J|A|C|S
PERSPECTIVES

Published on Web 08/18/2009

Synthesis at the Interface of Chemistry and Biology

Xu Wu[‡] and Peter G. Schultz*[,†,‡]

The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, and Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, California 92121

Received April 1, 2009; E-mail: schultz@scripps.edu

Abstract: As the focus of synthesis increasingly shifts from its historical emphasis on molecular structure to function, improved strategies are clearly required for the generation of molecules with defined physical, chemical, and biological properties. In contrast, living organisms are remarkably adept at producing molecules and molecular assemblies with an impressive array of functions — from enzymes and antibodies to the photosynthetic center. Thus, the marriage of Nature's synthetic strategies, molecules, and biosynthetic machinery with more traditional synthetic approaches might enable the generation of molecules with properties difficult to achieve by chemical strategies alone. Here we illustrate the potential of this approach and overview some opportunities and challenges in the coming years.

Introduction

The feature that perhaps most distinguishes chemistry from the rest of the sciences is the ability of chemists to control the structure of matter at the molecular level — from complex natural products like vancomycin to nanoparticles and whole genomes. Indeed there have been remarkable advances in the fields of total synthesis and synthetic methods over the past 50 years. Unfortunately, we are not nearly as adept at the synthesis of molecules with defined functions as we are at the synthesis of molecules with defined structures. As the focus of chemistry increasingly shifts from structure to function, chemists will need to develop better strategies to efficiently generate molecules, and systems of molecules, with desired physical, chemical, or biological properties in order to meet the biomedical, energy, and environmental needs of the future. Indeed this challenge represents one of the great opportunities for synthesis in the coming years. One direction we can turn for help is Mother Nature — after all, living organisms carry out a remarkable array of complex functions using natural molecules and molecular assemblies, ranging from antibiotics and enzymes to the ribosome and photosynthetic center.

Organic chemists have spent considerable effort synthesizing molecules that attempt to mimic the functions found in Nature. Early examples include functionalized synthetic hosts,[1,2] iron−sulfur clusters,[3] and heme analogues.[4] These efforts attempted to replicate key functions of natural enzymes and receptors and, thereby, give new insight into their molecular mechanisms. As chemists became more sophisticated in their understanding of biomolecules and biological methods, there was an increasing shift in focus to the synthesis of biomolecular mimetics that directly modulate the activities of biological systems themselves. A pioneering example was the synthesis by Dervan and co-workers of polypyrrole-carboxamides that bind DNA in a sequence-specific manner much like transcriptional repressors (Figure 1).[5−7]

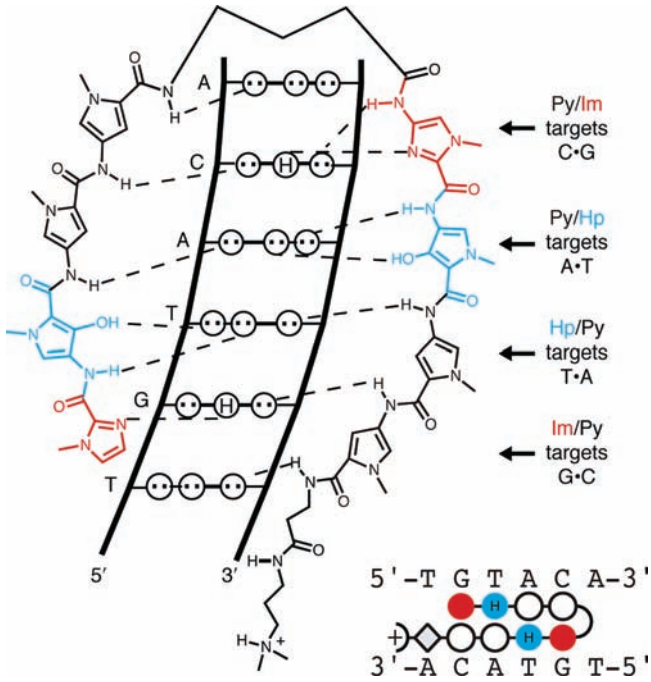


Figure 1. Synthetic molecules that sequence-specifically bind duplex DNA much like transcriptional repressors.[6] Image courtesy of Peter Dervan.

But one need not be limited to chemical synthesis alone to generate molecules with novel functions. One can exploit Nature itself, i.e., use the synthetic strategies, molecules, and biosynthetic machinery of living organisms together with more traditional chemical approaches to generate molecules with properties that might be difficult to realize by either strategy alone. Such an approach represents a marriage of traditional chemical synthesis with the emerging field of synthetic biology.[8] Early efforts in this direction included the generation of semisynthetic enzymes[9] and ion channels[10] by Kaiser and Erlanger, respectively, and the work of Orgel on DNA-directed chemical synthesis.[11] Today this approach is beginning to impact many areas of the chemical, biological, and materials sciences.

[‡] Genomics Institute of the Novartis Research Foundation.
[†] The Scripps Research Institute.

Examples include the biosynthesis of proteins and DNAs from unnatural building blocks,[12−19] the DNA- or protein-templated assembly of inorganic and organic materials,[20−24] combinatorial synthetic strategies inspired by Nature,[25−29]the construction of biohybrid thin films,[30] the use of enzymes in organic synthesis,[31,32] and the generation of orthogonal enzyme−inhibitor pairs.[33] We hope to illustrate the exciting opportunities that exist in synthesis at the interface of chemistry and biology through the examples detailed below, which, for reasons of space rather than significance, are derived largely from our own work.

Harnessing Nature's Biosynthetic Machinery

Nature has developed both templated and nontemplated biosynthetic machinery including the ribosome, DNA and RNA polymerases, polyketide and peptide synthases, and metabolic enzymes to make complex molecules with diverse functions. Moreover, the structure and properties of these molecules can be modified and enhanced by generating large numbers of analogues (through mutation, recombination, and amplification) and subjecting them to iterative selective processes. There are an increasing number of examples in which chemists have co-opted this natural biosynthetic machinery to create molecules with novel or enhanced functions. For example, the polyketide synthases are large multifunctional enzyme assemblies that consist of modules which encode ketosynthases, acyl transferases, ketoreductases, dehydrogenases, and enoyl reductases. Elegant studies by Khosla and others have shown that deletion of individual modules, alterations in the activity or specificity of a module, or addition of exogenous building blocks can lead to new macrolide antibiotics with enhanced activity[34−37] (which it may in the future be possible to further optimize through directed evolution). Similar strategies are being applied to the synthesis of novel nonribosomal and ribosomal-derived peptide antibiotics,[38−41] secondary metabolites,[42,43] glycopeptides,[44,45] and most recently biofuels.[46,47]

Other efforts focus on the generation of larger biomolecules with altered structures and functions. For example, DNAs are being synthesized *in vitro* by natural and engineered DNA polymerases that incorporate unnatural base pairs whose thermodynamic stabilities and fidelity of replication begin to rival or exceed those of the Watson−Crick A-T and G-C base pairs (Figure 2)[16−19,48−55] (the efficient *in vivo* replication and transcription of modified DNAs or RNAs with these unnatural bases remains a challenge for the future). Sugar and phosphate backbone replacements are also widely used to silence gene expression *in vivo*[56,57] and explore DNA structural constraints.[58] Natural proteins are also inspiring the synthesis of designed, folded polypeptides and, more recently, functional synthetic peptides and proteins.[59−64] In addition, a number of laboratories are focused on the synthesis of structurally defined, folded polymers entirely from unnatural building blocks (including *N*-alkylglycine and acyclic and cyclic β-amino acids[65−68]) — it may ultimately even be possible to produce these unnatural biopolymers enzymatically by templated synthesis. Of course, the above experiments were made possible by the pioneering work of Merrifield, Khorana, and Caruthers that enabled the efficient synthesis of large biological macromolecules.[69−71] In another exciting direction, the promiscuity of glycosyl transferases has been exploited to engineer glycoproteins which contain unnatural carbohydrates with orthogonal chemical handles to allow further chemical elaboration (including modification of cell surfaces).[72,73] This approach can also be applied to other protein and nucleic acid modifying enzymes and has

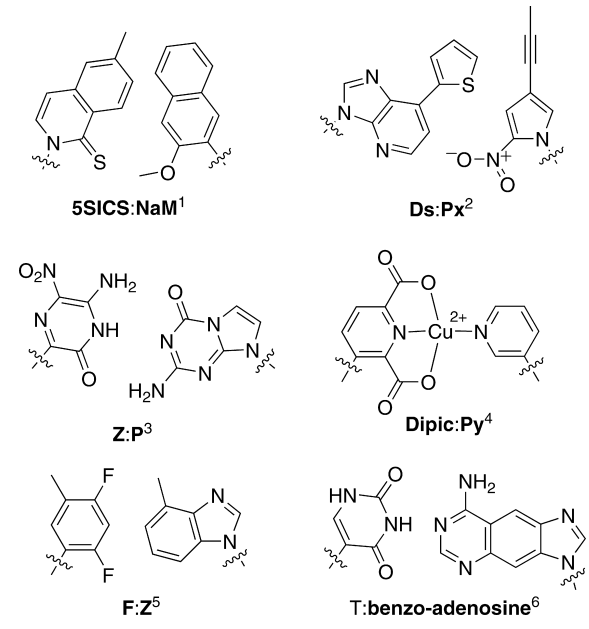


Figure 2. Representative unnatural base pairs. Only nucleobase analogues are shown; sugar and phosphate backbone have been omitted for clarity. Image courtesy of Floyd Romesberg.[18,50−55]

been used to selectively modify proteins with biophysical probes.[74−76] Perhaps in the future such methods will make it possible to control higher order cellular architectures in a defined way.

An Expanded Genetic Code

Synthesis of 21 Amino Acid Organisms. In our own work we asked the question whether our molecular level understanding and chemical/biological tools are sophisticated enough to begin to manipulate the genetic code itself, i.e., generate organisms that genetically encode 21 or more amino acids. Although the genetic codes of all known organisms specify the same 20 amino acids (with the rare exceptions of selenocysteine and pyrrolysine), it is clear that many proteins require additional chemistries associated with cofactors and post-translational modifications to carry out their natural functions. Therefore, although the functional groups contained in the 20 amino acid code might be sufficient for life, they might not be optimal. Consequently, the development of a general method that allows us to genetically encode additional amino acids beyond the canonical 20 might facilitate the evolution of proteins, or even entire organisms, with new or enhanced properties. Moreover, the ability to incorporate amino acids with defined steric/electronic properties and chemical reactivity at unique sites in proteins should provide powerful new tools for exploring protein structure and function, much the same way physical organic chemists use synthesis to understand the chemical reactivity of organic molecules.

Several approaches have been developed to insert unnatural amino acids into proteins. These include the use of native chemical ligation and intein-based methods for protein semisynthesis[77−79] and *in vitro* and cellular (microinjection) protein translation systems which make use of chemically aminoacylated tRNAs.[13,14,80−82] The latter have been used in detailed molecular studies of protein structure and function; notable examples include studies of the mechanism of ion channel gating,[83] and the quantification of the contribution of side-chain and backbone H-bonds to protein stability.[84,85] This methodology has even

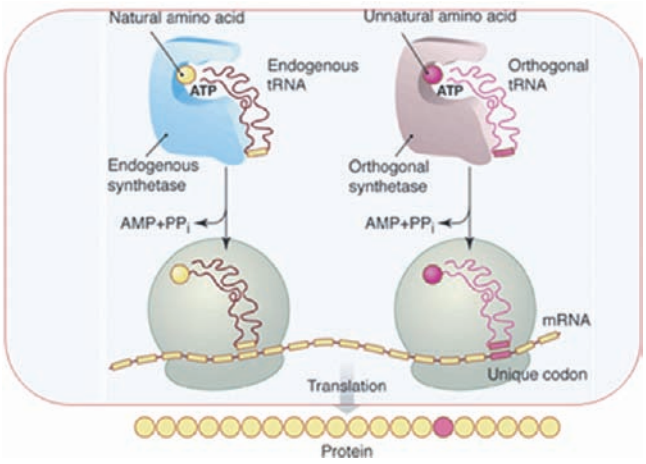


Figure 3. Protein biosynthesis with an expanded genetic code. Reprinted with permission from Ambrx.

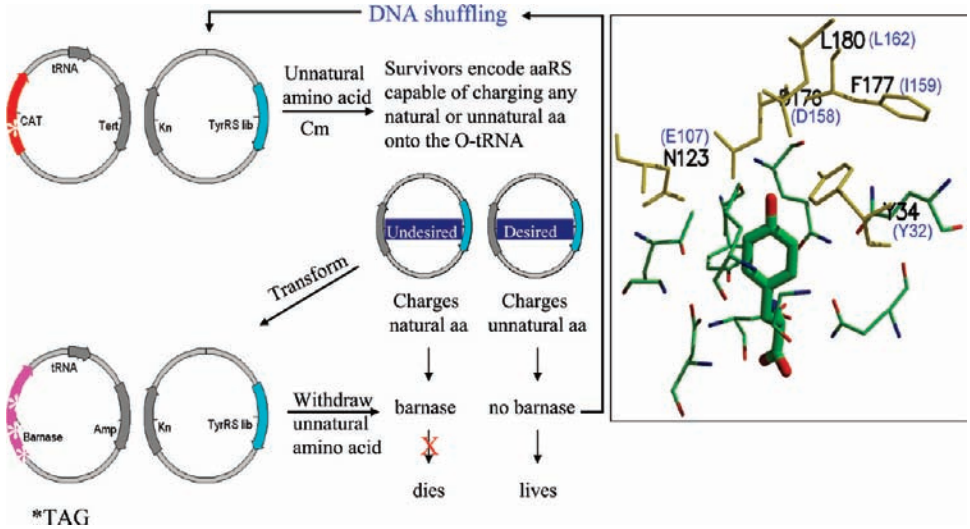
allowed the biosynthetic incorporation of D-amino acids into proteins.[86] However, these methods require the stoichiometric use of chemically aminoacylated tRNAs which results in relatively low yields of protein.[13,14] Alternatively, a number of *in vivo* methods have been developed to substitute the canonical amino acids in bacteria with unnatural amino analogues,[15,87−91] for example, substituting methionine with selenomethionine for phase determination in X-ray crystallography.[92] Perhaps the most elegant of these is that of Tirrell[15] in which rationally engineered mutations alter the specificity of aminoacyl-tRNA synthetases so that they can incorporate unnatural amino acids (*p*-azidophenylalanine, trifluoroisoleucine, etc.) in auxotrophic strains of *Escherichia coli* deficient in the biosynthesis of one of the 20 common amino acids. Although this approach overcomes the historical requirement that the unnatural amino acid be a close structural analogue of a common amino acid (so that it is misacylated by an endogenous aminoacyl-tRNA synthetase), the unnatural amino acid is still substituted for its corresponding canonical amino acid at all such sites throughout the proteome (i.e., one replaces one of the common 20 amino acids with an analogue; one does not add a new amino acid to the code). In addition, the use of auxotrophic strains can lead to partial incorporation of the unnatural amino acid in competition with endogenous amino acids. Nonetheless, these methods have found many exciting applications in the generation of novel biomaterials,[93] the surface immobilization of proteins,[94] the selective labeling of proteins with biophysical probes,[15,95] and the like.

The challenge we undertook was to develop a general method that makes it possible to *expand* the genetic code, i.e., incorporate additional amino acids (beyond the common 20) uniquely at any genetically specified site in a protein with the same high translational fidelity and efficiency characteristic of natural protein biosynthesis. The incorporation of additional amino acids into proteins directly in a living organism requires a unique tRNA:codon pair, a corresponding aminoacyl-tRNA synthetase, and significant intracellular levels of the unnatural amino acid (Figure 3).[12] First, the unnatural amino acid must be efficiently transported into the cytoplasm when added to the growth medium, or biosynthesized by the host, and it must be stable in the presence of endogenous metabolic enzymes (most unnatural amino acids meet these criteria). Next, to ensure that the unnatural amino acid is incorporated uniquely at the site specified by its codon, a tRNA must be constructed such that it is not recognized by the endogenous aminoacyl-tRNA synthetases (aaRS) of the host (21 in *E. coli*) but functions efficiently in translation (an orthogonal tRNA). Moreover, this tRNA must deliver the novel amino acid in response to a unique codon that does not encode any of the common 20 amino acids. This codon can be either one of the degenerate stop codons (e.g., an amber nonsense codon) or an efficient four-base frameshift codon. Another requirement for high fidelity is that the cognate aminoacyl-tRNA synthetase (an orthogonal synthetase) aminoacylates the orthogonal tRNA but does not aminoacylate any of the endogenous host tRNAs (86 in *E. coli*). Furthermore, this synthetase must aminoacylate the tRNA with only the desired unnatural amino acid, and not with any of the large number of endogenous amino acids of the host organism. Similarly, the unnatural amino acid cannot be a substrate for the endogenous synthetases if it is to be incorporated uniquely in response to its cognate codon. This set of severe specificity requirements represented the major synthetic challenge to the successful development of this methodology.

Fortunately, we have atomic-resolution structures of virtually the entire protein translational machinery,[96,97] so that the creation of new components with the requisite specificities is largely a chemical challenge. In brief, it was known that, for certain isoacceptor tRNAs, the identity elements of tRNAs from bacteria, archaea, and eukaryotes differ[98] so that, in theory, one could import a tRNA/aminoacyl-tRNA synthetase pair from archaea that would be functionally orthogonal in *E. coli*.[99] This proved not to be the case when an archael $tRNA^{Tyr}_{CUA}$ was imported into *E. coli*, so a large library of tRNA mutants was generated based on a consensus sequence analysis, and a general, two-step positive and negative selection scheme was developed to identify from this library orthogonal tRNA/aaRS pairs.[100] This strategy has proven a general approach for generating orthogonal tRNA/aaRS pairs; more recently, such pairs have also been generated from pyrrolysine-derived aminoacyl-tRNA synthetases.[101−103] To alter the substrate specificity of the orthogonal aminoacyl-tRNA synthetase so that it recognizes the desired unnatural amino acid and not any endogenous amino acids, a large, structure-based library ($\sim 10^9$ mutants) of active-site mutants was generated and subjected to a combination of positive and negative selections to identify a synthetase with the desired specificity (Scheme 1).[104] The positive selection is based on chloramphenicol resistance, which is conferred by the suppression of an amber mutation at a permissive site in the chloramphenicol acetyltransferase gene only in the presence of the unnatural amino acid. The negative selection uses the toxic barnase gene with amber mutations at permissive sites and is carried out in the absence of the unnatural amino acid to eliminate aaRS mutants that aminoacylate endogenous amino acids.

This selection scheme and more facile variants[105] have been used to develop orthogonal tRNA/aaRS pairs that are capable of selectively inserting one or more unnatural amino acids into proteins in *E. coli* in response to nonsense and/or four-base frameshift codons (with a cognate tRNA containing an expanded anticodon loop)[106] in good yields up to ~1 g/L and with high translational fidelities which rival those of natural protein biosynthesis. This system has been expanded to both yeast and mammalian cells by ourselves as well as in the laboratories of Yokoyama and Wang.[107−116] Experiments are now underway to generate a synthetic yeast, in which degenerate codons are freed up to encode unnatural amino acids, and to create transgenic worms and mice encoding a 21st amino acid of

Scheme 1[a]



[a] Right panel: Reprinted with permission from ref 104. Copyright 2001 American Association for the Advancement of Science.

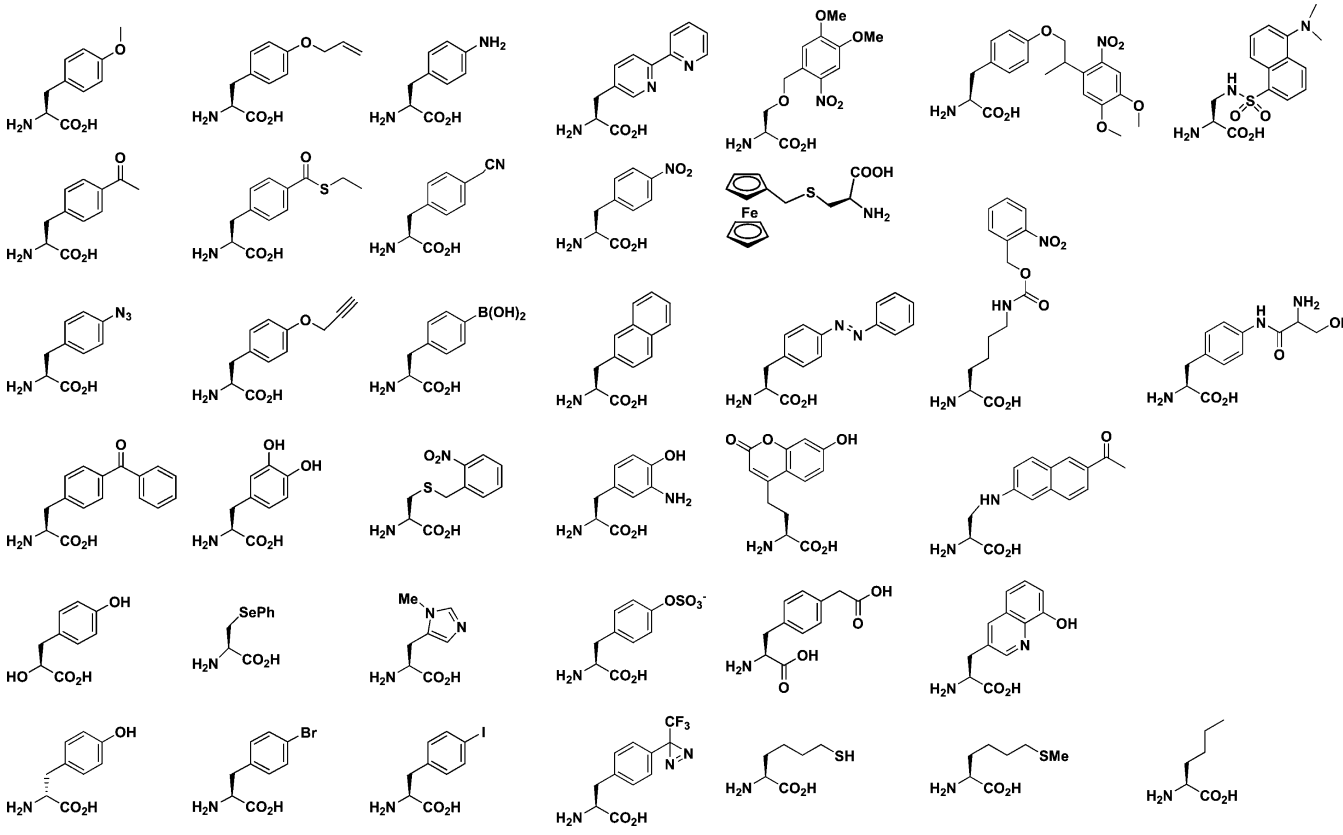


Figure 4. An expanding genetic code: examples of unnatural amino acids that have been genetically encoded in prokaryotic or eukaryotic organisms.[12]

choice. In addition, further improvements in the methodology are being pursued, including reducing context effects, enhancing suppression efficiency, and broadening substrate specificity through mutations in EF-Tu, tRNAs, the ribosome, release factors, etc.

Unnatural Amino Acids and Their Uses. Approximately 50 unnatural amino acids with novel chemical, biological, and physical properties have been genetically encoded in living organisms (Figure 4).[12,117] These include heavy atom containing amino acids to facilitate X-ray crystallographic studies;[118,119] amino acids with novel steric/packing and electronic properties for mechanistic studies;[116,120,121] photo-cross-linking amino acids which can be used to probe protein−protein and protein−nucleic acid interactions *in vitro* or *in vivo*[122−126] or to identify orphan ligands/receptors; keto, diketo, acetylene, azide, thioester, boronate, long-chain thiol- and dehydroalanine-containing amino acids that contain functional groups with unique chemical reactivity which can be used to site-specifically introduce a large number of biophysical probes, tags, toxins, and novel chemical functional groups into proteins *in vitro* or *in vivo* or to generate intra- or intermolecular cross-linked proteins and/or peptides;[127−133] redox-active amino acids to modulate electron

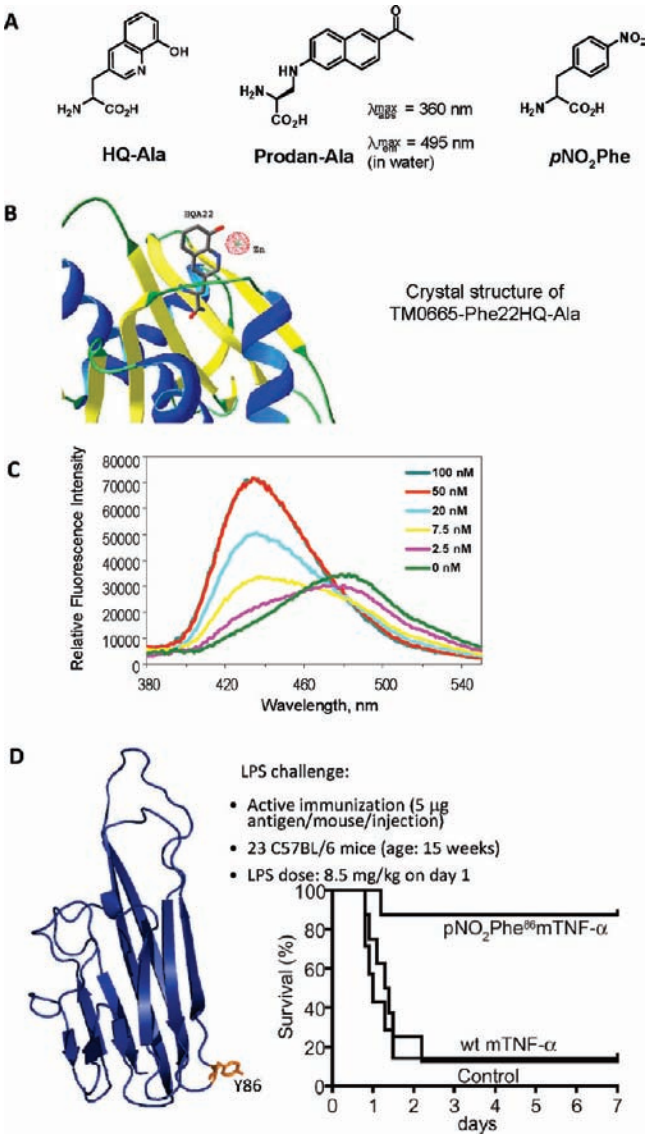


Figure 5. (A) Metal-ion-binding amino acid HQ-Ala, fluorescent amino acid Prodan-Ala, and immunogenic amino acid *p*-nitrophenylalanine. (B) Crystal structure of TM0665 Phe22→HQAla mutant.[118] (C) Fluorescence changes of an Asn160→Prodan-Ala mutant of glutamine binding protein upon addition of Gln.[163] (D) Substitution of the immunogenic amino acid *p*-nitrophenylalanine for Tyr86 in murine TNF-α and subsequent vaccination leads to a robust T-cell-driven immune response which cross-reacts with wild-type TNF-α and protects mice from LPS challenge.[159]

transfer in proteins;[134−136] photocaged and photoisomerizable amino acids to photoregulate cellular processes[137−141] such as signal transduction, protein trafficking, and transcription; metal-binding amino acids for catalysis, polypeptide self-assembly, X-ray phasing, and *in vivo* imaging (Figure 5A,B);[118,136,142,143] amino acids that contain NMR probes or fluorescent or IR-active side chains as local probes of protein structure and dynamics *in vitro* and *in vivo* (Figure 5A,C);[144−151] α-hydroxy acids and D-amino acids as probes of backbone conformation and hydrogen-bonding interactions[152] (raising the intriguing possibility of making folded "polyester" proteins); and sulfated amino acids and mimetics of phosphorylated amino acids as probes of protein post-translational modifications.[153,154] Clearly this list can likely be expanded to include many additional amino acids with novel chemical, physical, and biological properties. Importantly, we have solved the X-ray crystal structures of a number of these mutant aminoacyl-tRNA synthetases, and they reveal active sites with a high degree of plasticity, as evidenced by significant alterations in both active-site side chains and the polypeptide backbone to create new van der Waals packing and hydrogen-bonding interactions with the bound unnatural amino acid.[155,156]

There are many applications of these unnatural amino acids, including the following:

(1) The generation of therapeutic proteins with enhanced pharmacology (e.g., a long-lived, selectively pegylated human growth hormone made by Ambrx is in phase II clinical trials and is produced on greater than a 1000 L scale) — indeed this methodology promises for the first time to allow medicinal chemistry-like control over the structures of proteins to produce homogeneous therapeutic agents.[157] This is in contrast to the historical, relatively nonspecific methods for the chemical modification of therapeutic proteins with electrophilic moieties[158] or the selective modification of cysteine residues which is often complicated by the presence of multiple cysteines and/or their involvement in protein folding (e.g., immunoglobulins).

(2) The use of *p*-nitrophenylalanine mutants to break immunological self-tolerance and generate robust, long-lasting T- and B-cell-mediated immune responses that cross-react with wild-type proteins. This method for increasing the immunogenicity of self-proteins or weakly immunogenic pathogen proteins is currently being applied to the development of cancer and antiviral vaccines (Figure 5A,D).[159,160] Indeed the demonstration that simple nitration of phenylalanine at one site in a protein can break tolerance through T-cell-mediated mechanisms suggests that the enzymatic posttranslational generation of nitrotyrosine in proteins stimulated by local inflammation and cytokine release may be a general underlying initiating event in autoimmune disease.[161]

(3) The use of photocaged amino acids to photoactivate enzymatic activity or protein phosphorylation in living cells in a temporally and spatially defined fashion.[138,139]

(4) The generation of structurally-defined antibody conjugates including immunotoxins, antibody-based imaging agents, antibody−DNA conjugates, and bispecific antibodies as well as carrier−peptide conjugates with enhanced pharmacokinetics or targeted activities.[110]

(5) The use of environmentally sensitive, fluorescent amino acids (e.g., Prodan, dansyl, and 5-hydroxycoumarin side chains) for *in vitro* and cellular imaging of protein localization, biomolecular interactions, and conformational changes with the ability to place these small probes at virtually any site in the proteome.[148,162,163]

(6) The use of multidentate metal ion binding amino acids (e.g., bipyridyl and hydroxyquinoline side chains) to introduce mono- and bimetallic sites into proteins with redox and hydrolytic activities (either by selection or design) without the need to organize complex primary and secondary metal ion binding shells.[164]

(7) The use of redox amino acids as mechanistic probes of electron transfer in enzymes,[135] isotopically labeled amino acids as IR probes of protein dynamics,[146] and sterically modified amino acids as probes of ion channel activation.[116]

(8) The use of photo-cross-linking amino acids to map biomolecular interactions in cells and identify orphan ligands and receptors.[165−167]

(9) The use of uniquely reactive amino acids to introduce FRET pairs into proteins for single molecule spectroscopy studies of protein folding.[128]

(10) α-Hydroxy acids for protein purification with "traceless" affinity tags.[152]

In addition, we are beginning to examine the influence of an expanded genetic code on the evolution of peptides and proteins with new or enhanced properties. For example, a modified phage display system was used to evolve germline antibodies (with a randomized V_H CDR3 loop NNK library, where N = A, G, C, T and K = G or T) in strains that genetically encode sulfotyrosine or *p*-boronophenylalanine. We found that sulfotyrosine- and *p*-boronophenylalanine-containing Fab fragments emerged over all other clones in selections for gp120 and glucamine binding, respectively. In both cases, the selected mutant proteins utilized the unnatural amino acid to outcompete the other variants in the initial library, most of which contained only the 20 natural amino acids.[168,169] These results suggest that it may be possible to evolve antibodies with enhanced binding affinities to glycoproteins, serine proteases, metalloenzymes, and the like by incorporating amino acids with chemical "warheads" (hydroxamates, boronates, keto groups, etc.). A similar approach is now being applied to the synthesis and selection of acyclic and cyclic ribosomal peptides containing unnatural amino acids. It is likely that one can also either rationally design or evolve proteins with new or enhanced catalytic activities by introducing transition metal binding sites, or evolve novel folds from randomized codon libraries which encode both the canonical 20 as well as unnatural amino acids. The generation of orthogonal ribosomes by Chin and co-workers also promises to facilitate the ribosomal biosynthesis of unnatural biopolymers from nucleic acid templates by allowing mutations that do not affect natural translation.[170] Finally, we have also successfully "synthesized" an autonomous 21 amino acid bacterium that both biosynthesizes and genetically encodes the unnatural amino acid *p*-aminophenylalanine.[171] It will be of interest to compare its evolutionary fitness to that of wild-type *E. coli*. Thus, by seamlessly integrating the complex translational machinery of living cells with new chemistries and *in vitro* evolution methods, we have overcome an evolutionary constraint imposed by the universality of the genetic code. This advance may allow the generation of proteins and perhaps even living organisms with novel or enhanced properties, and underscores the power of co-opting (rather than mimicking) Nature to create novel new functions.

Molecular Diversity—Nature's Synthetic Strategy

Catalytic Antibodies. Another example of synergy between chemistry and biology in the generation of molecules with novel functions is the development and application of diversity-based synthetic strategies. This approach, inspired by Nature, involves the generation of large collections or "libraries" of molecules that are subsequently screened or selected on the basis of function. Indeed this represents one of living organisms' most powerful strategies for synthesizing molecules with desired properties. For example, the humoral immune system has developed highly sophisticated combinatorial and mutational mechanisms for generating large libraries of antibodies and selecting those that can recognize foreign antigens with high affinity and selectivity (Figure 6).[172] The notion that this natural diversity can be used to create novel chemical function was first illustrated with the generation of catalytic antibodies.[173−176] Rather than attempting to design a synthetic host that selectively binds a substrate of interest and then modify it with catalytic auxiliaries, it was realized that one could simply co-opt the immune system to generate a highly selective natural host in the form of an antibody combining site. To generate a selective catalyst rather than a selective receptor, stable transition-state analogues (rather than substrates) were used as antigens on the

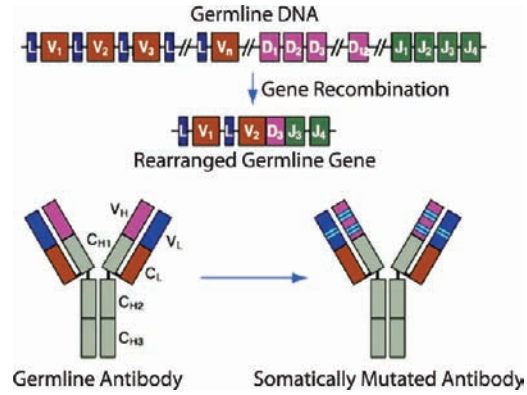


Figure 6. Combinatorial association of V, D, and J genes with recombination imprecision and subsequent somatic hypermutation during affinity maturation results in an enormous antibody repertoire capable of binding virtually any foreign molecule.[172]

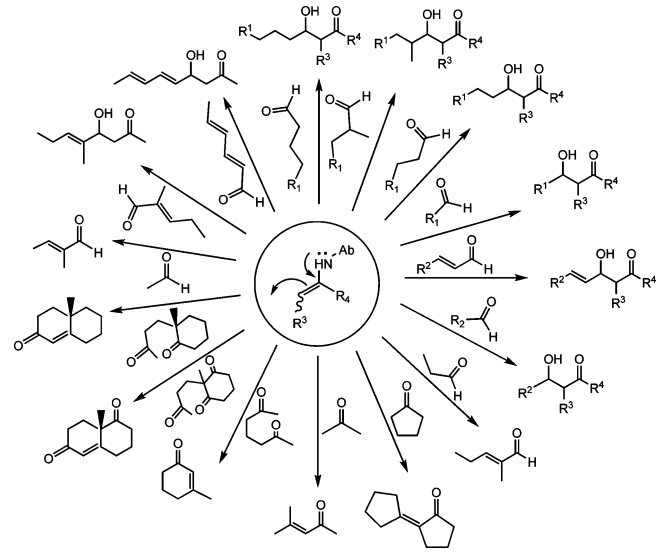


Figure 7. A catalytic antibody with broad substrate specificity that catalyzes aldol reactions through covalent catalysis. Reprinted with permission from ref 180. Copyright 1997 American Association for the Advancement of Science.

basis of the Pauling notion[177] that enzymes evolve maximum binding affinity to the transition state of a reaction. The early experiments by Lerner and co-workers and in our own laboratory involved the generation of esterolytic antibodies using phosphonate/phosphate transition-state analogues.[173,174] Other approaches have since been developed to generate catalytic antibodies, including covalent catalysis, proximity effects, and general acid−base catalysis[26] (thereby allowing us to dissect the contribution of each of these factors to biological catalysis). Using these approaches, antibodies have been generated that catalyze a wide array of chemical reactions, from acyl transfer and redox reactions to pericyclic and photochemical reactions with specificities and, in some cases, rates rivaling those of enzymes.[26,178,179] For example, Lerner and Barbas used a mechanism-based selection to generate antibodies that catalyze aldol reactions through covalent catalysis with catalytic efficiencies and selectivities remarkably similar to those of the Class I adolases (Figure 7).[180] Indeed one such antibody has made its way into clinical trials as a carrier for therapeutic peptides, which are bound covalently to the catalytic lysine residue. At the same time, mechanistic studies of this aldolase antibody led these investigators to the discovery that the simple

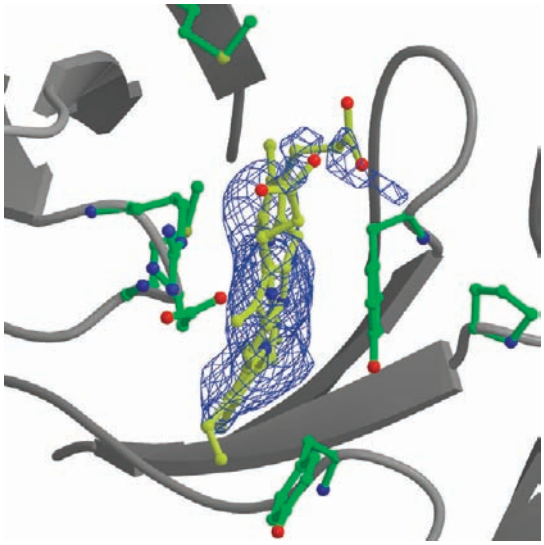


Figure 8. X-ray crystal structure of the Michaelis complex of the strained porphyrin substrate in a ferrochelatase antibody active site. Reprinted with permission from ref 185. Copyright 2003 The National Academy of Sciences, U.S.A.

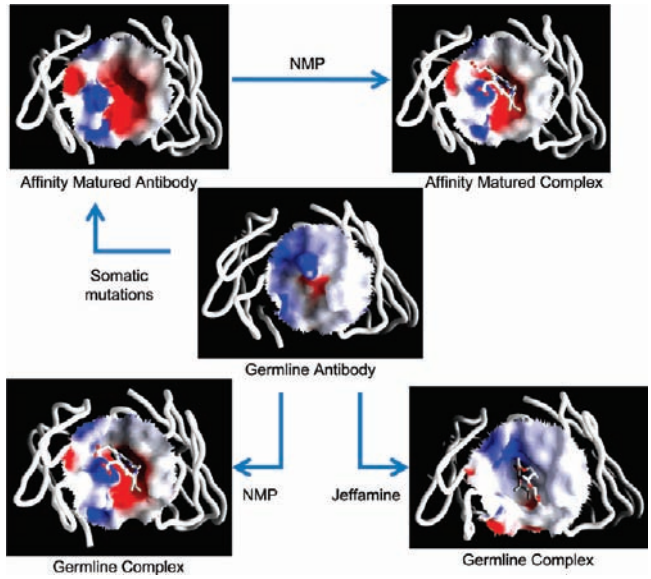


Figure 9. Structural plasticity of a germline antibody that binds *N*-methylmesoporphyrin (NMP). Somatic mutations during affinity maturation lock the optimal active site conformation.[188]

amino acid proline could act as an asymmetric organocatalyst.[181,182] Antibodies have also been generated to catalyze "difficult" chemical transformations such as kinetically disfavored exo-Diels–Alder and anti-Baldwin cyclization reactions.[183]

The detailed characterization of the immunological evolution, three-dimensional structures, and mechanisms of catalytic antibodies has also helped to dissect and quantify the relationship between binding energy and catalysis in the evolution of catalytic function. Indeed the use of transition-state analogues to elicit catalytic antibodies provided "proof by synthesis" of the Pauling notion of enzymatic catalysis. In another example, a "ferrochelatase" antibody, which catalyzes the efficient insertion of metal ions into porphyrin (the last step in heme biosynthesis), was generated against an *N*-methyl porphyrin, which mimics the distorted porphyrin ring of the putative transition state for metalation.[184] The crystal structure of the Michaelis complex (Figure 8)[185] indeed showed that the substrate is bound in a strained conformation, providing the first direct structural evidence for the theory of substrate strain proposed by Haldane over 70 years ago.[186]

The characterization of catalytic antibodies has also provided fundamental insight into the mechanisms by which the immune system itself evolves selective receptors. For example, the first detailed structural comparisons of germline and affinity-matured antibodies revealed the critical role of structural plasticity (in addition to genetic diversity) in determining the tremendous binding potential of the germline antibody repertoire (Figure 9).[187–189] Germline antibodies appear to have a high degree of intrinsic combining site conformational flexibility (reminiscent of the chemical instruction theory of the immune response proposed by Haurowitz[190] and Pauling[191]) which allows them to bind multiple, distinct ligands in different conformational states.[192] That conformational state which binds a specific antigen is then locked and further refined by somatic mutations which occur during affinity maturation (not protein folding as proposed by Pauling). Structural and biophysical analyses of the immunological evolution of catalytic antibodies also pointed to the critical role of mutations distal to the active site in controlling the binding and catalytic activity of proteins through complex networks of side chain and backbone interactions.[193] Indeed these studies underscore a key aspect of diversity-based synthetic strategies—detailed analyses of the relationship between molecular structure and properties in molecules obtained by combinatorial methods (properties which may be difficult to obtain by more conventional synthetic approaches) often lead to new chemical insight which further increases our ability to generate new molecular function from basic chemical principles.

Other Applications of Biological Diversity. The demonstration that the vast structural diversity of antibody molecules can be redirected with proper chemical instruction to generate selective catalysts illustrated the utility of molecular diversity (the antibody repertoire in this case) as a new, biologically inspired "synthetic strategy" to create novel chemical properties. Shortly thereafter, libraries of other biomolecules were designed and synthesized in order to identify molecules with new or enhanced functions. These included the use of phage display libraries[27,194,195] to generate peptides, proteins, and antibody fragments with novel specificities—for example, a peptide dimer with erythropoietin-like activity,[196] zinc finger proteins with new DNA binding specificities,[197–199] peptibodies, and more recently, polypeptides that template inorganic materials.[200] An example of the latter involves the selection of peptides that specifically bind semiconductors to direct nanoparticle assembly.[201] Libraries of random RNA sequences (including those containing unnatural bases with novel functional groups) have been transcribed and subjected to *in vitro* selections to identify RNAs that selectively bind ligands with high affinity (aptamers),[202,203] that catalyze chemical reactions such as acyl or phosphoryl transfers,[204–206] or whose structure and transcription is regulated by the binding of small synthetic molecules.[207] Indeed it may even be possible to use diversity-based approaches to construct completely synthetic viruses (e.g., small RNA viroids that do not encode proteins) that target specific cell types, or even self-replicating systems of molecules.[208–210] In addition, combinatorial-based diversity approaches such as DNA shuffling (in which positively selected point mutants are recombined to yield additive improvements in function) are also being used to evolve proteins with new catalytic activities, specificities, and regulation;[211,213]

and cellular complementation strategies are being used to evolve enzymes for the *in vivo* manipulation of biomolecules.[212] The inherent somatic mutation machinery of the B-cell itself has also been harnessed to evolve fluorescent proteins with enhanced photophysical properties.[214] Random libraries of amino acids or secondary structural elements have even been created to identify minimalist and novel protein folds.[215−218] Diversity-based approaches are also being used to create synthetic biological circuitry in which a defined input is biologically processed to give a specific output (e.g., bacteria migrate to illuminated regions of a plate, or a binding event initiates enzymatic amplification of a specific signal).[219,220] Finally, libraries of peptides, oligonucleotides, and unnatural biopolymers provide a powerful approach to identify molecules that facilitate the trafficking of therapeutic small molecules, polypeptides, DNAs, and siRNAs across cellular, endosomal, and intestinal membranes—a major challenge in medicine.[221−223] These various applications are beyond the scope of this Perspective but represent important new directions for synthesis at the interface of chemistry and biology.

Applications of Molecular Diversity to Chemistry

Materials Science. Today, combinatorial strategies, which in their most basic form involve the parallel synthesis of large numbers of chemically diverse structures around a central framework (typically chosen on the basis of theoretical or empirical considerations), are impacting many areas of chemistry.[25] This method is particularly valuable when theory has insufficient predictive power to guide molecular design with precision—it quickly provides large amounts of experimental data around initial candidate structures to either iterate with theoretical predictions or develop improved empirical models to guide additional experiments. One particularly illustrative example of how a synthetic concept, adopted from the natural process of mutation and selection, can impact a very distinct scientific discipline is the application of diversity-based approaches to the generation of solid-state materials with novel properties.[224−228] The properties of many functional materials, such as high-temperature superconductors, heterogeneous catalysts, ferroelectric materials, magnets, and even structural materials, arise from complex interactions involving the host structure, dopants, defects, and morphology, all of which are highly dependent on composition and processing. Unfortunately, our current level of theoretical understanding does not generally allow one to predict the structures and resulting properties of these materials. The situation is further complicated by the complex compositions and structures of many modern materials and size-dependent properties, and by the fact that materials synthesis, unlike organic synthesis, is generally not a kinetically controlled process. Given the large number of elements in the periodic table that can be used to make compositions consisting of up to six elements, the universe of possible new compounds with interesting physical and chemical properties remains largely unexplored; combinatorial synthetic methods represent a powerful way for experimentalists and theorists alike to more effectively mine this huge chemical space for interesting new materials properties.[229−232]

The first application of combinatorial methods to materials science involved the synthesis and screening of libraries of thin-film copper oxides to identify high-temperature superconductors.[224] Libraries of solid-state materials were synthesized to explore many different compositions of interest by sequentially sputtering precursors at different sites on a substrate using a

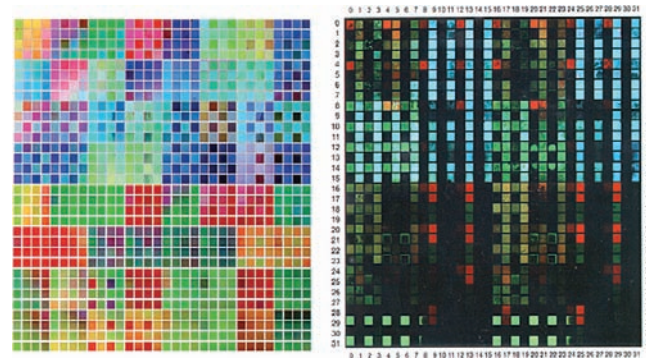


Figure 10. A library of novel luminescent molecules generated by laser ablation of metal oxides through a series of physical masks, under both ambient and UV irradiation. Reprinted with permission from ref 226. Copyright 1998 American Association for the Advancement of Science.

series of precisely positioned physical masks. Low-temperature annealing followed by high-temperature processing resulted in the formation of superconducting thin films. More recently, quaternary and shutter masking systems, together with photolithographic masking techniques and pulsed laser deposition, have made possible the synthesis of high-quality, diverse thin-film libraries of some 1000−10 000 samples (Figure 10). We and others have used such libraries to discover phosphors with novel luminescent properties and structures, new families of giant magnetoresistive materials, and ferroelectric materials.[224–227] Solution-based and bulk methods (e.g., ball milling) have also been applied to materials and catalyst library synthesis[233] (and, in an interesting twist, have been used to explore phase diagrams for crystalline proteins[234]). In addition, a large number of scanning or parallel detection systems have been developed for rapidly screening materials libraries for optical, electronic, magnetic, adsorptive, or catalytic properties of interest. This combinatorial approach to materials discovery, which was dramatically expanded in scope by Weinberg and co-workers at Symyx, is now practiced in many industries and has led to new olefin polymerization and oxidative catalysts, hydrogen storage materials, separations materials, dielectrics, phosphors, etc. and is now being applied to the optimization of complex integrated devices such as lithium ion batteries, solar cells, and computer chips.[228–233,235−238] Indeed with the challenges we now face for new environmentally friendly energy sources, combinatorial methods are likely to play a critical role in the development of enabling new materials—these include new hydrogen and methane storage materials, fuel cell catalysts, photovoltaic devices, CO_2 sequestrants, and high-energy-density batteries. This will likely be best achieved by a synergistic use of combinatorial approaches, more conventional solid-state chemistry, and theory.

Biologically Active Small Molecules. Another particularly powerful application of combinatorial strategies, pioneered by Ellman and co-workers with their work on benzodiazepines, involves the synthesis of diverse libraries of nonoligomeric synthetic molecules.[28,239,240] Just as large libraries of antibodies are genetically assembled from families of V, D, and J gene segments, it was realized that libraries of small organic molecules could be efficiently assembled from chemical building blocks. Although there are many examples of the rational design of biologically active small molecules (notable examples include mechanism-based inhibitors by Bloch,[241] reporters of cellular messengers by Tsien,[242] and the synthesis of captopril and Prozac by chemists at Squibb and Eli Lilly, respectively[243,244]),

Figure 11. Automated high-throughput screening systems and chemical, genomic, and protein libraries are enabling large-scale, cost-effective cellular screens.[245] Image courtesy of GNF Engineering.

it remains a challenge to design *a priori* molecules that selectively activate or inhibit a desired enzyme or receptor, or modulate a specific cellular signaling pathway, regulatory circuit, or transcriptional program. As a consequence, the screening of synthetic chemical libraries offers a highly effective approach to identify biologically active molecules, especially molecules with novel cellular activities which may not be predicted or even conceived of in hypothesis-driven experiments. Chemical leads from these screens can then form the basis for further optimization experiments (including structure-based design), just as natural products have long inspired medicinal chemists in the synthesis of new drugs. The real challenge for chemists here is not chemical synthesis itself, but rather the creation and integration of productive chemical libraries with well-designed cellular screens and screening strategies (transcriptional, pathway, image-based, and phenotypic screens) to find molecules that selectively modulate biological systems in interesting and/or even unprecedented ways, and the subsequent characterization of their mechanisms of action.

In particular, combinatorial strategies allow one to assimilate large libraries (>100 000 molecules) of diverse molecular structures relatively easily and inexpensively (in contrast to the historical large collections of synthetic compounds and natural products that were found almost exclusively in large pharmaceutical companies). In addition, high-throughput screening systems in which compound libraries can be screened in parallel in biochemical or cellular assays in <10 μL volumes (and in almost any format from FRET and ELISA assays to luminescent and high-content imaging screens) have reduced the cost of assays (as well as the amount of compound required) from dollars per data point to pennies per data point. These advances have now made small-molecule screens a viable research tool for many academic laboratories and not just the domain of industrial research efforts.[245,246] Recent innovations in micro- and nanofluidics and dispensing technologies have further miniaturized the screening format;[247] and automated cell culture systems in which 100+ cell lines can be passaged, plated, and assayed in a completely automated fashion allow massively parallel screening of protein families for selective ligands (Figure 11).[245] As a consequence, the number of small molecules that modulate protein or RNA targets or cellular pathways will grow enormously in the next 5−10 years. Indeed the screening database at the Genomics Institute of the Novartis Research Foundation (GNF) now exceeds 500 million data points, allowing one to rapidly correlate the activity of "hits" from any given screen to hundreds of other cell-based or biochemical screens to determine selectivity and/or potential cellular pathways that are affected through cluster analysis. For example, screening of >2 Mio molecules in a *Plasmodium falciparum* proliferation assay in red blood cells yielded >1000 active compounds which had antimalarial activity and minimal activity against human cells.[248] Cluster analysis revealed a number of novel scaffolds that targeted both known and novel biochemical pathways and were active against mutant strains. New chemoinformatic and data mining tools will be essential to fully exploit the data generated by these screens in the coming years.

Since it is impossible to represent all chemical space corresponding to low-molecular-weight (LMW) compounds (<500 Da), various arguments have been put forth regarding the structural motifs that should be represented in chemical libraries—from complex natural product-like molecules[249−251] to heterocycles such as purines, benzimidazoles, indoles, quinolines,[252] and other "privileged" scaffolds (Figure 12).[253] The latter have the advantage that they have historically yielded biologically active molecules that selectively modulate the activities of enzymes, GPCRs, nuclear receptors, ion channels, and the like. Moreover, they can be rapidly synthesized, and there exists a wealth of information on the pharmacology of these scaffolds—which is critical to the chemical optimization of these structures for *in vivo* studies. Indeed our experience is that these heterocycle libraries are an extremely rich source of selective and potent hits (in both cellular and biochemical assays) whose properties can be rapidly optimized. There are many other approaches being pursued toward library design, including the synthesis of metal-templated combinatorial libraries.[254]

One can ask why academic laboratories should screen small-molecule libraries for novel biological activities, given that this activity has historically been the focus of the pharmaceutical and biotechnology industries. Arguments have been made that unbiased cell-based screens of small molecules offer an alternative to genetic approaches to identify novel genes associated with a cellular phenotype.[255,256] Indeed novel gene products have been identified from synthetic molecules and natural products: the identification of FK506 binding protein independently by Schreiber and the Merck group using affinity-based methods represents one outstanding example.[257] However, the lack of selectivity and weak potency of many primary screen hits (in comparison to natural products), the need for structure−activity relationship information to optimize activity and generate affinity probes, and the complexities associated with either genomic (mRNA expression analysis, cDNA overexpression) or proteomic (phosphoprotein profiling and protein affinity arrays) target deconvolution make this approach challenging even for laboratories with considerable chemical and biological expertise. The availability of genome-wide cDNA, siRNA or shRNA, and noncoding RNA libraries in transfectable or retro-/lentiviral format provides a more straightforward approach to identify new gene function that requires no deconvolution and is generally far more selective in targeting a specific gene.[258−260] However, as improved methods for identifying the targets and mechanisms of biologically active small molecules are developed, cellular screens of chemical libraries will become increasingly important in cell biology.[261] Indeed such methods represent probably the most significant current challenge to the effective use of cell-based small-molecule screens to discover novel biology.

Nonetheless, cell-based screens of small-molecule libraries do allow one, in a single experiment, to interrogate large numbers of interacting proteins and nucleic acids (many with unknown function) in an unbiased fashion to identify synthetic

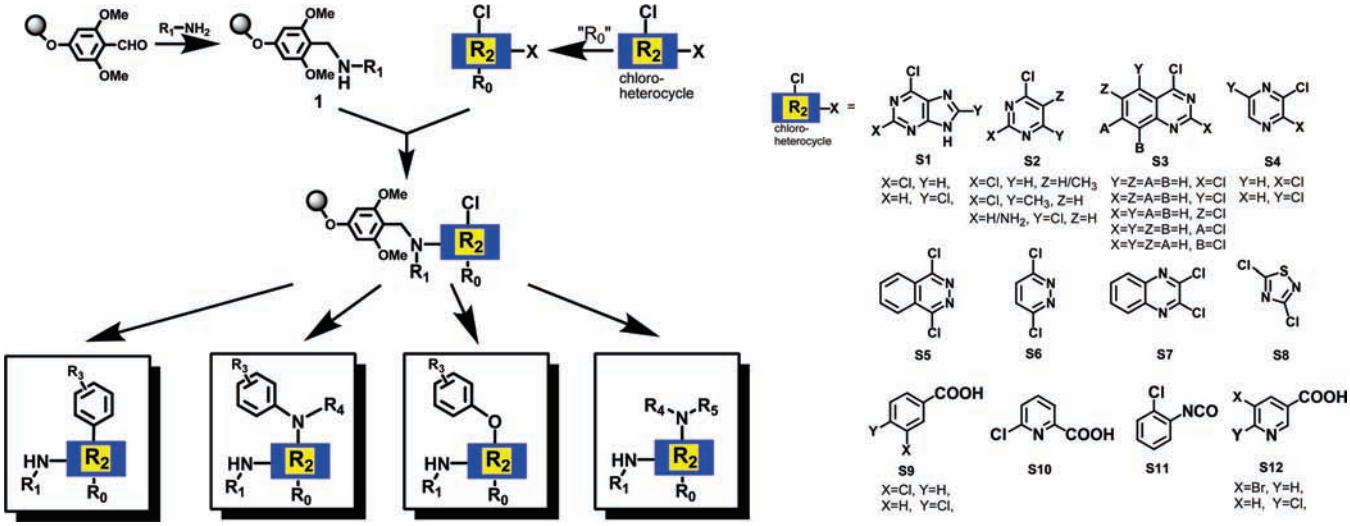


Figure 12. An efficient synthesis of heterocycle libraries.[252]

molecules that affect biological systems in novel ways. Furthermore, in contrast to most genetic methods, small molecules are able to perturb specific functions of a single protein (or multiple proteins) with temporal control and without affecting its other activities, which is an especially useful feature for dissecting signaling pathways where proteins often serve both catalytic and recruitment functions that cannot be addressed independently with genetic approaches. Indeed the power of cell-based screens dates back to the pioneering studies by Ehrlich, in which synthetic arsenicals were identified that selectively killed syphilis.[262] This cell biological approach was largely displaced by biochemical target-based approaches, but with large chemical libraries, highly sensitive screening and imaging technologies, and genomic and proteomic target deconvolution strategies, cellular screens once again offer a powerful strategy for finding synthetic molecules with novel biological and potentially therapeutic activities.

Another real advantage of the small molecules derived from such screens is that they can be relatively quickly translated to primary cells and *in vivo* models to test biological hypotheses in the complex setting of whole organism physiology (without the need to generate knockout or transgenic animals or to use viral gene or siRNA delivery) — ultimately with the hope of finding new medicines that operate by novel mechanisms. To do so, however, requires a knowledge of basic medicinal chemistry and pharmacology to optimize the properties of molecules including but not limited to potency, selectivity (indeed most molecules isolated from screens lack high selectivity), bioavailability, serum half-life, blood−brain barrier penetration, biodistribution, and the like.[263] This aspect of "chemical biology" has been largely ignored or less explored by many in the academic research community — an oversight that must be remedied if we are to realize the full biomedical value of the molecules that will result from such screening activities. Indeed it is likely that we will see a large increase in the number of therapeutic agents resulting from academic efforts as chemists become more sophisticated in *in vitro* and *in vivo* pharmacology, animal models, and cell biology.

Chemical Libraries and Stem Cell Biology. Given the increased availability and decreased cost of chemical libraries and the power of modern screening technologies, what opportunities should the academic chemistry community pursue with these new tools? One answer is to focus on those areas of biology which are still poorly understood and, as a consequence, there exists a real need for small molecules as *in vitro* and *in vivo* probes; another is to focus on major unmet medical needs that have been largely ignored by industrial research efforts due to perceived risk or financial considerations. A timely example (of both) is regenerative medicine, in which new cells (e.g., neurons, muscle, chondrocytes, etc.) are generated to replace tissues lost to degenerative diseases or aging. Small molecules offer an alternative to cell-based therapies or gene therapy, which carry with them the possibility of developing cancers. To this end, we and others have been carrying out cell-based screens to identify molecules that (1) control the self-renewal or lineage-specific differentiation of adult and embryonic stem cells (cells with the capacity to self-renew or differentiate into specialized tissues) and (2) control the proliferation or reprogramming of differentiated cells.[264−271] For example, we have carried out image-based screens (using cellular markers such as CD34, CD133, CD41, etc.) with one class of adult stem cells, hematopoietic stem cells (HSCs), for molecules that control self-renewal and differentiation (HSCs are adult stem cells that give rise to all the blood lineages such as macrophages, B and T cells, platelets, red blood cells, etc.). Molecules have been identified that induce the selective differentiation of human HSCs toward megakaryocytes; molecules have also been found that are able to significantly expand both human peripheral mobilized HSCs and umbilical cord blood HSCs in an undifferentiated state such that they efficiently engraft in a NOG mouse model and give rise to all blood cell lineages.[272] The latter molecules represent a new regenerative medicine to expand cord blood HSCs for the large number of cancer, blood, and autoimmune disease patients for which no matched donors exist, but for which matches do exist in cord blood banks (but the number of cells is too small for effective engraftment in adult patients). In other experiments with adult stem cells, we have used cell-based screens, coupled with immunostaining of neuronal markers, to identify a molecule that induces the selective neurogenesis of neural progenitor cells *in vitro* and *in vivo* in the rat dentate gyrus. This compound acts by selectively binding the centrosomal protein Tacc3, which has been previously implicated in regulating the balance between progenitor cell renewal and differentiation.[273,274] This and other such molecules may ultimately lead to new treatments for neurodegenerative disease. In other experiments, we have used alkaline

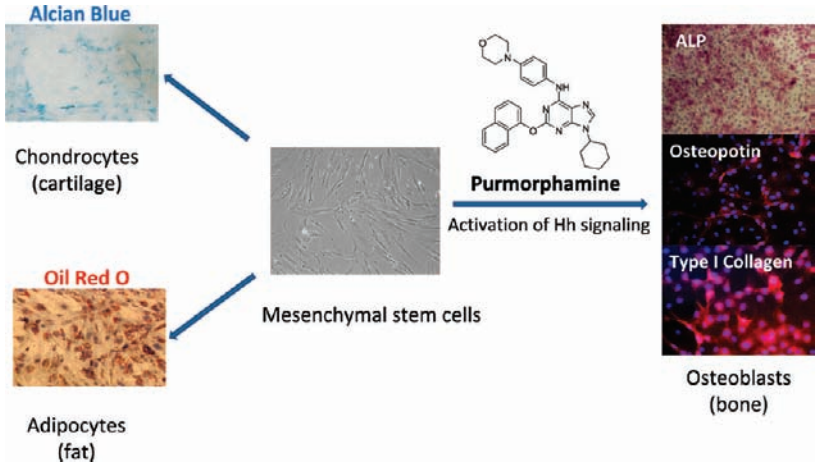


Figure 13. Purmorphamine induces the selective differentiation of mesenchymal stem cells to osteoblasts by activating the Hedgehog signaling pathway.[275]

phosphatase and Alcian blue-based screens to identify molecules that selectively induce mesenchymal stem cells (adult stem cells which normally give rise to osteoblasts, adipocytes, and chondrocytes) to undergo osteogenesis to form bone, or chondrogenesis to form cartilage (with the potential to treat early-stage osteoarthritis) on the basis of analysis of multiple tissue-specific markers, cellular morphology, and the like (Figure 13).[275,276] The former molecule, purmorphamine, functions as an agonist of Hedgehog signaling, a key developmental pathway.[275] Other laboratories have also recently identified small molecules that modulate stem cell fate. Examples include molecules that suppress osteoclastogenesis, enhance differentiation of osteoprogenitor cells to increase bone formation, affect myocardium repair, and modulate the neuronal differentiation of stem cells.[277−283] In addition, small-molecule agonists and antagonists of important developmental signaling pathways (Wnt, Hh, BMP, etc.) have been discovered and are serving as important tools to study stem cell biology and cancer.[284−288] Many additional exciting opportunities exist for chemists with adult stem cells, including screens for molecules that expand satellite cells, endothelial progenitors or neural stem cells for muscular dystrophies, cardiovascular repair, neurodegeneration, and the like. One can also attempt to identify molecules that selectively induce apoptosis or differentiation of cancer stem cells which have been implicated in a number of different cancers (stem-like cells from which bulk tumors are thought to arise and which are often resistant to chemotherapies).[289] For example, we have identified small molecules that kill prostate-tumor-initiating cells by blocking the PI3K signaling pathway.[290] In combination with the antitumor drug Taxotere, this inhibitor is extremely effective in regressing established tumors in a mouse xenograft model.

One can also screen for molecules that affect the self-renewal or differentiation of embryonic stem cells. For example, Oct4-green fluorescent protein (GFP) (Oct4 is a gene expressed in embryonic stem cells, ESCs) and ANF-luciferase (atrial natriuretic factor, ANF, is a cardiomyocyte-specific marker) transcriptional reporter screens have been used to identify small molecules that allow one to either expand ESCs in an undifferentiated state[291] or induce their differentiation to cardiomyocytes,[292] respectively. In the former case, affinity-based methods showed that the molecule acts by binding *both* RasGAP and Erk1/2 and thereby modulating both cellular proliferation and differentiation. This combinatorial effect (one molecule modulating two targets) was confirmed both by abrogating the activity of the compound by overexpressing cDNAs for both genes and by independently knocking down the activity of each gene with siRNAs or chemical inhibitors.[291] Recently, a molecule (TWS119) previously identified in our laboratory as a neurogenesis inducing agent in ESCs[293] was reported to also generate $CD8^+$ memory stem cells with proliferative and antitumor activities, and could have implications for the design of new vaccine strategies.[294] In another screen, which was based on the induction of SOX17 expression (an endodermal marker) in the presence of activin A, we identified a molecule that efficiently potentiates ESCs to differentiate into ectodermal, mesodermal, or endodermal lineages in high yield depending on the additional lineage specification factors present (in chemically defined media). This molecule functions by binding and blocking the nuclear translocation of the protein NME2, which leads to downregulation of c-Myc, a key ES self-renewal regulator and oncogene (Figure 14A).[295] In a related series of experiments by Melton and co-workers, it was recently shown that two small molecules (IDE1 and IDE2) can selectively induce human ESCs to differentiate into definitive endoderm,[296] and subsequent treatment with (−)-indolactam V can further mature the definitive endoderm cells to pancreatic progenitor cells (Figure 14B,C).[297] Combinations of molecules of this sort may facilitate the generation of hepatocytes, endocrine, and other cell types from ESCs in high yield under chemically defined conditions. One can also expand these activities beyond small molecules and screen libraries of purified, secreted proteins and peptides (the majority of which have unknown functions) to identify circulating polypeptides that affect stem cells in novel ways.[298] For example, we recently identified the protein pigment epithelium-derived factor that allows one to expand huESCs in defined media.[299]

Another exciting direction in the field of regenerative medicine is the reprogramming of somatic cells[269] to create both cellular models of genetic disease and potentially cells for clinical use in regenerative therapies. Recently, it was shown by Yamanaka and co-workers that the overexpression of four genes (*Oct4*, *Sox2*, *c-Myc*, and *Klf4*) in mouse embryonic fibroblasts (MEFs) or human fibroblasts reverts them to embryonic stem-like cells (termed induced pluripotent stem cells, or iPS cells) that can then differentiate into all three germ layers (Figure 15).[300−304] Clearly the identification of drug-like small molecules with such activities might allow

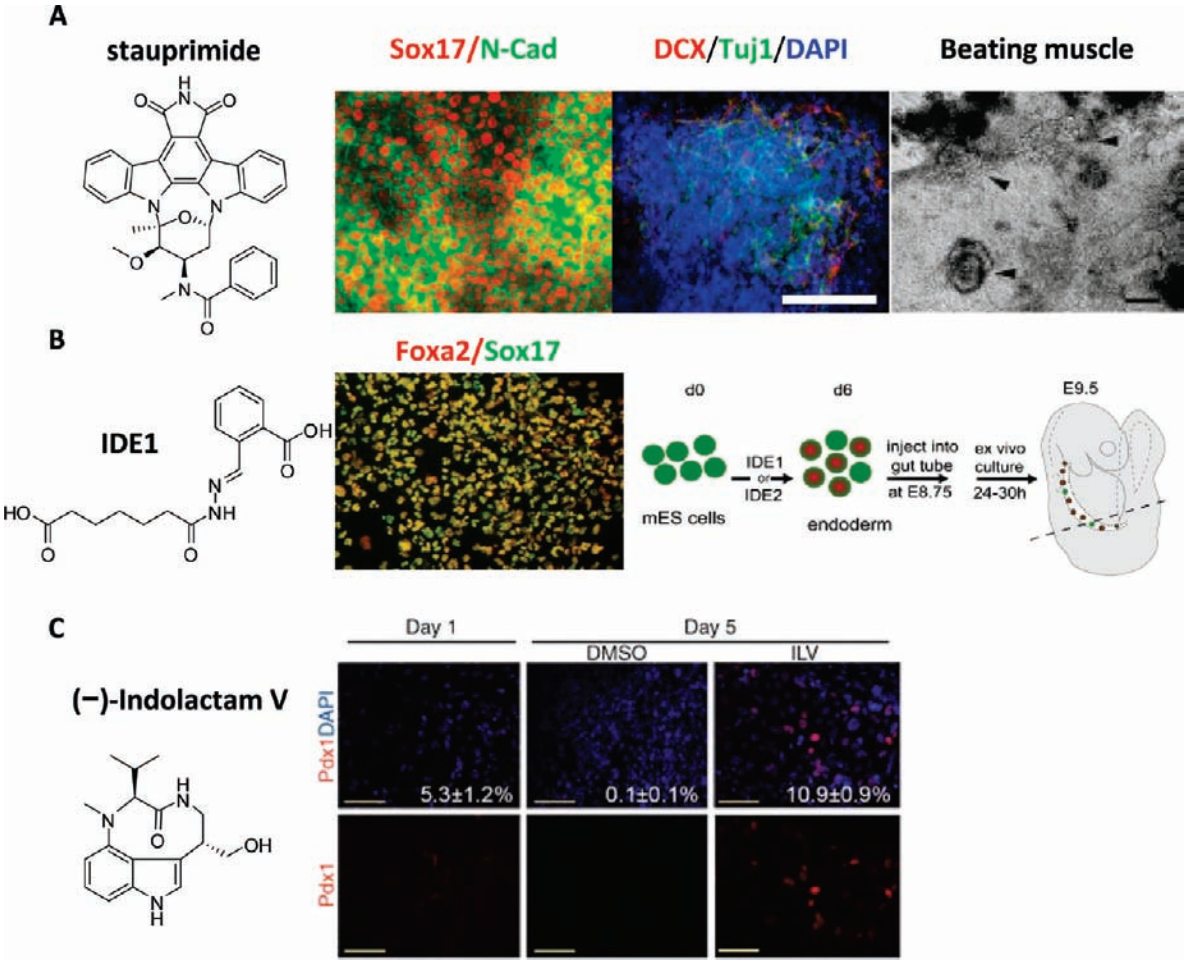


Figure 14. (A) Stauprimide potentiates the differentiation of ES cells toward multiple lineages (definitive endoderm, neural ectoderm, and mesoderm) when combined with lineage specification cues. Reprinted with permission from ref 295. Copyright 2009 Elsevier. (B) IDE1 and IDE2 induce lineage specification of ES cells to definitive endoderm. Chemically derived endoderm cells are functional *in vivo*. Reprinted with permission from ref 296. Copyright 2009 Elsevier. (C) (−)-Indolactam V (ILV) matures definitive endoderm cells to Pdx1 expressing pancreatic progenitor cells. Reprinted with permission from ref 297. Copyright 2009 Nature Publishing Group.

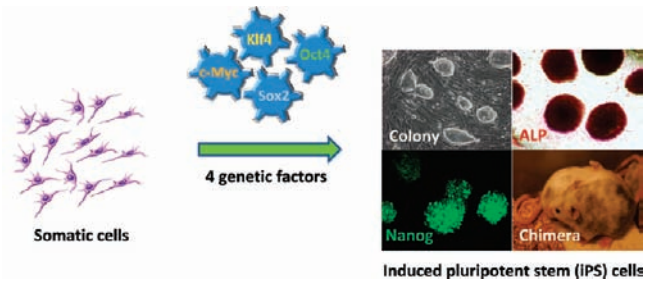


Figure 15. Cellular reprogramming. When a combination of genetic factors (for example, Oct4/Sox2/Klf4/c-Myc) is introduced into somatic cells, those cell could be reprogrammed to form induced pluripotent stem (iPS) cells. iPS cells show characteristics of pluripotent embryonic stem cells (tight colonies; expression of embryonic stem cell markers such as alkaline phosphatase (ALP) and Nanog; mouse iPS cells could form chimeric mice when injected into developing mouse embryos).

one to create desired cell types (e.g., those lost to degenerative disease: neurons, β-cells, etc.) from other tissues patient-specifically, without the need for nuclear transfer or embryonic stem cells. For example, we have identified a molecule (reversine) from a phenotypic screen that can reprogram myoblasts (muscle precursors) to a multipotent state which can then be differentiated into adipocytes (fat cells) or osteoblasts (bone cells) (Figure 16).[305,306] Clonal analysis demonstrated reprogramming rather than a side population artifact, and the relevant targets of reversine were found to be MEK1/2 and nonmuscle myosin II. Similarly, molecules have been identified that reprogram oligodendrocyte precursor cells (which normally give rise to glial cells in the brain) to neural stem-like cells, such that they can be further converted to neurons and function *in vivo*.[307,308] Many laboratories are now attempting to identify molecules that replace single factors or combinations of factors to generate iPS cells (in our case using a Nanog promoter driven luciferase reporter which allows >2 Mio compounds to be efficiently screened) — indeed we have identified molecules that independently replace *Sox2* and *Klf4* in reprogramming of embryonic fibroblasts.[309] Others have also demonstrated that small-molecule inhibitors of histone modification (methylation, deacetylation) and DNA methylation, agonists of L-type calcium channels, and kinase inhibitors (TGFβ receptor, GSK3β, or MEK), can enhance reprogramming efficiency, replace certain genetic factors in generating iPS cells, or maintain rat iPS cells.[304,310−313] Future efforts in this regard will likely include screens for small molecules that reprogram pancreatic exocrine cells or hepatocytes to insulin-producing β-cells, skin progenitors to various ectodermal lineages, and the like.

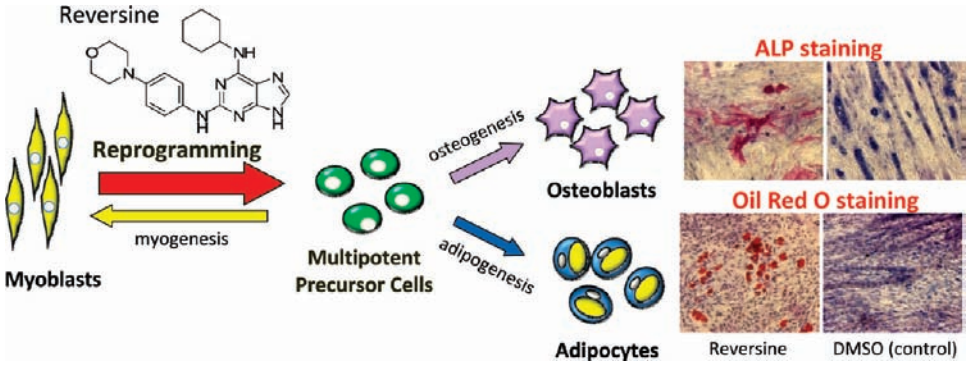


Figure 16. Reversine reprograms myoblasts (which normally differentiate into myotubes) to precursor cells that can be differentiated into osteoblasts or adipocytes. Adapted with permission from refs 305 (Copyright 2007 National Academy of Sciences, U.S.A.) and 306 (Copyright 2004 American Chemical Society).

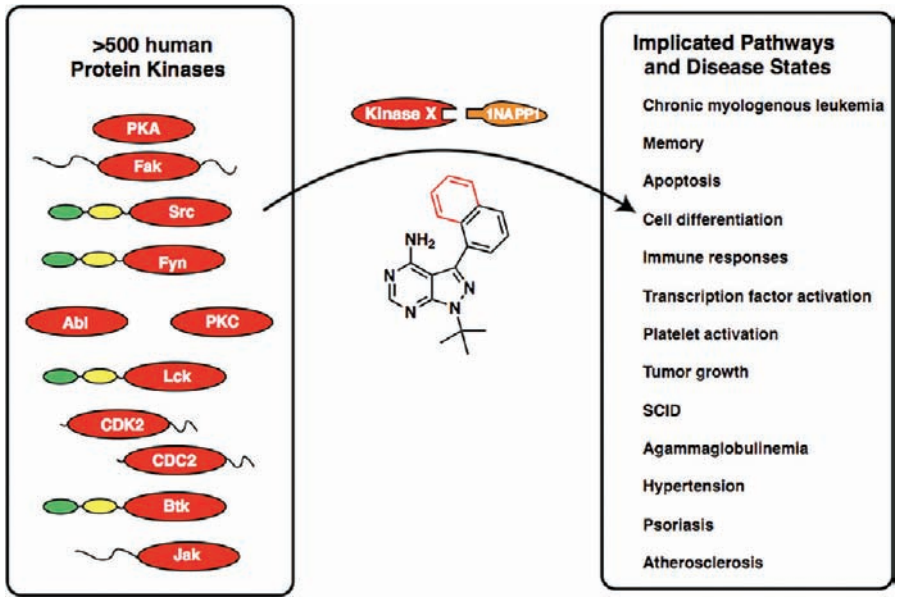


Figure 17. A kinase of interest can be genetically engineered to be selectively inhibited by an orthogonal kinase inhibitor which does not inhibit any wild-type protein kinases. The conservation of the ATP binding pocket across the kinome provides generality to the approach, as the residue that must be engineered is conserved.[33] Image courtesy of Kevan Shokat.

One can also ask whether new tissues can be regenerated simply by reversibly controlling cellular proliferation of differentiated cell types — after all, the liver regenerates after partial hepatectomy by simple division of existing hepatocytes, and pancreatic β-cells proliferate to meet metabolic demand during pregnancy or after partial pancreatectomy.[314] To this end, we have carried out a cell-based screen of murine pancreatic β-cells (reversibly immortalized with large-T antigen to generate the large number of cells required for the screen) for molecules that reversibly proliferate β-cells and maintain their ability to produce insulin. We identified a number of known compounds, including Wnt agonists and L-type calcium channel agonists, as well as novel compounds, which allow us to reversibly expand rodent β-cells (and more recently molecules that expand primary human islets) that are currently being tested in *in vivo* rodent models of type I diabetes.[315] Others have demonstrated that the Wnt signaling agonist BIO also promotes proliferation of cardiomyocytes.[316] Thus, the controlled and reversible cellular proliferation of terminally differentiated cells which are normally growth-arrested (e.g., neurons, cardiomyocytes, hepatocytes, kidney epithelial cells, etc.) may possibly offer an *ex vivo* (or even *in vivo*) alternative to stem cells as a replacement for lost tissue, especially for tissues where no adult stem cells have yet been identified. Clearly the identification of molecules that control cell fate is providing new insight into the complex factors that regulate stem cell biology but, equally importantly, may form the basis for a whole new therapeutic paradigm for synthetic molecules in medicine.

Additional Examples of Chemical Libraries in Biology and Medicine. Another exciting opportunity for the academic community to exploit chemical libraries and screening technologies that is not generally competitive with pharmaceutical or biotechnology research interests is in the area of orphan and neglected diseases. For example, there exist both a large research opportunity and a major unmet medical need with respect to molecules that kill persistent *Mycobacterium tuberculosis* (the biology of persistors is largely unknown), or molecules that target nonessential host factors that are required for viral replication (HIV, HCV, Dengue, etc.) but which will not mutate rapidly. In addition, there are a large number of orphan diseases (type I diabetes, muscular dystrophies, spinal muscular atrophy, childhood cancers, Rett syndrome, Fragile X, Huntington disease, etc.) for which no good treatments exist. The identification of molecules that modulate these disease processes may ultimately lead to new therapies as well as provide new insight into the novel biology of many of these diseases, including

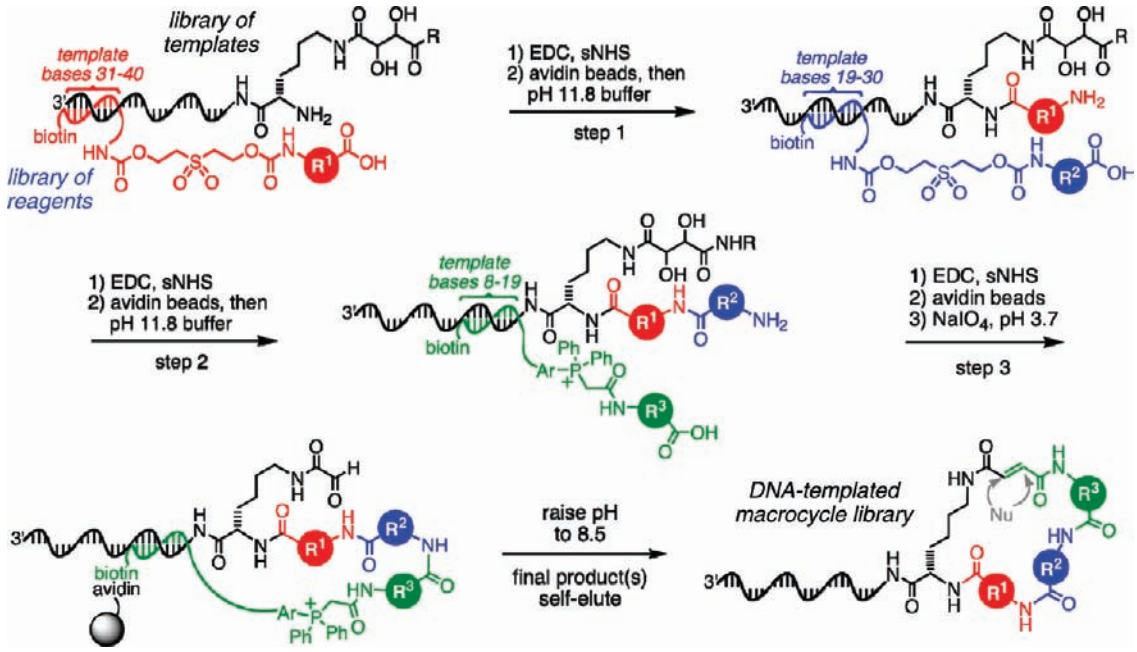


Figure 18. Template-directed synthesis of libraries of macrocycles.[328] Image courtesy of David Liu. Adapted with permission from ref 22. Copyright 2004 American Association for the Advancement of Science.

epigenetic DNA modifications, alternative splicing, protein translocations, protein misfolding, and the like. Exciting examples include the identification of compounds that selectively inhibit cell death in neuronal cells expressing the mutant huntingtin gene,[317,318] the isolation of molecules that affect snRNP assembly by causing intramolecular disulfide cross-links of SMN,[319] and the identification of molecules that modulate protein misfolding in transthyretin amyloidoses.[320] In our own laboratory, we have recently identified molecules that affect SMN2 splicing, fetal γ-hemoglobin expression, and the activity of translocated kinases in childhood cancers.[321]

Chemical libraries have many other applications in chemistry, biology, and medicine. They can be used to find molecules to explore the physiological effects of new biology—such as the roles of autophagy in neurodegenerative diseases[322] or sirtuins in aging.[323] Chemical libraries have also been used in synthetic lethal screens, synergy screens to identify interacting pathways and regulatory mechanisms, and multicellular organism screens for developmental defects[286] or effects on longevity[324] (although the latter require relatively large amounts of compound). They have also afforded molecules that complement engineered mutations in proteins, including transcription factors and growth factor receptors.[325,326] One can rationally design such molecules as well. For example, in a beautiful series of experiments, Shokat and co-workers mutated kinase active sites to introduce "holes" (e.g., Thr338Gly in Src kinase) that allow them to bind synthetic inhibitors and substrates containing "bumps" (side chains) not recognized by wild-type kinases (Figure 17).[33] This strategy, in which one combines a genetic mutation with a synthetic mutation, allows one to selectively inhibit the function of one member of a structurally homologous family of kinases to determine its cellular function and thereby avoid the arduous task of synthesizing highly selective kinase inhibitors. A similar approach has been applied to natural mutations in nuclear receptors[327] and can likely be applied to other enzymes and receptors, such as lipid kinases and methyltransferases.

A number of laboratories, including those of Liu, Harbury, and Pederson, have created libraries of small molecules by templated organic synthesis in which a DNA template is used to direct the stepwise synthesis of molecules using duplex formation to create high effective molarities of reactants (Figure 18).[22,23,328,329] This approach is being used to search for selective ligands to biological receptors, as well as to identify novel chemical reactions, and may one day lead to synthetic libraries which can be encoded, amplified, chemically mutated, and selected in a process that mimics natural protein evolution. Diversity-based approaches are also being applied to the synthesis of libraries of carbohydrates[330] and larger biomolecules.[331−340] For example, sequence-defined libraries of oligonucleotide, polypeptide, carbamate, and peptoid probes are being chemically fabricated (in some cases using a combination of solid-phase synthesis and photolithographic methods) and used to interrogate mRNA expression on a genome-wide level and probe single nucleotide polymorphisms in genetic studies,[335] identify selective modulators of receptors and enzymes,[337−340] or define substrates for phosphatases and other protein interfaces.[336] In addition, self-organizing dynamic libraries of ligands are being used to assemble supramolecular structures with defined properties.[341] Finally, the use of modern mass spectrometric and separation tools together with highly sensitive phenotypic or reporter-based cellular screens opens the opportunity to reexamine the library of natural small-molecule metabolites derived from mammalian cells or the microbiome to identify molecules that affect cellular processes such as stem cell self-renewal or differentiation or bind and activate orphan receptors[342,343] (including, for example, endogenous thyroid hormone derivatives with unusual activities[344]).

Conclusion

Chemistry continues to evolve from its historical focus on molecular structure, reactivity, and synthesis to take on the challenge of making small and large molecules and even systems of molecules with tailored properties and functions. This requires improved theoretical and analytical tools, as well as innovative

new synthetic strategies. Given the remarkable array of functions found in biological molecules, Mother Nature offers help in this regard through an approach to synthesis that seamlessly interfaces biology and chemistry. However, to fully exploit this opportunity, chemists must become highly proficient in the tools and concepts of modern biology without sacrificing their traditional understanding of molecular structure and reactivity. This represents a major challenge for the traditional chemistry education, from secondary education through postdoctoral studies. It will require increased scientific partnerships between the disciplines of chemistry, biology, and medicine in both education and research, and a reevaluation of the most important concepts and materials that we teach at the undergraduate and graduate levels. At the same time, we can inspire future generations of chemists with the exciting and highly relevant opportunities that exist at the interface of the chemical and biological sciences.

Acknowledgment. This work was supported by the National Institutes of Health, Department of Energy, Skaggs Institute for Chemical Biology, Juvenile Diabetes Research Foundation, California Institute for Regenerative Medicine, and Novartis Research Foundation. P.G.S. would also like to thank the remarkable group of co-workers he has had the opportunity to work with over the years — who themselves are having a major impact on the direction of biological chemistry.

Note Added after ASAP Publication. The version of this paper published on August 18, 2009, had an error in ref 163. The corrected version was published on September 2, 2009.

Supporting Information Available: Complete refs 145, 234, 245, 248, 258, 261, 281, 303, and 309. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Lehn, J. M. *Chem. Soc. Rev.* **2007**, *36*, 151–160.
(2) Breslow, R.; Dong, S. D. *Chem. Rev.* **1998**, *98*, 1997–2012.
(3) Venkateswara Rao, P.; Holm, R. H. *Chem. Rev.* **2004**, *104*, 527–559.
(4) Collman, J. P.; Boulatov, R.; Sunderland, C. J.; Fu, L. *Chem. Rev.* **2004**, *104*, 561–588.
(5) Schultz, P. G.; Dervan, P. B. *J. Am. Chem. Soc.* **1983**, *105*, 7748–7750.
(6) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468–471.
(7) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science* **1994**, *266*, 646–650.
(8) Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533–543.
(9) Kaiser, E. T. *Ann. N.Y. Acad. Sci.* **1987**, *501*, 14–20.
(10) (a) Erlanger, B. F. *Annu. Rev. Biochem.* **1976**, *45*, 257–283. (b) Qi, D.; Tann, C. M.; Haring, D.; Distafano, M. D. *Chem. Rev.* **2001**, *101*, 3081–3111.
(11) Orgel, L. E. *Acc. Chem. Res.* **1995**, *28*, 109–118.
(12) Xie, J.; Schultz, P. G. *Nat. Rev. Mol. Cell. Biol.* **2006**, *7*, 775–782.
(13) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621–633.
(14) Dougherty, D. A. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645–652.
(15) Link, A. J.; Mock, M. L.; Tirrell, D. A. *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609.
(16) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936–943.
(17) Seo, Y. J.; Hwang, G. T.; Ordoukhanian, P.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 3246–3252.
(18) Tae, E. L.; Wu, Y.; Xia, G.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 7439–7440.
(19) Benner, S. A. *Science* **2004**, *306*, 625–626.
(20) Alivisatos, A. P.; Johnsson, K. P.; Peng, X.; Wilson, T. E.; Loweth, C. J.; Bruchez, M. P., Jr.; Schultz, P. G. *Nature* **1996**, *382*, 609–611.
(21) Rosi, N. L.; Mirkin, C. A. *Chem. Rev.* **2005**, *105*, 1547–1562.
(22) Gartner, Z. J.; Tse, B. N.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R. *Science* **2004**, *305*, 1601–1605.
(23) Wrenn, S. J.; Weisinger, R. M.; Halpin, D. R.; Harbury, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 13137–13143.
(24) Seeman, N. C. *Nature* **2003**, *421*, 427–431.
(25) Liu, D. R.; Schultz, P. G. *Angew. Chem., Int. Ed.* **1999**, *38*, 36–54.
(26) Schultz, P. G.; Lerner, R. A. *Science* **1995**, *269*, 1835–1842.
(27) Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391–410.
(28) Bunin, B. A.; Ellman, J. A. *J. Am. Chem. Soc.* **1992**, *114*, 10997–10998.
(29) Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233–1251.
(30) Feinberg, A. W.; Feigel, A.; Shevkoplyas, S. S.; Sheehy, S.; Whitesides, G. M.; Parker, K. K. *Science* **2007**, *317*, 1366–1370.
(31) Hanson, S.; Best, M.; Bryan, M. C.; Wong, C. H. *Trends Biochem. Sci.* **2004**, *29*, 656–663.
(32) Wong, C. H.; Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*; Pergamon: Oxford, 1994.
(33) 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.; Shokat, K. M. *Nature* **2000**, *407*, 395–401.
(34) Khosla, C.; Harbury, P. B. *Nature* **2001**, *409*, 247–252.
(35) Khosla, C.; Gokhale, R. S.; Jacobsen, J. R.; Cane, D. E. *Annu. Rev. Biochem.* **1999**, *68*, 219–253.
(36) Kalaitzis, J. A.; Izumikawa, M.; Xiang, L.; Hertweck, C.; Moore, B. S. *J. Am. Chem. Soc.* **2003**, *125*, 9290–9291.
(37) Thomas, M.; Bixby, K.; Shen, B. *Anticancer Agents Nat. Prod.* **2005**, 519–551.
(38) Walsh, C. T. *Science* **2004**, *303*, 1805–1810.
(39) Fischbach, M. A.; Lai, J. R.; Roche, E. D.; Walsh, C. T.; Liu, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11951–11956.
(40) Cooper, L. E.; McClerren, A. L.; Chary, A.; van der Donk, W. A. *Chem. Biol.* **2008**, *15*, 1035.
(41) Tavassoli, A.; Benkovic, S. *Angew. Chem., Int. Ed.* **2005**, *44*, 2760–2763.
(42) Cane, D. E.; Walsh, C. T.; Khosla, C. *Science* **1998**, *282*, 63–68.
(43) Walsh, C. T. *Acc. Chem. Res.* **2008**, *41*, 4–10.
(44) Blanchard, S.; Thorson, J. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 263–271.
(45) Thibodeaux, C. J.; Melancon, C. E., III; Liu, H. W. *Angew. Chem., Int. Ed.* **2008**, *47*, 9814–9859.
(46) Atsumi, S.; Hanai, T.; Liao, J. C. *Nature* **2008**, *451*, 86–89.
(47) Stephanopoulos, G. *Science* **2007**, *315*, 801–804.
(48) Czarnik, A.; Leonard, N. *J. Am. Chem. Soc.* **1982**, *104*, 2624–2631.
(49) Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555.
(50) Kimoto, M.; Kawai, R.; Mitsuit, T.; Yokoyama, S.; Hirao, I. *Nucleic Acids Res.* **2009**, *37*, e14.
(51) Havermann, S. A.; Hoshika, S.; Hutter, D.; Benner, S. A. *Nucleosides Nucleotides Nucleic Acids* **2008**, *27*, 261–278.
(52) Meggers, E.; Holland, P. L.; Tolman, W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 10714–10715.
(53) Guckian, K. M.; Morales, J. C.; Kool, E. T. *J. Org. Chem.* **1998**, *63*, 9652–9656.
(54) Scopes, D. I. C.; Barrio, J. R.; Leonard, N. J. *Science* **1977**, *195*, 296–298.
(55) Liu, H.; Gao, J.; Maynard, L.; Saito, Y. D.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 1102–1109.
(56) Summerton, J.; Weller, D. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 187–195.
(57) Nielsen, P. E. *Mol. Biotechnol.* **2004**, *26*, 233–248.
(58) Eschenmoser, A. *Science* **1999**, *284*, 2118–2124.
(59) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. *Science* **1993**, *262*, 1401–1407.
(60) Hodges, R. S.; Saund, A. K.; Chong, P. C.; St-Pierre, S. A.; Reid, R. E. *J. Biol. Chem.* **1981**, *256*, 1214–1224.
(61) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976–978.
(62) Wade, H.; Stayrook, S. E.; Degrado, W. F. *Angew. Chem., Int. Ed.* **2006**, *45*, 4951–4954.
(63) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364–1368.
(64) Fernandez-Lopez, S.; Kim, H. S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, A.; Kraehenbuehl, K.; Long, G.; Weinberger, D. A.; Wilcoxen, K. M.; Ghadiri, M. R. *Nature* **2001**, *412*, 452–455.
(65) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219–3232.
(66) Kirshenbaum, K.; Zuckermann, R. N.; Dill, K. A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 530–535.
(67) Petersson, E. J.; Craig, C. J.; Daniels, D. S.; Qiu, J. X.; Schepartz, A. *J. Am. Chem. Soc.* **2007**, *129*, 5344–5345.

(68) Seebach, D.; Beck, A. K.; Bierbaum, D. J. *Chem. Biodivers.* **2004**, *1*, 1111–1239.
(69) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
(70) Khorana, H. G. *Science* **1979**, *203*, 614–625.
(71) Caruthers, M. H. In *Synthesis and Applications of DNA and RNA*; Narang, S. A., Ed.; Academic Press: Orlando, FL, 1987; pp 47−94.
(72) Prescher, J. A.; Bertozzi, C. R. *Cell* **2006**, *126*, 851–854.
(73) Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13132–13137.
(74) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86–89.
(75) Yin, J.; Lin, A. J.; Buckett, P. D.; Wessling-Resnick, M.; Golan, D. E.; Walsh, C. T. *Chem. Biol.* **2005**, *12*, 999–1006.
(76) Fernandez-Suarez, M.; Baruah, H.; Martinez-Hernandez, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol.* **2007**, *25*, 1483–1487.
(77) Flavell, R. R.; Muir, T. W. *Acc. Chem. Res.* **2009**, *42*, 107–116.
(78) Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
(79) Loving, G.; Imperiali, B. *J. Am. Chem. Soc.* **2008**, *130*, 13630–13638.
(80) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1989**, *111*, 8013–8014.
(81) Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Hecht, S. M. *Tetrahedron* **1984**, *40*, 87–94.
(82) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357–359.
(83) Lummis, S. C.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. *Nature* **2005**, *438*, 248–252.
(84) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314–11322.
(85) Thorson, J. S.; Chapman, E.; Murphy, E. C.; Schultz, P. G.; Judice, J. K. *J. Am. Chem. Soc.* **1995**, *117*, 1157–1158.
(86) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616–6617.
(87) Malkowski, M. G.; Quartley, E.; Friedman, A. E.; Babulski, J.; Kon, Y.; Wolfley, J.; Said, M.; Luft, J. R.; Phizicky, E. M.; DeTitta, G. T.; Grayhack, E. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6678–6683.
(88) Bentin, T.; Hamzavi, R.; Salomonsson, J.; Roy, H.; Ibba, M.; Nielsen, P. E. *J. Biol. Chem.* **2004**, *279*, 19839–19845.
(89) Suchanek, M.; Radzikowska, A.; Thiele, C. *Nat. Methods* **2005**, *2*, 261–267.
(90) Furter, R. *Protein Sci.* **1998**, *7*, 419–426.
(91) Doring, V.; Mootz, H. D.; Nangle, L. A.; Hendrickson, T. L.; de Crecy-Lagard, V.; Schimmel, P.; Marliere, P. *Science* **2001**, *292*, 501–504.
(92) Hendrickson, W. A.; Ogata, C. M. *Methods Enzymol.* **1997**, *276*, 494–523.
(93) Strable, E.; Prasuhn, D. E., Jr.; Udit, A. K.; Brown, S.; Link, A. J.; Ngo, J. T.; Lander, G.; Quispe, J.; Potter, C. S.; Carragher, B.; Tirrell, D. A.; Finn, M. G. *Bioconjugate Chem.* **2008**, *19*, 866–875.
(94) Carrico, I. S.; Maskarinec, S. A.; Heilshorn, S. C.; Mock, M. L.; Liu, J. C.; Nowatzki, P. J.; Franck, C.; Ravichandran, G.; Tirrell, D. A. *J. Am. Chem. Soc.* **2007**, *129*, 4874–4875.
(95) Beatty, K. E.; Xie, F.; Wang, Q.; Tirrell, D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14150–14151.
(96) Steitz, T. A. *Nat. Rev. Mol. Cell. Biol.* **2008**, *9*, 242–253.
(97) Yonath, A. *Mol. Cell* **2005**, *20*, 1–16.
(98) Fechter, P.; Rudinger-Thirion, J.; Tukalo, M.; Giege, R. *Eur. J. Biochem.* **2001**, *268*, 761–767.
(99) Giege, R.; Sissler, M.; Florentz, C. *Nucleic Acids Res.* **1998**, *26*, 5017–5035.
(100) Wang, L.; Schultz, P. G. *Chem. Biol.* **2001**, *8*, 883–890.
(101) Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Soll, D.; Nureki, O. *Nature* **2009**, *457*, 1163–1167.
(102) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Chem. Biol.* **2008**, *4*, 232–234.
(103) Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. *Chem. Biol.* **2008**, *15*, 1187–1197.
(104) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498–500.
(105) Melancon, C. E., III; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2009**, doi: 10.1016/j.bmcl.2009.04.007.
(106) Anderson, J. C.; Wu, N.; Santoro, S. W.; Lakshman, V.; King, D. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7566–7571.
(107) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 4052–4055.
(108) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964–967.
(109) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818–822.
(110) Young, T. S.; Ahmad, I.; Brock, A.; Schultz, P. G. *Biochemistry* **2009**, *48*, 2643–2653.
(111) Chen, S.; Schultz, P. G.; Brock, A. *J. Mol. Biol.* **2007**, *371*, 112–122.
(112) Cropp, T. A.; Anderson, J. C.; Chin, J. W. *Nat. Protoc.* **2007**, *2*, 2590–2600.
(113) Liu, W.; Brock, A.; Chen, S.; Schultz, P. G. *Nat. Methods* **2007**, *4*, 239–244.
(114) Ryu, Y. H.; Schultz, P. G. *Nat. Methods* **2006**, *3*, 263–265.
(115) Wang, Q.; Wang, L. *J. Am. Chem. Soc.* **2008**, *130*, 6066–6067.
(116) Wang, W.; Takimoto, J. K.; Louie, G. V.; Baiga, T. J.; Noel, J. P.; Lee, K. F.; Slesinger, P. A.; Wang, L. *Nat. Neurosci.* **2007**, *10*, 1063–1072.
(117) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–249.
(118) Lee, H. S.; Spraggon, G.; Schultz, P. G.; Wang, F. *J. Am. Chem. Soc.* **2009**, *131*, 2481–2483.
(119) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. *Nat. Biotechnol.* **2004**, *22*, 1297–1301.
(120) Brustad, E.; Bushey, M. L.; Brock, A.; Chittuluru, J.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6004–6006.
(121) Wang, L.; Brock, A.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 1836–1837.
(122) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.
(123) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.
(124) Chin, J. W.; Schultz, P. G. *ChemBioChem* **2002**, *3*, 1135–1137.
(125) Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A. V.; Schultz, P. G. *ChemBioChem* **2007**, *8*, 2210–2214.
(126) Lee, H. S.; Dimla, R.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5222–5224.
(127) Zeng, H.; Xie, J.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5356–5359.
(128) Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. *J. Am. Chem. Soc.* **2008**, *130*, 17664–17665.
(129) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 8220–8223.
(130) Deiters, A.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1521–1524.
(131) Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 6399–6401.
(132) Tsao, M. L.; Tian, F.; Schultz, P. G. *Chembiochem* **2005**, *6*, 2147–2149.
(133) Wang, J.; Schiller, S. M.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2007**, *46*, 6849–6851.
(134) Alfonta, L.; Zhang, Z.; Uryu, S.; Loo, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 14662–14663.
(135) Seyedsayamdost, M. R.; Xie, J.; Chan, C. T.; Schultz, P. G.; Stubbe, J. *J. Am. Chem. Soc.* **2007**, *129*, 15060–15071.
(136) Tippmann, E. M.; Schultz, P. G. *Tetrahedron* **2007**, *63*, 6182–6184.
(137) Bose, M.; Groff, D.; Xie, J.; Brustad, E.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 388–389.
(138) Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2006**, *45*, 2728–2731.
(139) Lemke, E. A.; Summerer, D.; Geierstanger, B. H.; Brittain, S. M.; Schultz, P. G. *Nat. Chem. Biol.* **2007**, *3*, 769–772.
(140) Peters, F. B.; Brock, A.; Wang, J.; Schultz, P. G. *Chem. Biol.* **2009**, *16*, 148–152.
(141) Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. *J. Am. Chem. Soc.* **2004**, *126*, 14306–14307.
(142) Xie, J.; Liu, W.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2007**, *46*, 9239–9242.
(143) Ye, S.; Kohrer, C.; Huber, T.; Kazmi, M.; Sachdev, P.; Yan, E. C.; Bhagat, A.; RajBhandary, U. L.; Sakmar, T. P. *J. Biol. Chem.* **2008**, *283*, 1525–1533.
(144) Deiters, A.; Geierstanger, B. H.; Schultz, P. G. *ChemBioChem* **2005**, *6*, 55–58.
(145) Cellitti, S. E.; et al. *J. Am. Chem. Soc.* **2008**, *130*, 9268–9281.
(146) Groff, D.; Thielges, M. C.; Cellitti, S. E.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2009**, *48*, 3478–881.
(147) Schultz, K. C.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 13984–13985.
(148) Wang, J.; Xie, J.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 8738–8739.
(149) Tsao, M. L.; Summerer, D.; Ryu, Y.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 4572–4573.
(150) Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. *J. Org. Chem.* **2003**, *68*, 174–176.
(151) Jackson, J. C.; Hammill, J. T.; Mehl, R. A. *J. Am. Chem. Soc.* **2007**, *129*, 1160–1166.

(152) Guo, J.; Wang, J.; Anderson, J. C.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 722–725.
(153) Xie, J.; Supekova, L.; Schultz, P. G. *ACS Chem. Biol.* **2007**, *2*, 474–478.
(154) Liu, C. C.; Brustad, E.; Liu, W.; Schultz, P. G. *J. Am. Chem. Soc.* **2007**, *129*, 10648–10649.
(155) Turner, J. M.; Graziano, J.; Spraggon, G.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6483–6488.
(156) Zhang, Y.; Wang, L.; Schultz, P. G.; Wilson, I. A. *Protein Sci.* **2005**, *14*, 1340–1349.
(157) Deiters, A.; Cropp, T. A.; Summerer, D.; Mukherji, M.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5743–5745.
(158) DeSantis, G.; Jones, J. B. *Curr. Opin. Biotechnol.* **1999**, *10*, 324–330.
(159) Grünewald, J.; Hunt, G. S.; Dong, L.; Niessen, F.; Wen, B. G.; Tsao, M. L.; Perera, R.; Kang, M.; Laffitte, B. A.; Azarian, S.; Ruf, W.; Nasoff, M.; Lerner, R. A.; Schultz, P. G.; Smider, V. V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4337–4342.
(160) Neumann, H.; Hazen, J. L.; Weinstein, J.; Mehl, R. A.; Chin, J. W. *J. Am. Chem. Soc.* **2008**, *130*, 4028–4033.
(161) Doyle, H. A.; Mamula, M. J. *Trends Immunol* **2001**, *22*, 443–449.
(162) Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J. W.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9785–9789.
(163) Lee, H. S.; Guo, J.; Lemke, E.; Dimla, R.; Schultz, P. G. *J. Am. Chem. Soc.* In press.
(164) Lee, H. S.; Schultz, P. G. *J. Am. Chem. Soc.* **2008**, *130*, 13194–13195.
(165) Chen, H. T.; Warfield, L.; Hahn, S. *Nat. Struct. Mol. Biol.* **2007**, *14*, 696–703.
(166) Hino, N.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. *Nat. Protoc.* **2006**, *1*, 2957–2962.
(167) Weibezahn, J.; Tessarz, P.; Schlieker, C.; Zahn, R.; Maglica, Z.; Lee, S.; Zentgraf, H.; Weber-Ban, E. U.; Dougan, D. A.; Tsai, F. T.; Mogk, A.; Bukau, B. *Cell* **2004**, *119*, 653–665.
(168) Liu, C. C.; Mack, A. V.; Tsao, M. L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17688–17693.
(169) Liu, C. C.; Mack, A. V.; Brustad, E. M.; Mills, J. H.; Groff, D.; Smider, V. V.; Schultz, P. G. *J. Am. Chem. Soc.* In press.
(170) Rackham, O.; Chin, J. W. *Nat. Chem. Biol.* **2005**, *1*, 159–166.
(171) Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 935–939.
(172) Burnet, F. M. *The Clonal Selection Theory of Acquired Immunity*; Vanderbilt University Press: Nashville, TN, 1959.
(173) Pollack, S. J.; Jacobs, J. W.; Schultz, P. G. *Science* **1986**, *234*, 1570–1573.
(174) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566–1570.
(175) Schultz, P. G.; Lerner, R. A. *Nature* **2002**, *418*, 485.
(176) Schultz, P. G.; Yin, J.; Lerner, R. A. *Angew. Chem., Int. Ed.* **2002**, *41*, 4427–4437.
(177) Pauling, L. *Nature* **1948**, *161*, 707–709.
(178) Jacobsen, J. R.; Prudent, J. R.; Kochersperger, L.; Yonkovich, S.; Schultz, P. G. *Science* **1992**, *256*, 365–367.
(179) Hilvert, D.; Carpenter, S. H.; Nared, K. D.; Auditor, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4953–4955.
(180) Barbas, C. F., III; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Bjornestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A. *Science* **1997**, *278*, 2085–2092.
(181) List, B.; Lerner, R. A.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2000**, *122*, 2395–2396.
(182) Lelais, G.; MacMillan, D. W. C. *Aldrichim. Acta* **2006**, *39*, 79–87.
(183) Schultz, P. G.; Lerner, R. A. *Acc. Chem. Res.* **1993**, *26*, 391–395.
(184) Cochran, A. G.; Schultz, P. G. *Science* **1990**, *249*, 781–783.
(185) Yin, J.; Andryski, S. E.; Beuscher, A. E.; Stevens, R. C.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 856–861.
(186) Haldane, J. D. S. *Enzymes*; Longmans Green: London, 1930.
(187) Patten, P. A.; Gray, N. S.; Yang, P. L.; Marks, C. B.; Wedemayer, G. J.; Boniface, J. J.; Stevens, R. C.; Schultz, P. G. *Science* **1996**, *271*, 1086–1091.
(188) Yin, J.; Beuscher, A. E., IV; Andryski, S. E.; Stevens, R. C.; Schultz, P. G. *J. Mol. Biol.* **2003**, *330*, 651–656.
(189) Zimmermann, J.; Oakman, E. L.; Thorpe, I. F.; Shi, X.; Abbyad, P.; Brooks, C. L., III; Boxer, S. G.; Romesberg, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13722–13727.
(190) Breinl, F.; Haurowitz, F. Z. *Physiol. Chem.* **1930**, *192*, 45–57.
(191) Pauling, L. *J. Am. Chem. Soc.* **1940**, *62*, 2643.
(192) Foote, J.; Milstein, C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10370–10374.
(193) Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2003**, *301*, 1196–1202.
(194) Cwirla, S. E.; Peters, E. A.; Barrett, R. W.; Dower, W. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6378–6382.
(195) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. *Science* **1990**, *249*, 404–406.
(196) Wrighton, N. C.; Farrell, F. X.; Chang, R.; Kashyap, A. K.; Barbone, F. P.; Mulcahy, L. S.; Johnson, D. L.; Barrett, R. W.; Jolliffe, L. K.; Dower, W. J. *Science* **1996**, *273*, 458–464.
(197) Choo, Y.; Sanchez-Garcia, I.; Klug, A. *Nature* **1994**, *372*, 642–645.
(198) Greisman, H. A.; Pabo, C. O. *Science* **1997**, *275*, 657–661.
(199) Wu, H.; Yang, W. P.; Barbas, C. F., III. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 344–348.
(200) Dickerson, M. B.; Sandhage, K. H.; Naik, R. R. *Chem. Rev.* **2008**, *108*, 4935–4978.
(201) Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. *Nature* **2000**, *405*, 665–668.
(202) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
(203) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
(204) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611–647.
(205) Joyce, G. F. *Annu. Rev. Biochem.* **2004**, *73*, 791–836.
(206) Unrau, P. J.; Bartel, D. P. *Nature* **1998**, *395*, 260–263.
(207) Mandal, M.; Breaker, R. R. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 451–463.
(208) Hager, A. J.; Pollard, J. D.; Szostak, J. W. *Chem. Biol.* **1996**, *3*, 717–725.
(209) Johnston, W. K.; Unrau, P. J.; Lawrence, M. S.; Glasner, M. E.; Bartel, D. P. *Science* **2001**, *292*, 1319–1325.
(210) Lincoln, T. A.; Joyce, G. F. *Science* **2009**, *323*, 1229–1232.
(211) Stemmer, W. P. *Nature* **1994**, *370*, 389–391.
(212) Gallagher, S. S.; Cornish, V. W. *Encyclopedia for Chemical Biology*; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2008.
(213) Tracewell, C. A.; Arnold, F. H. *Curr. Opin. Chem. Biol.* **2009**, *13*, 3–9.
(214) Wang, L.; Jackson, W. C.; Steinbach, P. A.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16745–16749.
(215) Davidson, A. R.; Sauer, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2146–2150.
(216) Graziano, J. J.; Liu, W.; Perera, R.; Geierstanger, B. H.; Lesley, S. A.; Schultz, P. G. *J. Am. Chem. Soc.* **2008**, *130*, 176–185.
(217) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. *Science* **1993**, *262*, 1680–1685.
(218) Keefe, A. D.; Szostak, J. W. *Nature* **2001**, *410*, 715–718.
(219) You, L.; Cox, R. S., III; Weiss, R.; Arnold, F. H. *Nature* **2004**, *428*, 868–871.
(220) Levskaya, A.; Chevalier, A. A.; Tabor, J. J.; Simpson, Z. B.; Lavery, L. A.; Levy, M.; Davidson, E. A.; Scouras, A.; Ellington, A. D.; Marcotte, E. M.; Voigt, C. A. *Nature* **2005**, *438*, 441–442.
(221) Brooks, H.; Lebleu, B.; Vives, E. *Adv. Drug Delivery Rev.* **2005**, *57*, 559–577.
(222) Murphy, J. E.; Uno, T.; Hamer, J. D.; Cohen, F. E.; Dwarki, V.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1517–1522.
(223) (a) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13003–13008. (b) Kolonko, E. M.; Pontrello, J. K.; MMangold, S. L.; Kiessling, L. L. *J. Am. Chem. Soc.* **2009**, *131*, 7327–7333.
(224) Xiang, X. D.; Sun, X.; Briceno, G.; Lou, Y.; Wang, K. A.; Chang, H.; Wallace-Freedman, W. G.; Chen, S. W.; Schultz, P. G. *Science* **1995**, *268*, 1738–1740.
(225) Briceño, G.; Chang, H.; Sun, X.; Schultz, P. G.; Xiang, X. D. *Science* **1995**, *270*, 273–275.
(226) Wang, J.; Yoo, Y.; Gao, C.; Takeuchi, I. I.; Sun, X.; Chang, H.; Xiang, X.; Schultz, P. G. *Science* **1998**, *279*, 1712–1714.
(227) Danielson, E.; Devenney, M.; Giaquinta, D. M.; Golden, J. H.; Haushalter, R. C.; McFarland, E. W.; Poojary, D. M.; Reaves, C. M.; Weinberg, W. H.; Wu, X. D. *Science* **1998**, *279*, 837–839.
(228) Arriola, D. J.; Carnahan, E. M.; Hustad, P. D.; Kuhlman, R. L.; Wenzel, T. T. *Science* **2006**, *312*, 714–719.
(229) Maier, W. F.; Stowe, K.; Sieg, S. *Angew. Chem., Int. Ed.* **2007**, *46*, 6016–6067.
(230) Turner, H. W.; Volpe, A. F., Jr.; Weinberg, W. H. *Surf. Sci.* **2009**, *603*, 1763–1769.
(231) Rajan, K. *Annu. Rev. Mater. Res.* **2008**, *38*, 299–322.
(232) Xiang, X. D.; Takeuchi, I. *Combinatorial Materials Synthesis*; Dekker: New York, 2003.
(233) Bailey, M.; Kaye, S. Unpublished results.
(234) Lesley, S. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11664–11669.
(235) Jandeleit, B.; Schaefer, D. J.; Powers, T. S.; Turner, H. W.; Weinberg, W. H. *Angew. Chem., Int. Ed.* **1999**, *38*, 2494–2532.

(236) Lewis, G. J.; Sachtler, J. W. A.; Low, J. J.; Lesch, D. A.; Dosek, P.; Faheem, S.; Knight, L. M.; Halloran, L. *DOE Hydrogen Program Peer Review*; U.S. Department of Energy: Washington, DC, 2008.
(237) Bonakdarpour, A.; Hewitt, K. C.; Hatchard, T. D.; Fleischauer, M. D.; Dahn, J. R. *Thin Solid Films* **2003**, *440*, 11–18.
(238) Danielson, E.; Golden, J. H.; McFarland, E. W.; Reaves, C. M.; Weinberg, W. H.; Di Wu, X. *Nature* **1997**, *389*, 944–948.
(239) Thompson, L. A.; Ellman, J. A. *Chem. Rev.* **1996**, *96*, 555–600.
(240) Gordon, E. M.; Barrett, R. W.; Dower, W. J.; Fodor, S. P.; Gallop, M. A. *J. Med. Chem.* **1994**, *37*, 1385–1401.
(241) Bloch, K. *Acc. Chem. Res.* **1969**, *2*, 193–202.
(242) Tsien, R. Y. In *Calcium as a cellular regulator*; Carafoli, E., Klee, C., Eds.; Oxford University Press: New York, 1999, pp 28−54.
(243) Bauer, J. H. *Am. J. Hypertens.* **1990**, *3*, 331–337.
(244) Wong, D. T.; Perry, K. W.; Bymaster, F. P. *Nat. Rev. Drug. Discov.* **2005**, *4*, 764–774.
(245) Melnick, J. S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3153–3158.
(246) Goldstein, D. M.; Gray, N. S.; Zarrinkar, P. P. *Nat. Rev. Drug. Discov.* **2008**, *7*, 391–397.
(247) Hong, J.; Edel, J. B.; deMello, A. J. *Drug Discov. Today* **2009**, *14*, 134–146.
(248) Plouffe, D.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9059–9064.
(249) Blackwell, H. E.; Perez, L.; Stavenger, R. A.; Tallarico, J. A.; Cope Eatough, E.; Foley, M. A.; Schreiber, S. L. *Chem. Biol.* **2001**, *8*, 1167–1182.
(250) Clemons, P. A.; Koehler, A. N.; Wagner, B. K.; Sprigings, T. G.; Spring, D. R.; King, R. W.; Schreiber, S. L.; Foley, M. A. *Chem. Biol.* **2001**, *8*, 1183–1195.
(251) Arya, P.; Joseph, R.; Chou, D. T. *Chem. Biol.* **2002**, *9*, 145–156.
(252) Ding, S.; Gray, N. S.; Wu, X.; Ding, Q.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 1594–1596.
(253) Bunin, B. A.; Plunkett, M. J.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4708–4712.
(254) Bregman, H.; Williams, D. S.; Atilla, G. E.; Carroll, P. J.; Meggers, E. *J. Am. Chem. Soc.* **2004**, *126*, 13594–13595.
(255) Stockwell, B. R. *Nat. Rev. Genet.* **2000**, *1*, 116–125.
(256) Knight, Z. A.; Shokat, K. M. *Cell* **2007**, *128*, 425–430.
(257) Schreiber, S. L. *Science* **1991**, *251*, 283–287.
(258) Chanda, S. K.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12153–12158.
(259) Willingham, A. T.; Orth, A. P.; Batalov, S.; Peters, E. C.; Wen, B. G.; Aza-Blanc, P.; Hogenesch, J. B.; Schultz, P. G. *Science* **2005**, *309*, 1570–1573.
(260) Huang, Q.; Raya, A.; DeJesus, P.; Chao, S. H.; Quon, K. C.; Caldwell, J. S.; Chanda, S. K.; Izpisua-Belmonte, J. C.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3456–3461.
(261) Ong, S. E.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4617–4622.
(262) Bosch, F.; Rosich, L. *Pharmacology* **2008**, *82*, 171–179.
(263) Hardman, J. G.; Limbird, L. E.; Gilman, A. G. *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, *10th ed.*; McGraw-Hill: New York, 2001.
(264) Ding, S.; Schultz, P. G. *Nat. Biotechnol.* **2004**, *22*, 833–840.
(265) Ding, S.; Schultz, P. G. *Curr. Top. Med. Chem.* **2005**, *5*, 383–395.
(266) Fuchs, E.; Segre, J. A. *Cell* **2000**, *100*, 143–155.
(267) Weissman, I. L. *Cell* **2000**, *100*, 157–168.
(268) Murry, C. E.; Keller, G. *Cell* **2008**, *132*, 661–680.
(269) Yamanaka, S. *Cell Stem Cell* **2007**, *1*, 39–49.
(270) Zhou, Q.; Melton, D. A. *Cell Stem Cell* **2008**, *3*, 382–388.
(271) Xu, Y.; Shi, Y.; Ding, S. *Nature* **2008**, *453*, 338–344.
(272) Boitano, A.; Cooke, M.; Schultz, P. G. Presented at the 50th ASH Annual Meeting and Exposition, San Francisco, CA, Dec 6−9, 2008; American Society of Hematology: Washington, DC, 2008.
(273) Warashina, M.; Min, K. H.; Kuwabara, T.; Huynh, A.; Gage, F. H.; Schultz, P. G.; Ding, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 591–593.
(274) Wurdak, H.; Zhu, S.; Schultz, P. G. Unpublished results.
(275) Wu, X.; Walker, J.; Zhang, J.; Ding, S.; Schultz, P. G. *Chem. Biol.* **2004**, *11*, 1229–1238.
(276) Johnson, K.; Tremblay, M.; Schultz, P. G. Unpublished results.
(277) Diamandis, P.; Wildenhain, J.; Clarke, I. D.; Sacher, A. G.; Graham, J.; Bellows, D. S.; Ling, E. K.; Ward, R. J.; Jamieson, L. G.; Tyers, M.; Dirks, P. B. *Nat. Chem. Biol.* **2007**, *3*, 268–273.
(278) Sadek, H.; Hannack, B.; Choe, E.; Wang, J.; Latif, S.; Garry, M. G.; Garry, D. J.; Longgood, J.; Frantz, D. E.; Olson, E. N.; Hsieh, J.; Schneider, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6063–6068.
(279) Saxe, J. P.; Wu, H.; Kelly, T. K.; Phelps, M. E.; Sun, Y. E.; Kornblum, H. I.; Huang, J. *Chem. Biol.* **2007**, *14*, 1019–1030.
(280) Schneider, J. W.; Gao, Z.; Li, S.; Farooqi, M.; Tang, T. S.; Bezprozvanny, I.; Frantz, D. E.; Hsieh, J. *Nat. Chem. Biol.* **2008**, *4*, 408–410.
(281) Buckbinder, L.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10619–10624.
(282) Hao, J.; Daleo, M. A.; Murphy, C. K.; Yu, P. B.; Ho, J. N.; Hu, J.; Peterson, R. T.; Hatzopoulos, A. K.; Hong, C. C. *PLoS ONE* **2008**, *3*, e2904.
(283) Kawatani, M.; Okumura, H.; Honda, K.; Kanoh, N.; Muroi, M.; Dohmae, N.; Takami, M.; Kitagawa, M.; Futamura, Y.; Imoto, M.; Osada, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11691–11696.
(284) Chen, J. K.; Taipale, J.; Young, K. E.; Maiti, T.; Beachy, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14071–14076.
(285) Stanton, B. Z.; Peng, L. F.; Maloof, N.; Nakai, K.; Wang, X.; Duffner, J. L.; Taveras, K. M.; Hyman, J. M.; Lee, S. W.; Koehler, A. N.; Chen, J. K.; Fox, J. L.; Mandinova, A.; Schreiber, S. L. *Nat. Chem. Biol.* **2009**, *5*, 154–156.
(286) Yu, P. B.; Hong, C. C.; Sachidanandan, C.; Babitt, J. L.; Deng, D. Y.; Hoyng, S. A.; Lin, H. Y.; Bloch, K. D.; Peterson, R. T. *Nat. Chem. Biol.* **2008**, *4*, 33–41.
(287) Lee, J.; Wu, X.; Pasca di Magliano, M.; Peters, E. C.; Wang, Y.; Hong, J.; Hebrok, M.; Ding, S.; Cho, C. Y.; Schultz, P. G. *ChemBioChem* **2007**, *8*, 1916–1919.
(288) Zhang, Q.; Major, M. B.; Takanashi, S.; Camp, N. D.; Nishiya, N.; Peters, E. C.; Ginsberg, M. H.; Jian, X.; Randazzo, P. A.; Schultz, P. G.; Moon, R. T.; Ding, S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7444–7448.
(289) Visvader, J. E.; Lindeman, G. J. *Nat. Rev. Cancer* **2008**, *8*, 755–768.
(290) Dubrovska, A.; Kim, S.; Salamone, R. J.; Walker, J. R.; Maira, S. M.; Garcia-Echeverria, C.; Schultz, P. G.; Reddy, V. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 268–273.
(291) Chen, S.; Do, J. T.; Zhang, Q.; Yao, S.; Yan, F.; Peters, E. C.; Scholer, H. R.; Schultz, P. G.; Ding, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17266–17271.
(292) Wu, X.; Ding, S.; Ding, Q.; Gray, N. S.; Schultz, P. G. *J. Am. Chem. Soc.* **2004**, *126*, 1590–1591.
(293) Ding, S.; Wu, T. Y.; Brinker, A.; Peters, E. C.; Hur, W.; Gray, N. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7632–7637.
(294) Gattinoni, L.; Zhong, X. S.; Palmer, D. C.; Ji, Y.; Hinrichs, C. S.; Yu, Z.; Wrzesinski, C.; Boni, A.; Cassard, L.; Garvin, L. M.; Paulos, C. M.; Muranski, P.; Restifo, N. P. *Nat. Med.* **2009**, doi: 10.1182/blood-2008-12-192419.
(295) Zhu, S.; Wurdak, H.; Wang, J.; Lyssiotis, C. A.; Peters, E. C.; Cho, C.; Wu, X.; Schultz, P. G. *Cell Stem Cell* **2009**, *4*, 416–426.
(296) Borowiak, M.; Maehr, R.; Chen, S.; Chen, A. E.; Tang, W.; Fox, J. L.; Schreiber, S. L.; Melton, D. A. *Cell Stem Cell* **2009**, *4*, 348–358.
(297) Chen, S.; Borowiak, M.; Fox, J. L.; Maehr, R.; Osafune, K.; Davidow, L.; Lam, K.; Peng, L. F.; Schreiber, S. L.; Rubin, L. L.; Melton, D. *Nat. Chem. Biol.* **2009**, *5*, 258–265.
(298) Lin, H.; Lee, E.; Hestir, K.; Leo, C.; HHuang, M.; Bosch, E.; Halenbeck, R.; Wu, G.; Zhou, A.; Behrens, D.; Hollenbaugh, D.; Linnemann, T.; Qin, M.; Wong, J.; Chu, K.; Doberstein, S. K.; Williams, L. T. *Science* **2008**, *320*, 807.
(299) Gonzales, R.; Schultz, P. G. Unpublished results.
(300) Takahashi, K.; Yamanaka, S. *Cell* **2006**, *126*, 663–676.
(301) Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. *Cell* **2007**, *131*, 861–872.
(302) Park, I. H.; Zhao, R.; West, J. A.; Yabuuchi, A.; Huo, H.; Ince, T. A.; Lerou, P. H.; Lensch, M. W.; Daley, G. Q. *Nature* **2008**, *451*, 141–146.
(303) Kim, J. B.; et al. *Cell* **2009**, *136*, 411–419.
(304) Wernig, M.; Meissner, A.; Foreman, R.; Brambrink, T.; Ku, M.; Hochedlinger, K.; Bernstein, B. E.; Jaenisch, R. *Nature* **2007**, *448*, 318–324.
(305) Chen, S.; Takanashi, S.; Zhang, Q.; Xiong, W.; Zhu, S.; Peters, E. C.; Ding, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10482–10487.
(306) Chen, S.; Zhang, Q.; Wu, X.; Schultz, P. G.; Ding, S. *J. Am. Chem. Soc.* **2004**, *126*, 410–411.
(307) Lyssiotis, C. A.; Walker, J.; Wu, C.; Kondo, T.; Schultz, P. G.; Wu, X. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14982–14987.
(308) Liu, A.; Han, Y. R.; Li, J.; Sun, D.; Ouyang, M.; Plummer, M. R.; Casaccia-Bonnefil, P. *J. Neurosci.* **2007**, *27*, 7339–7343.
(309) Lyssiotis, C. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 8912–8917.
(310) Shi, Y.; Desponts, C.; Do, J. T.; Hahm, H. S.; Scholer, H. R.; Ding, S. *Cell Stem Cell* **2008**, *3*, 568–574.
(311) Huangfu, D.; Maehr, R.; Guo, W.; Eijkelenboom, A.; Snitow, M.; Chen, A. E.; Melton, D. A. *Nat. Biotechnol.* **2008**, *26*, 795–797.

(312) Li, W.; Wei, W.; Zhu, S.; Zhu, J.; Shi, Y.; Lin, T.; Hao, E.; Hayek, A.; Deng, H.; Ding, S. *Cell Stem Cell* **2009**, *4*, 16–19.
(313) Shi, Y.; Do, J. T.; Desponts, C.; Hahm, H. S.; Scholer, H. R.; Ding, S. *Cell Stem Cell* **2008**, *2*, 525–528.
(314) Zaret, K. S.; Grompe, M. *Science* **2008**, *322*, 1490–1494.
(315) Wang, W.; Walker, J. R.; Wang, X.; Tremblay, M. S.; Lee, J. W.; Wu, X.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 1427–1432.
(316) Tseng, A. S.; Engel, F. B.; Keating, M. T. *Chem. Biol.* **2006**, *13*, 957–963.
(317) Bauer, A.; Stockwell, B. *Chem. Rev.* **2008**, *108*, 1774–1786.
(318) Varma, H.; Voisine, C.; DeMarco, C. T.; Cattaneo, E.; Lo, D. C.; Hart, A. C.; Stockwell, B. R. *Nat. Chem. Biol.* **2007**, *3*, 99–100.
(319) Wan, L.; Ottinger, E.; Cho, S.; Dreyfuss, G. *Mol. Cell* **2008**, *31*, 244–254.
(320) Johnson, S. M.; Wiseman, R. L.; Sekijima, Y.; Green, N. S.; Adamski-Werner, S. L.; Kelly, J. W. *Acc. Chem. Res.* **2005**, *38*, 911–921.
(321) Galkin, A. V.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 270–275.
(322) Sarkar, S.; Perlstein, E. O.; Imarisio, S.; Pineau, S.; Cordenier, A.; Maglathlin, R. L.; Webster, J. A.; Lewis, T. A.; O'Kane, C. J.; Schreiber, S. L.; Rubinsztein, D. C. *Nat. Chem. Biol.* **2007**, *3*, 331–338.
(323) Howitz, K. T.; Bitterman, K. J.; Cohen, H. Y.; Lamming, D. W.; Lavu, S.; Wood, J. G.; Zipkin, R. E.; Chung, P.; Kisielewski, A.; Zhang, L. L.; Scherer, B.; Sinclair, D. A. *Nature* **2003**, *425*, 191–196.
(324) Petrascheck, M.; Ye, X.; Buck, L. B. *Nature* **2007**, *450*, 553–556.
(325) Guo, Z.; Zhou, D.; Schultz, P. G. *Science* **2000**, *288*, 2042–2045.
(326) Lin, Q.; Barbas, C. F., III; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 612–613.
(327) Hassan, A. Q.; Koh, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 8868–8874.
(328) Gartner, Z. J.; Liu, D. R. *J. Am. Chem. Soc.* **2001**, *123*, 6961–6963.
(329) Halpin, D. R.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, E174.
(330) Liang, R.; Yan, L.; Loebach, J.; Ge, M.; Uozumi, Y.; Sekanina, K.; Horan, N.; Gildersleeve, J.; Thompson, C.; Smith, A.; Biswas, K.; Still, W. C.; Kahne, D.*Science* **1996**, *274*, 1520–1522.
(331) Geysen, H. M.; Mason, T. J. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 397–404.
(332) Lam, K. S.; Hruby, V. J.; Lebl, M.; Knapp, R. J.; Kazmierski, W. M.; Hersh, E. M.; Salmon, S. E. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 419–424.
(333) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84–86.
(334) Ohlmeyer, M. H.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922–10926.
(335) Chee, M.; Yang, R.; Hubbell, E.; Berno, A.; Huang, X. C.; Stern, D.; Winkler, J.; Lockhart, D. J.; Morris, M. S.; Fodor, S. P. *Science* **1996**, *274*, 610–614.
(336) Garaud, M.; Pei, D. *J. Am. Chem. Soc.* **2007**, *129*, 5366–5367.
(337) Udugamasooriya, D. G.; Dineen, S. P.; Brekken, R. A.; Kodadek, T. *J. Am. Chem. Soc.* **2008**, *130*, 5744–5752.
(338) Winssinger, N.; Harris, J. L.; Backes, B. J.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2001**, *40*, 3152–3155.
(339) Cho, C. Y.; Moran, E. J.; Cherry, S. R.; Stephans, J. C.; Fodor, S. P.; Adams, C. L.; Sundaram, A.; Jacobs, J. W.; Schultz, P. G. *Science* **1993**, *261*, 1303–1305.
(340) Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J. A.; Craik, C. S. *J. Biol. Chem.* **2006**, *281*, 12824–12832.
(341) Nitschke, J. R.; Lehn, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11970–11974.
(342) Backhed, F.; Ley, R. E.; Sonnenburg, J. L.; Peterson, D. A.; Gordon, J. I. *Science* **2005**, *307*, 1915–1920.
(343) Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3698–3703.
(344) Scanlan, T. S. *Endocrinology* **2009**, *150*, 1108–1111.

JA9026067