, because overexpression of wild-type PTEN significantly decreased Elk-1 phosphorylation in response to chronic insulin treatment.

Based on the current understanding of PTEN's role in diabetes, the idea has recently gained credibility that the inhibition of its lipid phosphatase activity might

increase glucose uptake triggered by the increase of PtdIns(3,4,5)P₃ (8, 18). Therefore, small molecules that inhibit PtdIns(3,4,5)P₃ phosphatase activity have the potential of enhancing insulin sensitivity and overcoming insulin resistance, which would be beneficial for the development of diabetes therapeutics. Because of its role in insulin-dependent signaling, as demonstrated by knock-out studies, PTEN has been proposed as a suitable drug target for antidiabetic treatment (8).

It has been shown that vanadate complexes are able to mimic a variety of insulin-like effects in both *in vitro* and *in vivo* systems (19–21). For example, these complexes enhance glucose uptake (22–26), which has been attributed to their broad inhibitory potency on PTPs (27–30). However, we recently discovered that these compounds are able to inhibit PTEN's phosphatase activity with much higher potency (31). This led us to design, synthesize, and test a PTEN-specific inhibitor, judged by the increase in cellular PtdIns(3,4,5)P₃ and PtdIns(3,4)P₂ levels, Akt phosphorylation, FoxO translocation, and glucose uptake, providing evidence that the insulin-mimetic properties correlate with PTEN inhibition.

RESULTS AND DISCUSSION

The PI 3-phosphatase PTEN is a member of the large family of cysteine-based phosphatases (CBPs), which also contains the tyrosine and dual-specificity phos-

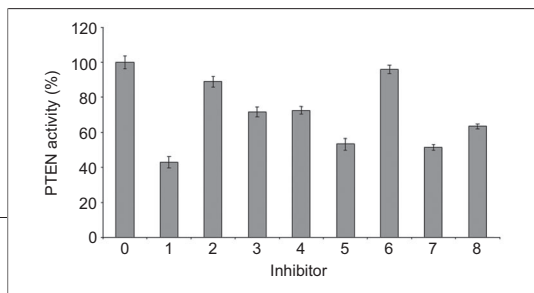


Figure 2. Residual PTEN activity after treatment with 50 nM vanadium compounds *in vitro*. Recombinant PTEN was pre-incubated with eight vanadium compounds for 5 min and then incubated with 3 nmol of PtdIns(3,4,5)P3 for 30 min. PTEN activity was measured by colorimetric determination of inorganic PO_4^{3-} levels and is shown relative to the activity of uninhibited PTEN \pm standard error (SE) of triplicate experiments. 0) solvent H_2O , 1) VO-OHpic, 2) bpV-OHpic, 3) bpV-pic, 4) VO-pic, 5) bpV-biguan, 6) VO-biguan, 7) bpV-phen, 8) bpV-isoqu.

effective inositol lipid 3-phosphatase. This discovery facilitated the elucidation of PTEN's influence on PI3K-dependent signaling pathways. In particular, this phosphatase controlled the Akt-dependent cell survival pathway, which correlated well with its antitumor properties (32). However, more recently, PTEN knock-outs demonstrated a role in insulin signaling and the development of diabetes and obesity, suggesting that PTEN would be a good target for antidiabetic drugs (8). The homology of PTEN with tyrosine phosphatases led us to reason that inhibitors of the latter could bind to the former, if the inhibitors were of sufficiently broad specificity within the tyrosine phosphatase family. In this context, bisperoxovanadates (bpVs) were the obvious choice because they possessed generic tyrosine phosphatase potencies and had antidiabetic properties (33). Our initial characterization revealed that some of these bpVs were more potent PTEN inhibitors than PTP inhibitors (31) but did not target most CBPs with similar affinity. This lack of specificity prompted us to exploit the wide substrate binding site present in the PTEN phosphatase to design a specific vanadium-based inhibitor.

Design and Synthesis of a Library of Small Molecules as Potential PTEN Inhibitors. Vanadium complexes are well known for their ability to mimic phosphoesters and phosphates and thus inhibit a broad range of phosphatases. In particular, bpVs and vanadyl complexed to a variety of organic ligands possess good inhibitory potency against many members of the large tyrosine phosphatase family. While the phosphoinositide 3-phosphatase PTEN is highly homologous to other members of the family of CBPs, including tyrosine phosphatases (Figure 1, panel a), its wide catalytic pocket ($\sim 5 \text{ \AA}$ deep) with an elliptical opening ($\sim 5 \text{ \AA} \times 11 \text{ \AA}$) (34, 35) distinguishes it from all other CBPs (Figure 1, panel b). Therefore, we synthesized (see Methods for details) a range of vanadates and bpV complexes with general formulas $\text{V}(=\text{O})(\text{L}-\text{L}')_2$ (abbreviated herein as VO-ligand) and

phosphatases. Not surprisingly, PTEN was initially thought to act on proteins. This notion was subsequently corrected by Maehama *et al.* (3), who discovered that PTEN is in fact a very

$\text{K}_n[\text{V}(=\text{O})(\text{O}_2)_2(\text{L}-\text{L}')]_n$ (abbreviated herein as bpV-ligand), respectively, and subsequently test their PTEN inhibitor potency.

VO-OHpic Is a Potent and Specific Inhibitor *in Vitro*.

In order to determine the most potent inhibitor, eight vanadium complexes with ligands of different sizes were tested for their ability to inhibit the PtdIns(3,4,5)P3-

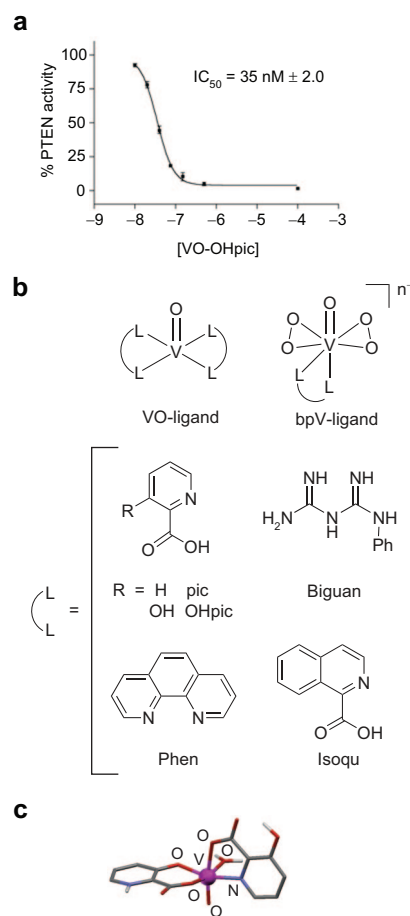


Figure 3. BpV activity and structure a) Inhibition of recombinant PTEN by VO-OHpic *in vitro*. Recombinant PTEN was pre-incubated with between 10 nM and 10 μM VO-OHpic for 5 min and then incubated with 3 nmol of PtdIns(3,4,5)P3 for 30 min. PTEN activity was measured by colorimetric determination of inorganic PO_4^{3-} levels and is shown \pm SE of triplicate experiments. b) Schematic representation of the different compounds under study (NOTE: only the bpVs were prepared with phenanthroline (phen) and 1-isoquinoline (isoqu)). c) Representation of the X-ray structure of VO-OHpic previously reported by Nakai (36).

phosphatase activity of recombinant PTEN at a concentration of 50 nM. The weakest inhibitor is the phenylbiguanide vanadium(V) complex VO-biguan (Figure 2), whereas the 3-hydroxypicolinate vanadium(IV) complex VO-OHpic is the most potent inhibitor (IC₅₀ = 35 nM; Figure 3) of the PTEN lipid phos-

phatase activity; therefore it was further characterized with respect to its specificity. We determined the inhibitory potency of VO-OHpic, as well as the seven other vanadium compounds, against the enzyme activities of four other recombinant CBPs, PTP- β , SAC1, myotubularin (MTM1), and SopB, *in vitro*. The IC₅₀ values of VO-OHpic for the tested mammalian CBP family members (Table 1) are all in the micromolar range, whereas the IC₅₀ for the bacterial SopB is in the high nanomolar range. Since the latter enzyme is not of mammalian origin and needs $>16 \times$ higher doses, one can conclude that VO-OHpic is highly specific for PTEN in mammals. The other vanadium compounds possess broader specificity. Interestingly, MTM seems to prefer vanadium(V) complexes with small ligands, whereas tyrosine phosphatases seem to prefer V-phenanthroline (phen), which is in agreement with earlier observations (31). In comparison, VO-OHpic is an encouragingly specific (Table 1) and potent PTEN inhibitor (Figure 2).

VO-OHpic is a water-soluble vanadium(IV) complex in which the metal center is coordinated to two OHpic ligands, a water molecule, and an oxo ligand, yielding a complex with asymmetric octahedral geometry. Recently, Nakai *et al.* (36) have shown by X-ray crystallography that in the solid state this compound contains one N,O-coordinated and one O,O-coordinated OHpic ligand (see the X-ray structure in Figure 3, panel c). A complex of these dimensions would fit well into the catalytic pocket of PTEN but would be too large for the narrow active site pockets of the other tested phosphatases, such as PTB-1 β and MTM (34, 37).

TABLE 1. *In vitro* specificity of VO-OHpic for recombinant PTEN compared with other CBPs^a

	PTP β (PNPP) μ M	SAC (PI4P) μ M	MTM (PI3P) μ M	SopB (PI4,5P2) nM
VO-OHpic	57.5 \pm 9.4	>10 ^b	4.03 \pm 0.04	588 \pm 16
bpV-OHpic	4.9 \pm 0.9	0.06 \pm 0.01	0.35 \pm 0.02	33 \pm 7
bpV-pic	12.7 \pm 3.2	0.99 \pm 0.01	0.24 \pm 0.01	^c
VO-pic	589 \pm 33	0.34 \pm 0.07	6.35 \pm 3.92	125 \pm 2
bpV-biguan	640 \pm 32	1.00 \pm 0.01	1.81 \pm 0.81	798 \pm 41
VO-biguan	112 \pm 5	0.77 \pm 0.03	4.37 \pm 0.94	811 \pm 88
bpV-phen	0.24 \pm 0.01	0.08 \pm 0.01	0.41 \pm 0.04	^c
bpV-isoqu	349 \pm 27	0.09 \pm 0.01	0.87 \pm 0.19	80 \pm 10

^aEnzyme activity was measured by colorimetric determination of inorganic PO₄³⁻ levels in the presence of 4 mM magnesium chloride and 50 μ M lipid substrate (in parenthesis). IC₅₀ values are shown \pm SE of duplicate experiments. ^bThe IC₅₀ value for VO-OHpic was not determined because there was <50% inhibition at 10 μ M concentration. ^cNot determined.

A wider active site cleft arises from the position of the TI loop unique to the PTEN phosphatase (34). The active site dimensions described by Lee *et al.* (34) are consistent with those of the VO-OHpic compound even in the absence of any induced fit in the phosphatase. In contrast, other PTPs require selectivity for phosphotyrosine/phosphothreonine and therefore have much narrower catalytic clefts unable to accommodate the bulkier VO-OHpic (38).

This proves that the spatial differences in the catalytic pockets can be exploited to generate a potent inhibitor with a promising specificity, thus validating our approach of designing vanadates with space-intensive ligands as potential specific PTEN inhibitors. Since SopB is a bacterial enzyme, one can conclude that the designed inhibitor (VO-OHpic) is selectively targeting mammalian PTEN in the nanomolar concentration range.

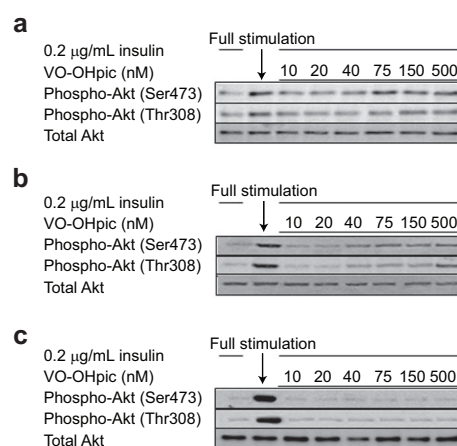


Figure 4. Dose-dependent inhibition of PTEN by VO-OHpic *in vivo*. Starved a) NIH 3T3 and b) L1 fibroblasts and c) UmUc-3 epithelial cells were prestimulated with 0.2 μ g mL⁻¹ insulin for 5 min, and indicated concentrations of VO-OHpic were applied for 15 min. For comparison, full stimulation of cells was carried out by addition of 10 μ g mL⁻¹ insulin. Cell lysates were analyzed by Western blotting for Akt phosphorylation on both Ser473 and Thr308 and equal loading (total Akt).

VO-OHpic Treatment Increases Akt Phosphorylation in a PTEN-Dependent Fashion *in Vivo*.

The data presented above confirm that VO-OHpic is a specific inhibitor for PTEN *in vitro*. To assess the applicability of this inhibitor in the cellular environment (NIH 3T3 and L1 fibroblasts, as well as the PTEN-negative UmUc-3 carcinoma cell line), we investigated the influence of VO-OHpic on the PTEN effector Akt. The activation of Akt, which can be monitored by detecting phosphorylation at two sites (Thr308 and Ser473), is strictly PtdIns(3,4,5)P3-dependent (39, 40). The Akt phosphorylation will thus be proportional to the changes of the cellular levels of this lipid, which is the major substrate of PTEN *in vivo* (32, 41). However, because PTEN is downstream of PI3K, it was necessary to prestimulate quiescent cells with low amounts of insulin ($0.2 \mu\text{g mL}^{-1}$), which would slightly raise the activity of PI3K above its basal levels and thus produce sufficient levels of the PTEN substrate, without stimulating Akt phosphorylation (Figure 4, lane 1). Consequently, when VO-OHpic was applied on starved cells without insulin, Akt phosphorylation is not increased (data not shown). However, under PI3K primed conditions most of the generated 3-phosphorylated lipids are quickly metabolized by the PI phosphatases such as PTEN, thus producing an environment where the phosphatase activity is dominant. The PTEN inhibitor increased the phosphorylation of Akt on both sites (Figure 4, panels a and b). This effect is detectable at 40 nM inhibitor concentration and reaches saturation at 75 nM VO-OHpic with a 2-fold increased phosphorylation of Akt as compared with the control (Supplementary Table 1). Not surprisingly, this effect is absent in the PTEN-negative UmUc-3 cells (Figure 4, panel c), providing the proof that this PTEN inhibitor is also very specific *in vivo* (5). The similarity of the dose responses on the tested fibroblasts to the IC50 values obtained with recombinant PTEN implies that VO-OHpic possesses cell permeability and potency. As expected, VO-OHpic had no effect on insulin-stimulated tyrosine phosphorylation at concentrations up to $10 \mu\text{M}$, which is well below its IC50 for the tyrosine phosphatase PTP- β (data not shown). It was not possible to test concentrations above the IC50 for the PTP- β due to cytotoxic effects at concentrations at or higher than $100 \mu\text{M}$ (data not shown). Taking all these observations together, one can conclude that this compound is a specific inhibitor of PTEN *in vivo*.

VO-OHpic Causes the Translocation of Phospho-Akt to the Plasma Membrane.

Inhibiting PTEN in cells should not only increase Akt phosphorylation, which is a prerequisite for activating downstream targets of Akt, but also induce translocation of this protein kinase to the plasma membrane and subsequently the perinuclear/nuclear region (42). Therefore we employed immunofluorescence microscopy to investigate the spatial distribution of phospho-Akt in response to VO-OHpic. Applying the PTEN inhibitor resulted in an Akt translocation and Ser473 phosphorylation that is comparable to insulin-stimulated cells (Figure 5, panel a). Low amounts of insulin do not cause Akt phosphorylation or translocation, whereas higher amounts of insulin or inhibition of PTEN by VO-OHpic increases the Ser473 phosphorylation of Akt localized on the plasma membrane and the nucleus, which is in agreement with the Western blotting experiments shown above (see Figure 4). The VO-OHpic-induced Akt phosphorylation is PI3K-dependent, since it is abolished by pretreatment with the PI3K inhibitor wortmannin (Figure 5, panel a). These data confirm that the PTEN inhibitor is activating Akt in a similar fashion to growth factors as judged by the ability to induce phosphorylation and translocation to membrane compartments. About 90% of the cells on the coverslip reacted in the same way (data not shown), indicating that the potential of the inhibitor in cell lines is independent of cell cycle, because the cells have not been synchronized. In addition, the cells looked healthy and showed no obvious morphological changes as compared with untreated cells.

VO-OHpic Treatment Increases PtdIns(3,4,5)P3 Levels in Fibroblasts.

The inhibition of PTEN results in the translocation and activation of Akt as shown above, which indicates that cellular levels of 3-phosphorylated inositol lipids must be elevated. The activation of Akt is known to rely on the increase of cellular PtdIns(3,4,5)P3 or PtdIns(3,4)P2 levels or both (43). These lipids can be readily identified by Pleckstrin homology (PH) domains (44), such as the one present in Akt, which recognizes PtdIns(3,4,5)P3 and PtdIns(3,4)P2 (45). While it is possible to image the relative amount and localization of these lipids in the living cell by overexpression of a fluorescently labeled Akt PH domain (46, 47), we preferred to employ this lipid recognition domain in a similar fashion to an antibody in fixed fibroblasts. This postfixation procedure avoids interference by the recognition domain on PtdIns(3,4,5)P3-dependent signaling path-

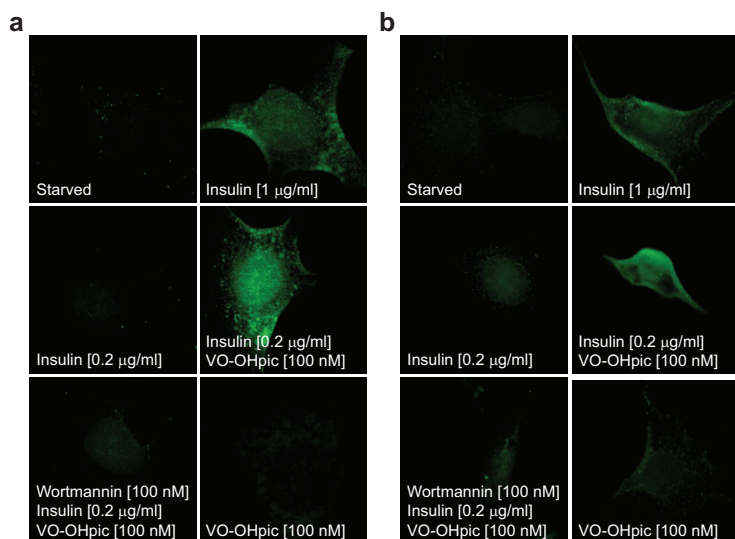


Figure 5. PTEN inhibition leads to increased PtdIns(3,4,5)P3 levels and Akt translocation. **a)** Immunofluorescence of Ser473 phospho-Akt. Starved NIH 3T3 fibroblasts were treated as indicated. Fixed cells were immunostained with a Ser473-phospho-specific antibody labeled with an Alexa Fluor 488 dye (green). Fluorescence was imaged in the FITC channel using a Nikon TE2000 inverted fluorescence microscope showing that VO-OHpic induced the phosphorylation of Akt (green) under PI3K primed conditions using a low concentration of insulin ($0.2 \mu\text{g mL}^{-1}$ insulin), which was not sufficient to activate Akt on its own in contrast to the positive control ($1 \mu\text{g mL}^{-1}$ insulin). Wortmannin treatment abolished any Akt phosphorylation, and VO-OHpic alone had no effect on Akt phosphorylation. Results are representative of at least three independent experiments carried out in duplicates. **b)** Detection of PtdIns(3,4,5)P3 and PtdIns(3,4)P2 levels postfixation. Starved NIH 3T3 cells were treated as indicated, fixed, and probed with a recombinant PH[Akt] domain labeled with Alexa Fluor 488 dye (green). Treatment with 100 nM VO-OHpic inhibits PTEN, as indicated by the increase of PtdIns(3,4,5)P3 and PtdIns(3,4)P2 levels.

ways, which includes Akt activation and could cause compensation by other mechanisms (48, 49). Employing this procedure, we found that PtdIns(3,4,5)P3 and PtdIns(3,4)P2 levels on the plasma membrane in VO-OHpic-treated cells are comparable to those in fully insulin-stimulated cells. Inhibition with wortmannin or prestimulation with low amounts of insulin results in residual or negligible staining (Figure 5, panel b), which demonstrates that VO-OHpic is changing cellular inositol lipid levels in a fashion compatible with PTEN inhibition. The uniform distribution of lipid staining on the plasma membrane indicates that the inhibitor does not localize to certain compartments in cells but targets PTEN evenly throughout the cell, resulting in an increase in the cellular levels of its substrates, the products of

PI3K. Taken together, these results prove that VO-OHpic targets PTEN's lipid phosphatase activity *in vivo*. They also demonstrate the advantage of this compound for the exploration of specific PtdIns(3,4,5)P3 signaling pathways and their downstream effects without altering PI3K activity or expression, thus complementing the already well-employed PI3K inhibitors wortmannin and LY294002.

VO-OHpic Treatment Reduces Akt-Dependent Transcriptional Activity of FoxO3a

Having established that VO-OHpic inhibits PTEN activity and Akt function *in vivo*, we investigated whether increased phosphorylation of Akt upon PTEN inhibition results in a corresponding change in Akt effectors. FoxO3a is known to translocate from the nucleus to the cytoplasm upon phosphorylation by Akt (50) and was therefore employed as a readout for VO-OHpic's ability to propagate signaling downstream of Akt. Immunofluores-

cence of FoxO3a in NIH 3T3 cells (Figure 6, panel a) reveals that FoxO3a accumulates in the cytosol after PTEN inhibitor treatment or insulin stimulation ($1 \mu\text{g mL}^{-1}$), whereas in quiescent cells, FoxO3a is localized in the nucleus ($0.2 \mu\text{g mL}^{-1}$ insulin). We corroborated these findings by investigating the *bim* promoter activity, which is known to be under the control of the FoxO3a transcriptional activator (51). Due to translocation from the nucleus to the cytosol upon phosphorylation of Akt, FoxO3a becomes inactive as a transcription factor, which should therefore result in less *bim* promoter activity (51). As expected, VO-OHpic treatment reduces *bim* promoter activity (Figure 6, panel b, left) up to 6-fold, whereas a *bim* promoter with a mutated consensus sequence for FoxO3a (Figure 6, panel b, right) has no

Employing this small-molecule chemical inhibitor could be advantageous because the enzyme activity can be reduced in a dose-dependent manner without affecting its scaffolding function and ability to interact with other proteins such as MAGI-2 (58), which complements the results obtained from genetic knock-out studies.

When all these observations are taken together, it is evident that VO-OHPic is a potent small-molecule compound that specifically inhibits PTEN's cellular enzymatic activity, which in turn activates downstream targets such as Akt and FoxO3a. Importantly, the data provide evidence that glucose uptake into adipocytes is dramatically enhanced upon PTEN inhibition with VO-OHPic (Figure 7), suggesting that endogenous PTEN has an important regulatory role in glucose uptake and thus would be an interesting drug target. However, it remains to be seen whether the specific PTEN inhibitor can overcome insulin insensitivity and protect from developing type II diabetes, in particular, since it is not known whether PTEN or SH2-containing 5'-inositol phosphatase (SHIP2) controls glucose transporter translocation (16, 59, 60). The here presented PTEN inhibitor, which does not significantly reduce SHIP2 phosphatase activity *in vitro* (IC₅₀ > 10 μM; data not shown), could

be a useful tool to unravel the contributions of these two diabetes drug targets (18).

Inhibiting PTEN will also cause changes in the migratory and invasive properties of cells (4, 61, 62), a feature that has been linked to its well-documented tumor suppressor function. Our own tests reveal that the PTEN inhibitor accelerates wound healing in fibroblasts (data not shown), indicating that it has the potential to alter these PTEN functions as well. While these observations would cause some caution toward the applicability of PTEN inhibitors in general, it does not rule out a tissue-specific delivery system targeting PTEN. This latter notion is supported by recent reports demonstrating that PTEN loss in pancreatic β cells is not tumorigenic (63) and that selective PTEN deletion in skeletal muscle protects against the development of insulin resistance without the development of cancer (11). A targeted tissue-specific drug could therefore be considered for the treatment of diabetes without causing malignant cell growth. The here presented chemical PTEN-specific inhibitor will be an important and probably decisive research tool in unraveling these issues, since it will facilitate investigation of PTEN's role in primary tissues as well as immortalized cancer cell lines.

METHODS

Chemistry Experimental Details. General Procedures and

Materials. IR spectra were recorded on a Perkin Elmer FT-IR1720 Research Series spectrometer using KBr disks in the range 4000–500 cm⁻¹. Fast atom bombardment (FAB(+)) mass spectra were recorded by J. Barton at Imperial College London on a VG AutoSpec-Q using 3-nitrobenzyl alcohol as matrix. 1,10-Phenanthroline, picolinic acid, 3-hydroxypicolinic acid, 1-isoquinoline, 1-phenylbiguanide, and V₂O₅ (99.99%) were purchased from Sigma Aldrich Chemical Co. and used as received unless otherwise stated. The following complexes were prepared according to previously reported literature procedures (33, 36, 64, 65): V(=O)(pic)₂, K₂[V(=O)(O₂)₂(pic)], V(=O)(H₂O)(OHPic)₂, K₂[V(=O)(O₂)₂(OHPic)], K[V(=O)(O₂)₂(phen)] and K₂[V(=O)(O₂)₂(isoq)] (where pic = picolinic acid, OHPic = 3-hydroxypicolinic acid, phen = 1,10-phenanthroline, and isoqu = 1-isoquinolinecarboxylic acid).

LY294002 was purchased from Calbiochem. All other chemicals were purchased from Sigma unless otherwise stated in the respective section.

Synthesis of V(=O)(phenylbiguanide)₂ (VO-biguan). This reaction was carried out under an atmosphere of argon. A solution of VOSO₄·3H₂O (1.00 g, 4.60 mmol) in distilled and degassed H₂O (5 mL) was mixed with a solution containing 2 equiv of 1-phenylbiguanide (1.63 g, 9.20 mmol) in 5 mL of distilled and degassed H₂O. This was followed by adjustment of the pH to 12 by slow addition of 11 mL of a 2 M solution of NaOH (0.90 g, 23.0 mmol). After complete addition of NaOH, the solution was stirred overnight resulting in precipitation of a light purple solid. The precipi-

tate was filtered through a glass microfiber filter paper, washed with cold water and diethyl ether, and dried overnight under reduced pressure (yield 1.127 g, 58.1%). Elemental analyses calculated for VOC₁₆H₂₀N₁₀ (MW 419 g mol⁻¹): C, 45.83; H, 4.81; N, 33.40. Found: C, 45.74; H, 4.57; N, 33.06. FAB(+) MS *m/z*: 419 amu (M⁺). IR (KBr) in cm⁻¹: ν(NH₂) 3175, 3300, 3056; ν(C = N) 1626; ν(V = O) 958.

Synthesis of K[V(=O)(O₂)₂(phenylbiguanide)] (bpV-biguan).

V₂O₅ (1.43 g, 7.84 mmol) was dissolved in a solution of KOH (1.04 g, 18.53 mmol) in water (15 mL) and stirred for 5 min. To the resulting green solution, 2 mL of a 30% (w/v) aqueous solution of H₂O₂ was added with production of effervescence. The mixture was stirred for 25 min with a progressive change in color to orange and with complete dissolution of all solids. A further 12 mL of a 30% (w/v) aqueous solution of H₂O₂ was added, and the reaction mixture was stirred for 15 min. At this point, a solution of 1-phenylbiguanide (2.78 g, 15.68 mmol) in a mixture of water (20 mL) and ethanol (5 mL) was added to the reaction mixture, which was stirred at RT for 30 min, after which consistent cloudiness was noted. The resulting pale yellow precipitate was filtered through glass microfiber filter paper and dried under reduced pressure (yield 6.33 g, 91%). IR (KBr) in cm⁻¹: ν(NH₂) 3180; ν(CN) 1641; ν(VO) 942; ν(OO) 868; ν(VO) 616 and 526 cm⁻¹. FAB(+) MS *m/z*: 348 amu ([M + H]⁺). Elemental analyses calculated for C₈H₁₁N₅KVO₅ (MW = 347 g mol⁻¹): C, 27.67; H, 3.19; N, 20.17. Found: C, 27.53; H, 3.33; N, 19.90.

Plasmids and Protein Expression. The coding region of the respective DNA sequences (of human PTEN, MTM, or rat Sac (amino acids 1–1042), mouse Akt PH domain (amino acids 5–108, SwissProt P31750), or bacterial SopB) was cloned into a

pGEX-4T2 expression vector (Pharmacia). Protein expression was induced overnight in the *Escherichia coli* DH5 α strain using 100 μ M isopropyl- β -D-thiogalactopyranoside at 18 $^{\circ}$ C. Glutathione S-transferase (GST)-fusion protein was purified according to the manufacturer's manual using glutathione-Sepharose 4B (Pharmacia). Protein integrity was confirmed by Western blot using GST antibody (Novagen). PTP- β was purchased from Sigma.

Phosphate Release Assay. Phosphatase activities were determined using a phosphate release assay (31). For the detection of PTEN and Sac1 activities, bismuth was added to the phosphate release assay in order to improve its stability and sensitivity (66). All enzyme preparations were tested for linearity to ensure that suitable amounts of enzyme were employed in the inhibitor assays. Enzymes were incubated with inhibitors at various concentrations prior to starting the phosphatase reaction by adding the corresponding substrates presented in octylglucoside mixed micelles as described before (31).

Antibodies. The antibody used for immunostaining of phosphorylated Akt/PKB was a mouse monoclonal anti-phospho-Akt (Ser473) antibody from the IgG2b isotype (Cell Signaling 587F11) or anti-phospho-Akt (Thr308) (Cell Signaling 4G10). The antibody for immunodetection of total Akt/PKB was a rabbit anti-Akt antibody (Cell Signaling). Horseradish peroxidase (HRP)-conjugated goat anti-mouse or -rabbit antibody was purchased from BioRad. Mouse GST antibody was from Novagen. The anti-FoxO3a antibody is a rabbit polyclonal antiserum raised against a C-terminal-specific peptide of human FoxO3a.

Antibody Labeling. Antibodies were labeled with Alexa Fluor dyes (488 and 595) with a Zenon IgG Labeling Kit (Molecular Probes) according to manufacturer's instructions.

Tissue Culture. NIH 3T3 fibroblasts (mouse fibroblasts), 3T3-L1 (mouse fibroblasts or fibroblasts differentiated to adipocytes as described previously by Volchuk *et al.* (67), and UmUc-3 cells (human urinary bladder carcinoma) were purchased from ATCC. All cells were grown in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% (v/v) newborn calf serum (Invitrogen) in an atmosphere of 5% (v/v) CO $_2$ at 37 $^{\circ}$ C.

Immunofluorescence of Total Akt and Ser473 Phosphorylated Akt. If not otherwise indicated, cells were seeded on poly-L-lysine (Sigma) coated glass coverslips and starved overnight. Cells were preincubated with inhibitors and stimulated as indicated in the Results and Discussion section. Cells were fixed in 4% (w/v) paraformaldehyde (PFA) for 10 min, quenched with 50 mM NH $_4$ Cl/phosphate-buffered saline (PBS), permeabilized with 0.25% (v/v) Triton X-100 for 7 min and blocked with 1% (w/v) bovine serum albumin/PBS (fatty acid free; Sigma) for 1 h at RT. Cells were then incubated with the fluorescently labeled antibody (1:200 diluted in blocking solution) for 1 h followed by three extensive PBS washes and a second fixation step with 4% (w/v) PFA for 10 min. Where applicable, washed coverslips were incubated with a second antibody (1:200) for 1 h, washed extensively, and treated with 300 nM 4',6-diamidino-2-phenylindole (dapi) for nuclear counterstaining. Thoroughly washed coverslips were mounted on glass slides using Mowiol supplemented with 0.6% (w/v) 1,4-diazabicyclo-[2.2.2]octane.

Detection of PtdIns(3,4,5)P3. The dialyzed GST-tagged PH domain of Akt/PKB (amino acids 5–108, SwissProt P31750) was used for detection of PtdIns(3,4,5)P3 and PtdIns(3,4)P2 on fixed cells applying the same method as described above and in Byrne *et al.* (68) for immunofluorescence with slight modification. The recombinant peptide was diluted 1:1000 in blocking buffer and detected with an Alexa Fluor labeled anti-GST antibody (1:5000).

Cell preparations were observed under a Nikon TE 2000 fluorescence microscope using a 100 \times Fluor oil lens. Filters used in

the fluorescence experiments were bandpass for dapi, FITC, and tetramethyl rhodamine iso-thiocyanate (TRITC) with excitation wavelengths of 340–380, 465–495, and 540–580, respectively, and with emission wavelengths of 435–485, 515–555, and 572–605, respectively. Images were digitally acquired with a CCD camera (Hamamatsu) for each fluorophore separately and processed using IPLab software, v 3.65a. and ImageJ (National Institutes of Health, Bethesda, MD).

Western Blotting. Quiescent cells were treated as indicated and lysed in SDS sample buffer. Protein samples were separated by SDS-PAGE and blotted on polyvinylidene difluoride membrane. After blocking in 5% (w/v) milk, membranes were probed with the mouse monoclonal anti-phospho-Akt (Ser473) antibody (1:1000) or anti-phospho-Akt (Thr308) antibody and detected with an HRP-conjugated anti-mouse antibody. Stripped membranes were reprobed with a rabbit anti-Akt antibody and detected with an HRP-conjugated anti-rabbit antibody.

Transfection and Luciferase Assays. The wild-type and mutant *bim* promoter/reporter constructs have previously been described (51). NIH 3T3 cells were transfected using the calcium phosphate coprecipitation method as described previously (51). Briefly, calcium phosphate precipitates containing 1 μ g of wild-type *bim* promoter firefly-luciferase reporter plasmid (pGL2-hBim) or mutant plasmid (pGL2-mutant-hBim), together with 0.2 μ g of a *Renilla* luciferase transfection control (pRL-TK; Promega), were incubated overnight with subconfluent cell cultures in each well of a 24-well plate. The cells were then washed twice in PBS, treated, and harvested for firefly/*Renilla* luciferase assays using the Dual-Luciferase Reporter Assay System (Promega).

Glucose Uptake. Differentiated mouse adipocytes growing on coverslips were serum- and glucose-starved for 4 h. After preincubation with VO-OHpic and insulin, cells were treated with 500 μ M 2-NBDG fluorescent glucose analogue (Molecular Probes) for 10 min at 37 $^{\circ}$ C. A negative control preincubated with 10 mM D-glucose was included. Coverslips were washed extensively in PBS and counterstained with dapi. Cell preparations were viewed under a Nikon TE 2000 fluorescence microscope using a band pass filter for FITC in order to analyze the extent of glucose uptake.

Accession Codes: PTEN, AAD13528; MTM1, Q13496; PTP-1B, NP_002818; PTP- β , AAB26530; SopB, AAS76429; Sac, O15056.

Acknowledgments: This work was partially supported by EPSRC, MRC, and Cancer Research U.K. Thanks to D. Briggs for assistance in preparing Figure 1, panel b, to M. Serrano (ICIQ) for his help in preparing some of the previously reported vanadium complexes, to E. Galyov (Institute for Animal Health, U.K.) for providing the SopB plasmid, and to M. Clague (University of Liverpool, U.K.) for providing the MTM1 protein.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Hoffman, B. T., Nelson, M. R., Burdick, K., and Baxter, S. M. (2004) Protein tyrosine phosphatases: strategies for distinguishing proteins in a family containing multiple drug targets and anti-targets, *Curr. Pharm. Des.* 10, 1161–1181.
- Goberdhan, D. C. I., and Wilson, C. (2003) PTEN: tumour suppressor, multifunctional growth regulator and more, *Hum. Mol. Genet.* 12, R239–R248.
- Maehama, T., and Dixon, J. E. (1998) The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate, *J. Biol. Chem.* 273, 13375–13378.

4. Cai, X. M., Tao, B. B., Wang, L. Y., Liang, Y. L., Jin, J. W., Yang, Y., Hu, Y. L., and Zha, X. L. (2005) Protein phosphatase activity of PTEN inhibited the invasion of glioma cells with epidermal growth factor receptor mutation type III expression, *Int. J. Cancer* **117**, 905–912.
5. Hamilton, J. A., Stewart, L. M., Ajayi, L., Gray, I. C., Gray, N. E., Roberts, K. G., Watson, G. J., Kaisary, A. V., and Snary, D. (2000) The expression profile for the tumour suppressor gene PTEN and associated polymorphic markers, *Br. J. Cancer* **82**, 1671–1676.
6. Li, D. M., and Sun, H. (1998) PTEN/MMAC1/TEP1 suppresses the tumorigenicity and induces G1 cell cycle arrest in human glioblastoma cells, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15406–15411.
7. Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S. I., Puc, J., Miliareis, C., Rodgers, L., McCombie, R., Bigner, S. H., Giovanella, B. C., Iltmann, M., Tycko, B., Hibshoosh, H., Wigler, M. H., and Parsons, R. (1997) PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer, *Science* **275**, 1943–1947.
8. Lazar, D. F., and Saltiel, A. R. (2006) Lipid phosphatases as drug discovery targets for type 2 diabetes, *Nat. Rev. Drug Discovery* **5**, 333–342.
9. Hafen, E. (2004) Cancer, type 2 diabetes, and ageing: news from flies and worms, *Swiss Med. Wkly.* **134**, 711–719.
10. Ishihara, H., Sasaoka, T., Kagawa, S., Murakami, S., Fukui, K., Kawagishi, Y., Yamazaki, K., Sato, A., Iwata, M., Urakaze, M., Ishiki, M., Wada, T., Yaguchi, S., Tsuneki, H., Kimura, I., and Kobayashi, M. (2003) Association of the polymorphisms in the 5'-untranslated region of PTEN gene with type 2 diabetes in a Japanese population, *FEBS Lett.* **554**, 450–454.
11. Wijesekara, N., Konrad, D., Eweida, M., Jefferies, C., Liadis, N., Giacca, A., Crackower, M., Suzuki, A., Mak, T. W., Kahn, C. R., Klip, A., and Woo, M. (2005) Muscle-specific Pten deletion protects against insulin resistance and diabetes, *Mol. Cell. Biol.* **25**, 1135–1145.
12. Kurlawalla-Martinez, C., Stiles, B., Wang, Y., Devaskar, S. U., Kahn, B. B., and Wu, H. (2005) Insulin hypersensitivity and resistance to streptozotocin-induced diabetes in mice lacking PTEN in adipose tissue, *Mol. Cell. Biol.* **25**, 2498–2510.
13. Ning, K., Miller, L. C., Laidlaw, H. A., Burgess, L. A., Perera, N. M., Downes, C. P., Leslie, N. R., and Ashford, M. L. (2006) A novel leptin signalling pathway via PTEN inhibition in hypothalamic cell lines and pancreatic beta-cells, *EMBO J.* **25**, 2377–2387.
14. Nakashima, N., Sharma, P. M., Imamura, T., Bookstein, R., and Olefsky, J. M. (2000) The tumor suppressor PTEN negatively regulates insulin signaling in 3T3-L1 adipocytes, *J. Biol. Chem.* **275**, 12889–12895.
15. Thong, F. S. L., Dugani, C. B., and Klip, A. (2005) Turning signals on and off: GLUT4 Traffic in the insulin-signaling highway, *Physiology* **20**, 271–284.
16. Tang, X., Powelka, A. M., Soriano, N. A., Czech, M. P., and Guilherme, A. (2005) PTEN, but not SHIP2, suppresses insulin signaling through the phosphatidylinositol 3-kinase/Akt pathway in 3T3-L1 adipocytes, *J. Biol. Chem.* **280**, 22523–22529.
17. Mosser, V. A., Li, Y., and Quon, M. J. (2001) PTEN does not modulate GLUT4 translocation in rat adipose cells under physiological conditions, *Biochem. Biophys. Res. Commun.* **288**, 1011–1017.
18. Sasaoka, T., Wada, T., and Tsuneki, H. (2006) Lipid phosphatases as a possible therapeutic target in cases of type 2 diabetes and obesity, *Pharmacol. Ther.* **112**, 799–809.
19. Srivastava, A. K., and Mehdi, M. Z. (2005) Insulino-mimetic and anti-diabetic effects of vanadium compounds, *Diabetic Med.* **22**, 2–13.
20. Cam, M. C., Rodrigues, B., and McNeill, J. H. (1999) Distinct glucose lowering and beta cell protective effects of vanadium and food restriction in streptozotocin-diabetes, *Eur. J. Endocrinol.* **141**, 546–554.
21. Heyliger, C. E., Tahiliani, A. G., and McNeill, J. H. (1985) Effect of vanadate on elevated blood glucose and depressed cardiac performance of diabetic rats, *Science* **227**, 1474–1477.
22. Green, A. (1986) The insulin-like effect of sodium vanadate on adipocyte glucose transport is mediated at a post-insulin-receptor level, *Biochem. J.* **238**, 663–669.
23. Shechter, Y., Li, J., Meyerovitch, J., Gefel, D., Bruck, R., Elberg, G., Miller, D. S., and Shisheva, A. (1995) Insulin-like actions of vanadate are mediated in an insulin-receptor-independent manner via non-receptor protein tyrosine kinases and protein phosphotyrosine phosphatases, *Mol. Cell. Biochem.* **153**, 39–47.
24. Shisheva, A., and Shechter, Y. (1993) Mechanism of pervanadate stimulation and potentiation of insulin-activated glucose transport in rat adipocytes: dissociation from vanadate effect, *Endocrinology* **133**, 1562–1568.
25. Goldwaser, I., Gefel, D., Gershonov, E., Fridkin, M., and Shechter, Y. (2000) Insulin-like effects of vanadium: basic and clinical implications, *J. Inorg. Biochem.* **80**, 21–25.
26. Tolman, E. L., Barris, E., Burns, M., Pansini, A., and Partridge, R. (1979) Effects of vanadium on glucose metabolism *in vitro*, *Life Sci.* **25**, 1159–1164.
27. Gordon, J. A. (1991) Use of vanadate as protein-phosphotyrosine phosphatase inhibitor, *Methods Enzymol.* **201**, 477–482.
28. Huyer, G., Liu, S., Kelly, J., Moffat, J., Payette, P., Kennedy, B., Tsaprailis, G., Gresser, M. J., and Ramachandran, C. (1997) Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate, *J. Biol. Chem.* **272**, 843–851.
29. Bhattacharyya, S., and Tracey, A. S. (2001) Vanadium(V) complexes in enzyme systems: aqueous chemistry, inhibition and molecular modeling in inhibitor design, *J. Inorg. Biochem.* **85**, 9–13.
30. Cuncic, C., Desmarais, S., Detich, N., Tracey, A. S., Gresser, M. J., and Ramachandran, C. (1999) Bis(N,N-dimethylhydroxamido)hydroxovanadate inhibition of protein tyrosine phosphatase activity in intact cells: comparison with vanadate, *Biochem. Pharmacol.* **58**, 1859–1867.
31. Schmid, A. C., Byrne, R. D., Vilar, R., and Woscholski, R. (2004) Bis-peroxovanadium compounds are potent PTEN inhibitors, *FEBS Lett.* **566**, 35–38.
32. Maehama, T., and Dixon, J. E. (1999) PTEN: a tumour suppressor that functions as a phospholipid phosphatase, *Trends Cell Biol.* **9**, 125–128.
33. Posner, B. I., Faure, R., Burgess, J. W., Bevan, A. P., Lachance, D., Zhang-Sun, G., Fantus, I. G., Ng, J. B., Hall, D. A., and Lum, B. S. (1994) Peroxovanadium compounds. A new class of potent phosphotyrosine phosphatase inhibitors which are insulin mimetics, *J. Biol. Chem.* **269**, 4596–4604.
34. Lee, J. O., Yang, H., Georgescu, M. M., Di Cristofano, A., Maehama, T., Shi, Y., Dixon, J. E., Pandolfi, P., and Pavletich, N. P. (1999) Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association, *Cell* **99**, 323–334.
35. Begley, M. J., Taylor, G. S., Kim, S. A., Veine, D. M., Dixon, J. E., and Stuckey, J. A. (2003) Crystal structure of a phosphoinositide phosphatase, MTMR2: insights into myotubular myopathy and Charcot-Marie-Tooth syndrome, *Mol. Cell* **12**, 1391–1402.
36. Nakai, M., Sekiguchi, F., Obata, M., Ohtsuki, C., Adachi, Y., Sakurai, H., Orvig, C., Rehder, D., and Yano, S. (2005) Synthesis and insulin-mimetic activities of metal complexes with 3-hydroxypyridine-2-carboxylic acid, *J. Inorg. Biochem.* **99**, 1275–1282.
37. Clague, M. J., and Lorenzo, O. (2005) The myotubularin family of lipid phosphatases, *Traffic* **6**, 1063–1069.
38. De Lano, W. L. (2006) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
39. Alessi, D. R., James, S. R., Downes, C. P., Holmes, A. B., Gaffney, P. R., Reese, C. B., and Cohen, P. (1997) Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Balph α , *Curr. Biol.* **7**, 261–269.

40. Currie, R. A., Walker, K. S., Gray, A., Deak, M., Casamayor, A., Downes, C. P., Cohen, P., Alessi, D. R., and Lucocq, J. (1999) Role of phosphatidylinositol 3,4,5-trisphosphate in regulating the activity and localization of 3-phosphoinositide-dependent protein kinase-1. *Biochem. J.* **337**, 575–583.
41. Ramaswamy, S., Nakamura, N., Vazquez, F., Batt, D. B., Perera, S., Roberts, T. M., and Sellers, W. R. (1999) Regulation of G1 progression by the PTEN tumor suppressor protein is linked to inhibition of the phosphatidylinositol 3-kinase/Akt pathway. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2110–2115.
42. Meier, R., Alessi, D. R., Cron, P., Andjelkovic, M., and Hemmings, B. A. (1997) Mitogenic activation, phosphorylation, and nuclear translocation of protein kinase Bbeta. *J. Biol. Chem.* **272**, 30491–30497.
43. Franke, T. F., Kaplan, D. R., Cantley, L. C., and Toker, A. (1997) Direct regulation of the Akt proto-oncogene product by phosphatidylinositol-3,4-bisphosphate. *Science* **275**, 665–668.
44. Lemmon, M. A. (2003) Phosphoinositide recognition domains. *Traffic* **4**, 201–213.
45. Frech, M., Andjelkovic, M., Ingley, E., Reddy, K. K., Falck, J. R., and Hemmings, B. A. (1997) High affinity binding of inositol phosphates and phosphoinositides to the pleckstrin homology domain of RAC/protein kinase B and their influence on kinase activity. *J. Biol. Chem.* **272**, 8474–8481.
46. Gray, A., Van der Kaye, J., and Downes, C. P. (1999) The pleckstrin homology domains of protein kinase B and GRP1 (general receptor for phosphoinositides-1) are sensitive and selective probes for the cellular detection of phosphatidylinositol 3,4-bisphosphate and/or phosphatidylinositol 3,4,5-trisphosphate *in vivo*. *Biochem. J.* **344**, 929–936.
47. Vamai, P., and Balla, T. (2006) Live cell imaging of phosphoinositide dynamics with fluorescent protein domains. *Biochim. Biophys. Acta.* **1761**, 957–967.
48. Vamai, P., Bondeva, T., Tamas, P., Toth, B., Buday, L., Hunyady, L., and Balla, T. (2005) Selective cellular effects of overexpressed pleckstrin-homology domains that recognize PtdIns(3,4,5)P₃ suggest their interaction with protein binding partners. *J. Cell. Sci.* **118**, 4879–4888.
49. Irvine, R. (2004) Inositol lipids: to PHix or not to PHix? *Curr. Biol.* **14**, R308–R310.
50. Burgering, B. M. T., and Medema, R. H. (2003) Decisions on life and death: FOXO Forkhead transcription factors are in command when PKB/Akt is off duty. *J. Leukocyte Biol.* **73**, 689–701.
51. Sunters, A., Stahl, M., Brosens, J. J., Zoumpoulidou, G., Saunders, C. A., Coffey, P. J., Medema, R. H., Coombes, R. C., and Lam, E. W. (2003) FoxO3a transcriptional regulation of Bim controls apoptosis in paclitaxel-treated breast cancer cell lines. *J. Biol. Chem.* **278**, 49795–49805.
52. Sunters, A., Madureira, P. A., Pomeranz, K. M., Aubert, M., Brosens, J. J., Cook, S. J., Burgering, B. M. T., Coombes, R. C., and Lam, E. W. F. (2006) Paclitaxel-induced nuclear translocation of FOXO3a in breast cancer cells is mediated by c-Jun NH2-terminal kinase and Akt. *Cancer Res.* **66**, 212–220.
53. Lam, E. W., and Watson, R. J. (1993) An E2F-binding site mediates cell-cycle regulated repression of mouse B-myb transcription. *EMBO J.* **12**, 2705–2713.
54. Fernandez de, M. S., Essafi, A., Soeiro, I., Pietersen, A. M., Birkenkamp, K. U., Edwards, C. S., Martino, A., Nelson, B. H., Francis, J. M., Jones, M. C., Brosens, J. J., Coffey, P. J., and Lam, E. W. (2004) FoxO3a and BCR-ABL regulate cyclin D2 transcription through a STAT5/BCL6-dependent mechanism. *Mol. Cell. Biol.* **24**, 10058–10071.
55. Ono, H., Katagiri, H., Funaki, M., Anai, M., Inukai, K., Fukushima, Y., Sakoda, H., Oghara, T., Onishi, Y., Fujishiro, M., Kikuchi, M., Oka, Y., and Asano, T. (2001) Regulation of phosphoinositide metabolism, Akt phosphorylation, and glucose transport by PTEN (phosphatase and tensin homolog deleted on chromosome 10) in 3T3-L1 Adipocytes. *Mol. Endocrinol.* **15**, 1411–1422.
56. Yoshizaki, T., Maegawa, H., Egawa, K., Ugi, S., Nishio, Y., Imamura, T., Kobayashi, T., Tamura, S., Olefsky, J. M., and Kashiwagi, A. (2004) Protein phosphatase-2C alpha as a positive regulator of insulin sensitivity through direct activation of phosphatidylinositol 3-kinase in 3T3-L1 adipocytes. *J. Biol. Chem.* **279**, 22715–22726.
57. Mahadev, K., Wu, X., Motoshima, H., and Goldstein, B. J. (2004) Integration of multiple downstream signals determines the net effect of insulin on MAP kinase vs. PI 3'-kinase activation: potential role of insulin-stimulated H(2)O(2). *Cell. Signalling* **16**, 323–331.
58. Wu, X., Hepner, K., Castellino-Prabhu, S., Do, D., Kaye, M. B., Yuan, X. J., Wood, J., Ross, C., Sawyers, C. L., and Whang, Y. E. (2000) Evidence for regulation of the PTEN tumor suppressor by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4233–4238.
59. Blero, D., Zhang, J., Pesesse, X., Payrastre, B., Dumont, J. E., Schurmans, S., and Erneux, C. (2005) Phosphatidylinositol 3,4,5-trisphosphate modulation in SHIP2-deficient mouse embryonic fibroblasts. *FEBS J.* **272**, 2512–2522.
60. Vandeput, F., Backers, K., Villeret, V., Pesesse, X., and Erneux, C. (2006) The influence of anionic lipids on SHIP2 phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase activity. *Cell. Signalling* **18**, 2193–2199.
61. Leslie, N. R., Yang, X., Downes, C. P., and Weijer, C. J. (2005) The regulation of cell migration by PTEN. *Biochem. Soc. Trans.* **33**, 1507–1508.
62. Raftopoulou, M., Etienne-Manneville, S., Self, A., Nicholls, S., and Hall, A. (2004) Regulation of cell migration by the C2 domain of the tumor suppressor PTEN. *Science* **303**, 1179–1181.
63. Stiles, B. L., Kuralwalla-Martinez, C., Guo, W., Gregorian, C., Wang, Y., Tian, J., Magnuson, M. A., and Wu, H. (2006) Selective deletion of PTEN in pancreatic {beta} cells leads to increased islet mass and resistance to STZ-induced diabetes. *Mol. Cell. Biol.* **26**, 2772–2781.
64. Melchior, M., Thompson, K. H., Jong, J. M., Rettig, S. J., Shuter, E., Yuen, V. G., Zhou, Y., McNeill, J. H., and Orvig, C. (1999) Vanadium complexes as insulin mimetic agents: coordination chemistry and *in vivo* studies of oxovanadium(IV) and dioxovanadate(V) complexes formed from naturally occurring chelating oxazolinates, thiazolinates, or picolinates units. *Inorg. Chem.* **38**, 2288–2293.
65. Quilitzsch, U., and Wieghardt, K. (1979) Kinetics of the diperoxovanadate(V)-monoperoxovanadate(V) conversion in perchloric-acid media. *Inorg. Chem.* **18**, 869–871.
66. Cariani, L., Thomas, L., Brito, J., and del Castillo, J. R. (2004) Bis-muth citrate in the quantification of inorganic phosphate and its utility in the determination of membrane-bound phosphatases. *Anal. Biochem.* **324**, 79–83.
67. Volchuk, A., Wang, Q., Ewart, H. S., Liu, Z., He, L., Bennett, M. K., and Klip, A. (1996) Syntaxin 4 in 3T3-L1 adipocytes: regulation by insulin and participation in insulin-dependent glucose transport. *Mol. Biol. Cell* **7**, 1075–1082.
68. Byrne, R. D., Rosivatz, E., Parsons, M., Larjani, B., Parker, P. J., Ng, T., and Woscholski, R. (2006) Differential activation of the PI 3-kinase effectors AKT/PKB and p70 S6 kinase by compound 48/80 is mediated by PKCalpha. *Cell. Signalling*, in press.