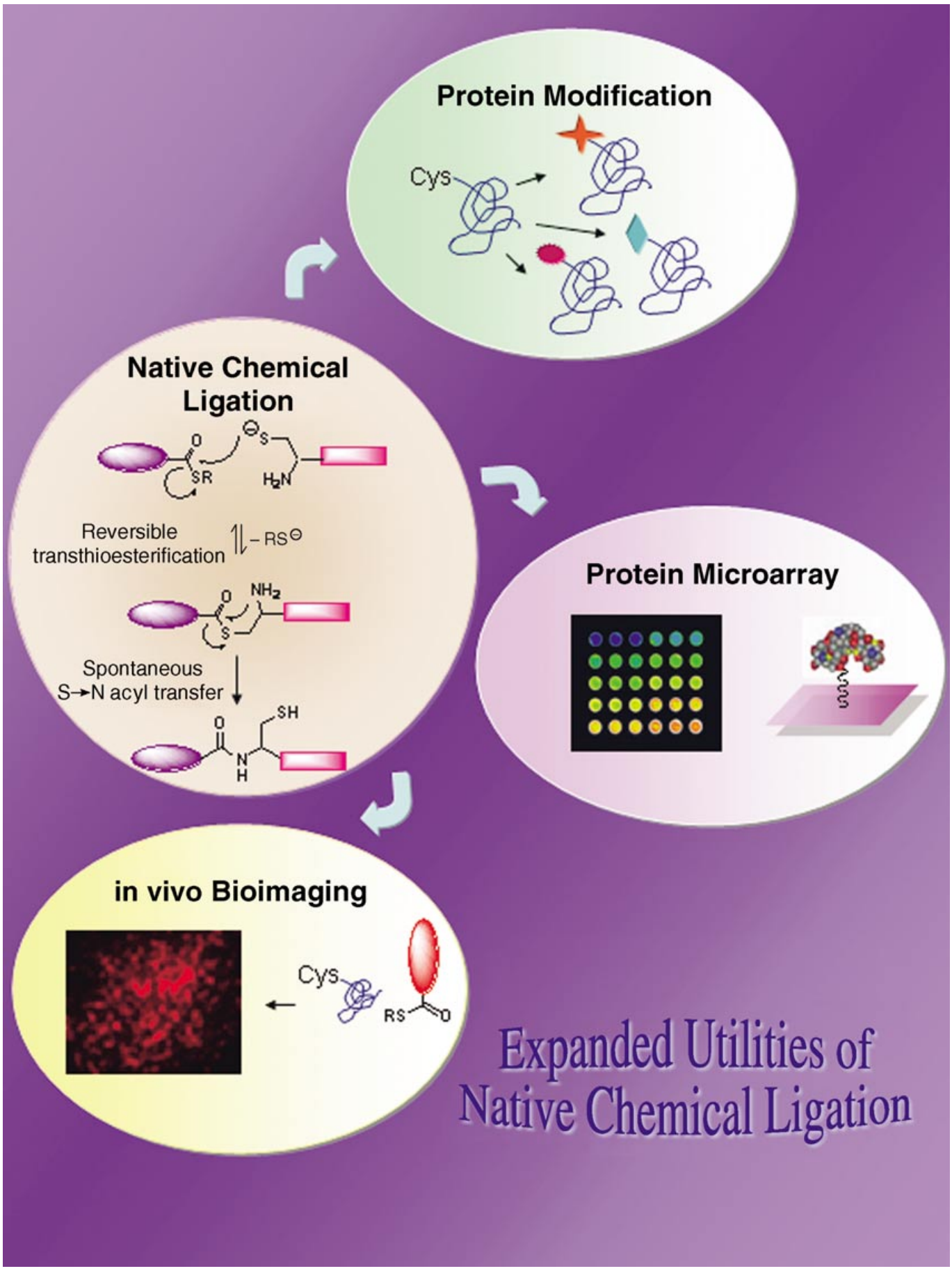


DOI: 10.1002/chem.200400414 *Chem. Eur. J.* **2004**, *10*, 4664 – 4672

Expanded Utility of the Native Chemical Ligation Reaction

Dawn S. Y. Yeo,[a] Rajavel Srinivasan,[b] Grace Y. J. Chen,[a, b] and Shao Q. Yao*[a, b]

Abstract: The post-genomic era heralds a multitude of challenges for chemists and biologists alike, with the study of protein functions at the heart of much research. The elucidation of protein structure, localization, stability, post-translational modifications, and protein interactions will steadily unveil the role of each protein and its associated biological function in the cell. The push to develop new technologies has necessitated the integration of various disciplines in science. Consequently, the role of chemistry has never been so profound in the study of biological processes. By combining the strengths of recombinant DNA technology, protein splicing, organic chemistry, and the chemoselective chemistry of native chemical ligation, various strategies have been successfully developed and applied to chemoselectively label proteins, both in vitro and in live cells, with biotin, fluorescent, and other small molecule probes. The site-specific incorporation of molecular entities with unique chemical functionalities in proteins has many potential applications in chemical and biological studies of proteins. In this article, we highlight recent progress of these strategies in several areas related to proteomics and chemical biology, namely, in vitro and in vivo protein biotinylation, protein microarray technologies for large-scale protein analysis, and live-cell bioimaging.

Keywords: bioimaging · biotinylation · fluorescence · intein · native chemical ligation · protein modifications

[a] D. S. Y. Yeo, G. Y. J. Chen, Prof. S. Q. Yao
Department of Biological Sciences
National University of Singapore
3 Science Drive 3
Singapore 117543 (Singapore)
Fax: (+65) 677-91-691
E-mail: chmyaosq@nus.edu.sg

[b] R. Srinivasan, G. Y. J. Chen, Prof. S. Q. Yao
Department of Chemistry
National University of Singapore
3 Science Drive 3
Singapore 117543 (Singapore)

Introduction

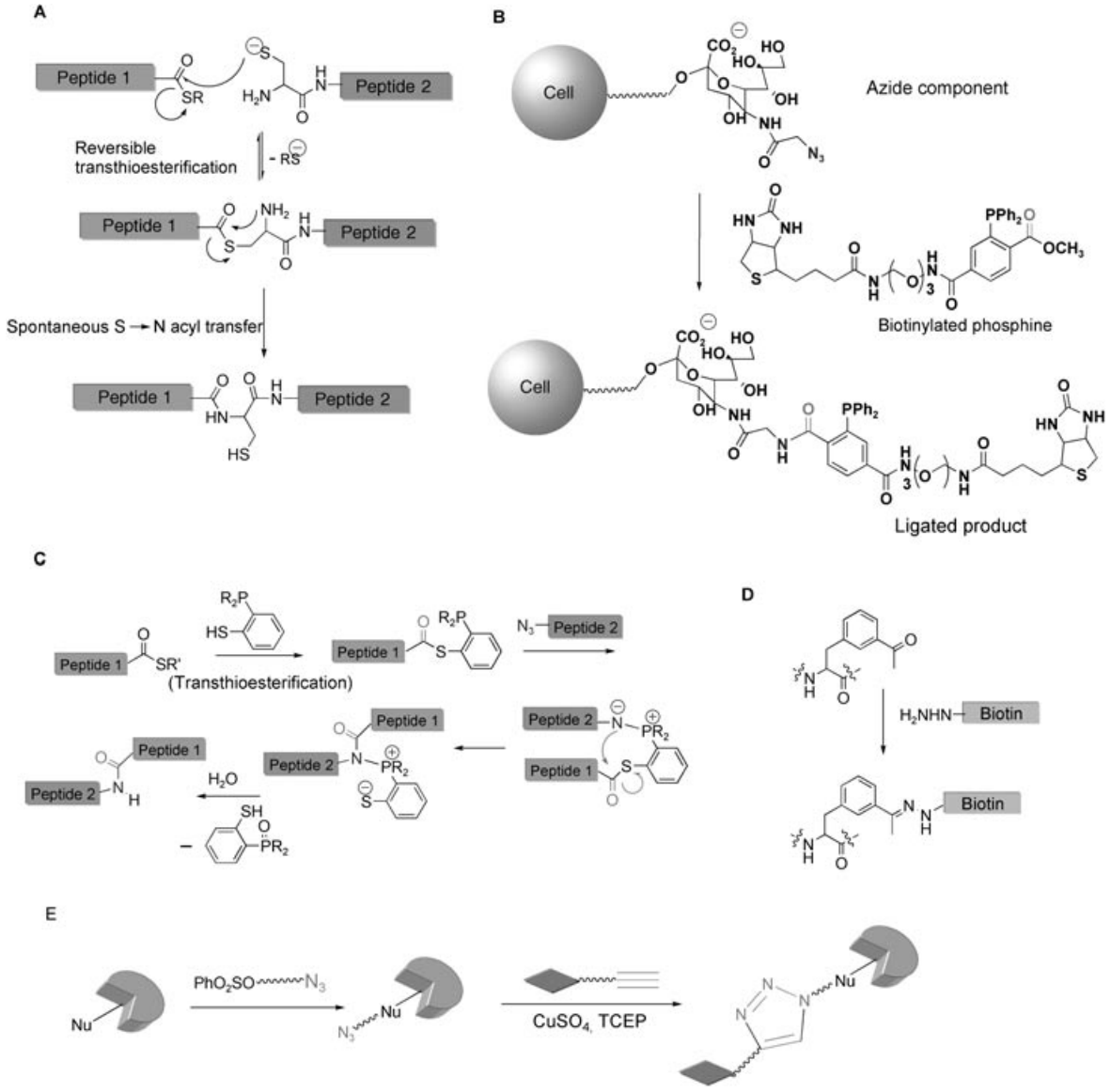
Covalent chemical reactions compatible with physiological environments and capable of achieving high selectivity have a myriad of applications in biotechnology, biomedical research, and chemical biology. Highly selective and in vivo compatible reactions have numerous biological applications, for example, for the functional assembly of complex biostructures, protein semisynthesis, and chemical targeting of biomolecules, including live cells. By chemoselectively modifying only subsets of predefined cellular components, such reactions are chemically inert toward all others and could therefore provide important information for a better understanding of numerous cellular processes. A number of important criteria need to be met in order for such reactions to work inside complex cellular environments. Firstly, such reactions have to proceed efficiently in aqueous conditions. Secondly, the two (or more) participating functional groups in the reaction should be carefully tuned, such that their reaction is highly specific and devoid of any interference from other chemical entities present in surrounding molecules (e.g., proteins, DNA/RNA, etc). Lastly, the reaction should be able to generate a product that is highly stable in its physiological environments. Very few such reactions are known to date, most of which will be briefly described in the following paragraphs. For more exhaustive coverage of some of these reactions, interested readers are referred to a number of excellent reviews.[1] Herein, we focus on one such aqueous-compatible reaction—the native chemical ligation reaction—and some of the recent advances made using this reaction from the authors' laboratory.

In Vivo Compatible Reactions

The well-established chemistry of the native chemical ligation reaction was first proposed by Wieland et al.[2a] and Brenner et al.,[2b] and subsequently demonstrated by Dawson et al. in 1994 as a general synthetic route for the semisynthesis of native proteins.[2c] The field has recently been reviewed.[1c] While many other ligation chemistries exist that result in the formation of a nonnative bond at the ligation site of the protein,[1a] the native chemical ligation is

 DOI: 10.1002/chem.200400414

one of the very few nonenzymatic reactions known that efficiently joins two unprotected peptide segments, containing appropriately installed chemical functionalities, to generate a ligated peptide/protein product with a native peptide bond at the reaction site. This highly chemoselective reaction occurs in an aqueous solution at physiological pH and involves a peptide fragment with an N-terminal cysteine residue and a second peptide fragment containing a C-terminal thioester group. The essence of the native chemical ligation reaction lies in the transthioesterification step between the thioester in one peptide and the sulfydryl group from the N-terminal cysteine residue in the other to generate a ligated thioester intermediate, which undergoes spontaneous S→N acyl rearrangement to give rise to the final ligated product containing a native peptide bond at the ligation junction (Scheme 1A). The transthioesterification reaction (i.e., the first step) is catalyzed by a suitable thiol additive (i.e., 2-mercaptoethanesulfonic acid), and is reversible under physiological conditions. The subsequent intramolecular nucleophilic attack by the α-amino group of the N-terminal cysteine to form the final amide bond is irreversible, and highly favorable due to the intramolecular five-membered ring formation and the subsequent generation of the thermodynamically stable amide bond. Consequently, all of the freely equilibrating thioester intermediates (e.g., from the first step) will eventually be depleted by the irreversible reaction in the second step, giving rise to only a single stable, ligated product. A key feature of this reaction is that it is highly chemoselective; the reaction occurs exclusively at the N-terminal cysteine of the peptide, even in the presence of other unprotected side-chain residues including internal cysteine residues.[1c,2c] The advantage of this will become evident in our later discussions. It should be noted that a number of strategies have recently been developed that have success-



Scheme 1. In vivo compatible chemical reactions. A) Principle of native chemical ligation. B) Reaction of biotinylated phosphine and azido sialic acids on the cell surface. C) Modified Staudinger ligation developed by Raines et al.[5c] D) In vivo labeling of a protein containing m-acetyl-L-phenylalanine with a hydrazide derivative. E) In vivo labeling strategy of enzymes using Click Chemistry.

fully extended the native chemical ligation reaction to accommodate noncysteine residues at the ligation site.[3]

The classical Staudinger reaction occurs between an azide and a phosphine to yield aza–ylide.[4] The aza–ylide hydrolyzes spontaneously in the presence of water to give an amine and a phosphine oxide. The reaction is specific, high yielding, water-compatible, and friendly to cellular environments. Bertozzi et al. modified the above reaction such that an amide bond was formed instead of an amine (B in Scheme 1).[5a,4] This was achieved by engineering the phosphine such that it contained an electrophilic trap, for example, a methyl ester group, next to the phosphine to capture the nucleophilic aza–ylide intermediate. The aza–ylide intermediate then underwent an intramolecular cyclization ultimately resulting in the formation of a stable amide bond. They successfully demonstrated that this modified Staudinger reaction worked efficiently on the surface of mammalian cells: by growing cells in the presence of *N*-azidoacetylmannosamine, azido groups were introduced on the cell surface through the sialic acid biosynthetic pathway.[5a,4] Upon treatment with a water-soluble biotinylated phosphine, as shown in Scheme 1B, the chemoselective reaction occurred between the phosphine and azido groups on the cell surface generating an unnatural, biotinylated cell surface on which the ligated product was joined by a stable amide bond. Recently, Bertozzi[5b] and Raines[5c] independently developed modified Staudinger reactions in which a native peptide bond was formed between a thioester- and an azide-containing peptide, in the presence of phosphinobenzenethiol or other suitable thiols (Scheme 1C). These two modified strategies based on the Staudinger reaction, although similar to the native chemical ligation reaction described earlier,[2c] are highlighted by the fact that no cysteine residue is required at the ligating site. Despite their great potentials, the two methods have yet to be extensively tested for their compatibility with the semisynthesis of proteins.[5d]

Aldehydes and ketones react selectively with hydrazides, aminooxy compounds, and thiosemicarbazides to give chemoselective products.[1a] The reaction is in vivo compatible and has been utilized by various groups including Schultz,[6a] Bertozzi,[7] and so forth. Recently, Schultz et al. developed a method for the genetic incorporation of unnatural amino acids site-specifically into proteins expressed in *E. coli* in response to the amber nonsense codon.[6b] By utilizing an orthogonal tRNA-TyrRS pair, which selectively and efficiently incorporates *m*-acetyl-L-phenylalanine into a protein overexpressed in *E. coli*, they successfully introduced a "ketone handle" into the target protein; this was subsequently modified, in living cells, by hydrazide-containing small molecules by means of ketone/hydrazide chemistry (Scheme 1D).[6a] The labeling reaction was selective and in general proceeded with relatively good yields (e.g., $>75\%$ ligation product).

Another aqueous-compatible reaction is the so-called "Click Chemistry", developed by the Sharpless group.[1b,8] The reaction occurs between an azide and a terminal alkyne in the presence of a Cu^{I} catalyst, generating the resulting 1,4-disubstituted 1,2,3-triazole in quantitative yields (Scheme 1E). Both reactants and the product are stable towards a wide range of chemicals and reaction conditions. The high specificity, compatibility with water, and high-yielding nature of this reaction makes it potentially applicable for a variety of in vivo applications. Recently, by utilizing this reaction, Cravatt et al. reported a two-step approach for the activity-based protein profiling (ABPP) of enzymes in live cells and animals.[9]

Intein-Mediated Protein Ligation Reactions and Their Expanded Applications

Protein splicing is a cellular processing event that occurs post-translationally at a polypeptide level.[10] The initial nonfunctional protein precursor undergoes a series of intramolecular reactions and rearrangements, resulting in the excision of an internal polypeptide fragment, termed the intein, and the concurrent ligation of the two flanking polypeptide sequences, namely, the exteins. The product of ligated exteins is a functionally mature protein. The biochemical mechanism of protein splicing, despite being distinct in its own ways, does share many critical features with the native chemical ligation reaction described earlier,[2c] including the formation of a (thio)ester intermediate and the final step of a S→O or S→N acyl shift to form the final amide-linked product. This extremely complex process is autocatalytic, requiring neither cofactors nor auxiliary enzymes. The elucidation of many reaction steps involved in protein splicing has made it possible to engineer modified inteins that undergo highly specific self-cleavage and protein ligation reactions. Consequently, these intein-mediated recombinant approaches are finding increasing applications in protein engineering.[11–19]

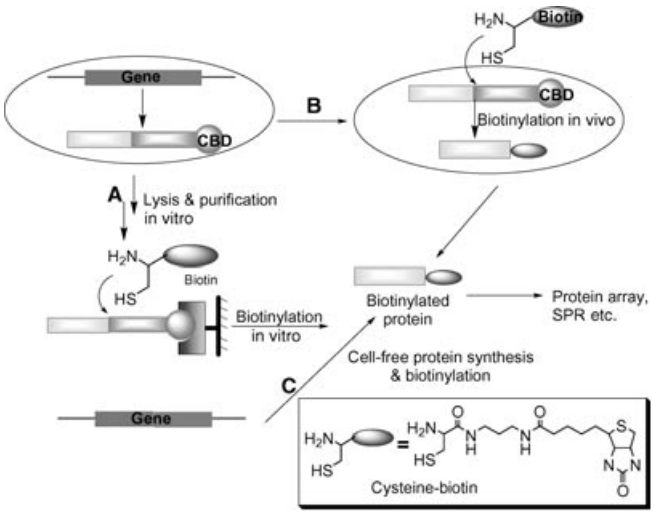
Current applications: Intein-mediated recombinant approaches provide a biological alternative to the aforementioned native chemical ligation reaction for the semi-synthesis of proteins.[1c] This so-called "expressed protein ligation (EPL)" strategy, or intein-mediated protein ligation, has been extensively reviewed recently.[1d] Briefly, by generating proteins that contain either a C-terminal thioester or an N-terminal cysteine residue by using protein expression systems with self-cleavable affinity tags based on modified inteins, it is now possible to introduce unnatural functional groups into large proteins by means of semisynthetic approaches. Many important proteins have been successfully synthesized in vitro, including the 600 amino acid N-terminal segment of the σ^{70} subunit of *E. coli* RNA polymerase,[11a] membrane proteins,[11b] and ion channels.[11c] The use of inteins in combination with proteases has led to the so-called expressed enzymatic ligation approach.[11d] *trans*-Splicing inteins, in which the functionally mature inteins are split into two smaller intein pieces, regain their activity upon reconstitution of the fragments.[10] They have found a variety of applications in vitro, including protein semisynthesis[12] and segmental isotopic labeling.[13] Split inteins have been used to cyclize proteins in vivo,[14] and to study protein–protein interactions in living cells.[15] Indeed, protein *trans*-splicing has many of the attributes necessary for the semisynthesis of

unnatural proteins in vivo. Recently, Muir et al. cleverly adopted the *trans*-splicing property of the *Ssp* DnaE intein for the semisynthesis of proteins in live cells, in which the specific incorporation of chemical probes into the protein was successfully demonstrated.[16]

Expanded applications: We recently expanded the intein-mediated recombinant approaches into several areas related to proteomics and chemical biology, namely, the in vitro and in vivo protein biotinylation,[17] protein microarray technologies for large-scale protein analysis,[17a,18] and live-cell bioimaging.[19] Herein, we summarize our recent findings and, at the same time, look ahead to some future developments.

Protein biotinylation: The avidin–biotin interaction has gained much prominence in protein science due to the remarkable affinity of the protein avidin and its small molecule ligand, biotin (vitamin H, 0.24 kDa). With a K_d of 10^{-15} M, this is one of the strongest known noncovalent interactions in nature.[20] An essential prerequisite for the success of avidin–biotin technologies is the incorporation of the biotin moiety into the protein of interest. Before site-specific methods for protein biotinylation were available, simple bioconjugation techniques with biotin-containing reactive chemicals were used to biotinylate proteins nonspecifically, in many cases leading to the inactivation of proteins.[21] Cronan et al. identified small peptide sequences that could be sufficiently biotinylated by biotin ligase, a 35.5 kDa monomeric enzyme, upon fusion to a protein.[22] This provided, for the first time, a means to site-specifically biotinylate proteins both in vitro and in vivo. Early peptide sequences identified were relatively long (>63 amino acids), and thus may potentially interfere with the native activity of the fused protein. Subsequent optimizations revealed that smaller peptide tags of 15 to 30 amino acids could also be sufficiently biotinylated. In general, in vivo biotinylation of proteins using this approach is often inefficient and toxic to the host cell.[22,23] Other problems include proteolytic degradation of the tag sequence and the inhibitory effects of commonly used reagents towards biotin ligase.[23] To overcome some of these drawbacks, our group recently developed an intein-based system for site-specific protein biotinylation.[17] We have thus far shown that the strategy is highly efficient and versatile, equally applicable for both in vitro and in vivo experiments with different host cells (e.g., bacteria and mammalian cells). It is also compatible with large-scale protein biotinylation for potential proteomic experiments.[17b]

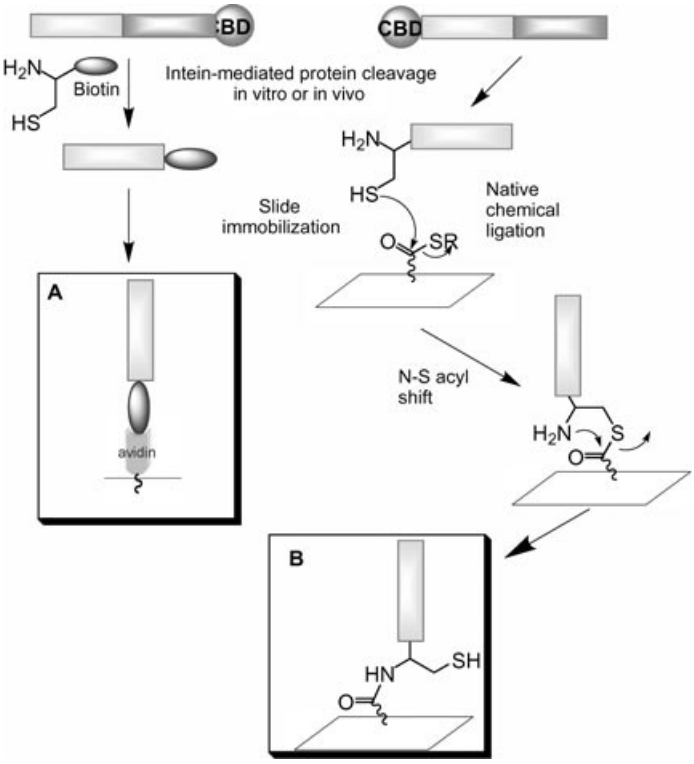
In our strategy (Scheme 2), the protein of interest was fused, through its C-terminus, to an intein.[17b] Upon expression in a host cell, the fusion protein underwent intramolecular rearrangement under appropriate conditions, which was catalyzed by the fused intein, and resulted in the generation of a new protein–intein fusion joined by means of a thioester linkage. The thioester-containing fusion protein was subsequently purified and biotinylated in vitro, in a single step, to generate the final protein in which the biotin moiety was exclusively and covalently attached to the C-terminus of the target protein (Method A in Scheme 2). We showed that the strategy was highly versatile and capable of biotinylating a variety of proteins having different primary sequences and biological functions. The approach is simple yet efficient, thus amendable to potential high-throughput biotinylation of proteins in large-scale proteomic experiments. We also showed that the intein-mediated protein biotinylation approach worked in vivo with both bacterial and mammalian cells (Method B in Scheme 2). Since the bioitinylating reagent used in our approach, that is, cysteine biotin, was a cell-permeable small molecule, by simply treating host cells expressing the thioester-containing fusion protein we were able to show that protein biotinylation, which is very similar to the native chemical ligation reaction, occurred inside live bacteria and mammalian cells. We also investigated and successfully demonstrated that the intein-mediated protein biotinylation strategy worked in a cell-free system (Method C in Scheme 2).[17b] A cell-free system has many advantages over traditional recombinant methods used for protein expression.[24] It eliminates problems of protein toxicity and potential proteolytic degradation of the protein by endogenous proteases. It may also prevent the formation of inclusion bodies, a problem typically encountered when expressing eukaryotic proteins in prokaryotic hosts.



Scheme 2. Three intein-mediated protein biotinylation strategies. Method A: In vitro biotinylation of column-bound proteins; Method B: In vivo biotinylation in live cells; Method C: Cell-free biotinylation of proteins.[17b]

Protein microarray applications: One of the key applications of the above-described site-specific biotinylation approach is in the domain of protein microarrays.[25] Avidin, being an extremely stable protein, is an excellent candidate for slide derivatization and immobilization. Each avidin/streptavidin molecule can bind rapidly and almost irreversibly up to four molecules of biotin, thus doing away with the long incubation time that alternative methods typically need for the critical immobilization step.[25] Avidin also acts as a molecular layer that minimizes nonspecific binding of proteins to the slide surface, thereby eliminating blocking procedures and minimizing background signals in downstream screenings. Our group was the first to employ the biotin–avidin

system in the fabrication of peptide[26] and protein microarrays.[17a] Our strategy for protein microarray generation utilized the aforementioned intein-mediated protein biotinylation method (Method A in Scheme 2). Within a single column-purification step, the expressed C-terminal fusion protein was purified and biotinylated on chitin beads, and subsequently immobilized in a site-specific manner onto an avidin-functionalized glass slide to generate the corresponding protein microarray (A in Scheme 3).[17,18] This highly robust and novel protein array features uniformly oriented proteins, which may help immobilized proteins to retain their full biological activities. The use of biotin–avidin interaction for immobilization also allows proteins to withstand chemically harsh conditions used for downstream screenings, thus making the protein array compatible with most biochemical assays. We further demonstrated that cells expressing in vivo biotinylated proteins could be lysed and directly spotted to generate the corresponding protein microarray.[18] Overall, although the in vivo biotinylation of proteins with an intein tag was less efficient relative to the in vitro system, it nevertheless provided an alternative method for researchers to generate in vivo, without further processing (e.g. purification and elution, etc.), a large array of biotinylated, ready-to-spot proteins in a truly high-throughput, high-content fashion. Protein biotinylation using the aforementioned cell-free synthesis system (Method C in Scheme 2) may provide yet another attractive alternative to obtain spotting-ready protein samples for protein microarray applications for the following reasons:



Scheme 3. Two intein-mediated strategies for the site-specific protein immobilization to generate the corresponding protein arrays by using A) biotin–avidin interaction and B) native chemical ligation reaction.

1) Proteins may be readily biotinylated site-specifically at their C-termini using the cell-free method.
2) Minute quantities of proteins generated in cell-free system are sufficient for spotting and detection in a protein array.
3) The method could be easily adopted in 96- and 384-well formats with a conventional PCR machine for potential high-throughput protein synthesis.[24]

Another strategy we have developed for site-specific immobilization of proteins is through the highly chemoselective native chemical ligation. We previously immobilized N-terminal cysteine-containing peptides onto glass slides chemically derivatized with thioester moieties to generate the corresponding peptide microarray.[26] We recently showed that the same immobilization strategy may be applied to protein microarray generation as well (B in Scheme 3). By extending the intein-mediated expression system, N-terminal cysteine-containing proteins were readily generated; these were subsequently immobilized onto thioester-functionalized glass slides to generate the corresponding protein microarray.[18] The N-terminal cysteine residue of the protein chemoselectively reacted with the thioester on the slide through the formation of a thioester intermediate, followed by an N→S acyl shift to form a native peptide bond. Once again, site-specific immobilization of proteins in this strategy allowed the full retention of their biological activities. We also showed that the strategy is versatile, and applicable to the immobilization of N-terminal cysteine proteins, which were either purified prior to spotting or present in crude cell lysates (e.g., unpurified).[18] There are alternative methods of generating N-terminal cysteine-containing proteins, for example, the ubiquitin fusion in eukaryotic systems that possess endogenous deubiquitinating enzymes which cleave specifically after the last residue of ubiquitin.[27]

Bioimaging: Site-specific biotinylation is just one of the many protein modifications one can perform with the intein system. Some others include the attachment of small organic fluorophores, photocleavable tags, photocrosslinking groups, isotope tags, and so forth, so long as they are conjugated to an appropriate chemical group.[28] We have extended the utility of native chemical ligation and protein splicing to site-specific in vivo labeling of proteins for bioimaging applications.[19]

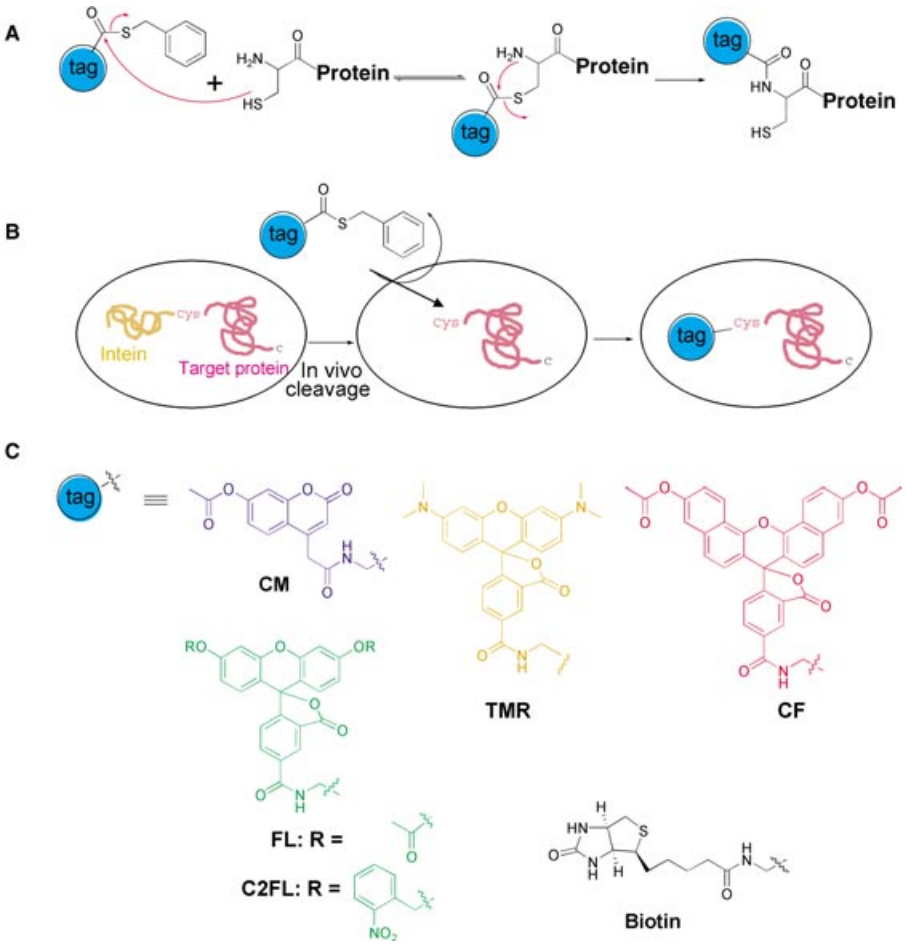
Studying the dynamic movement and interactions of proteins inside living cells is critical for a better understanding of cellular mechanisms and functions. Traditionally this has been done by in vitro labeling of proteins with fluorescent and other molecular probes, followed by monitoring them inside live cells.[29a,b] Recent advances in fluorescent proteins have allowed for the direct generation and visualization of fluorescently labeled proteins in live cells or even in live animals.[29c] However, inherent drawbacks of fluorescent proteins related to bioimaging include their relative large sizes (e.g., ~27 kDa for green fluorescent protein, or GFP), dimer/tetramer formation in certain variants, and so forth, all of which may affect the native biological activity of the fused protein. In addition, relatively few "colors" are availa-

ble amongst existing fluorescent proteins and they are not always ideal fluorophores, as many have broad excitation/emission spectra, low quantum yields, and are susceptible to photobleaching. Lastly, the use of fluorescent proteins limits protein labeling to only fluorescent and not any other molecular tags (e.g., biotin).

Several small molecule-based strategies have recently been developed that allow site-specific labeling of proteins with small molecules and subsequently imaging of them inside living cells.[30–34] Tsien's group designed biarsenical probes with a high affinity for a tetracysteine motif (CCXXCC)[30] and demonstrated their utility in live-cell imaging of the translocation of connexin in and out of gap junctions.[30b] Noncovalent interactions of small molecule ligands with streptavidin- and antibody-conjugated fusions have been used for in vivo labeling of proteins.[31,32] Johnsson et al. described an enzymatic approach to label proteins fused to the human O^6-alkylguanine-DNA alkyltransferase (hAGT) with small molecular substrates.[33] The most recent report of small molecule probes for live-cell labeling is based on the noncovalent interaction between *E. coli* dihydrofolate reductase (DHFR) and methotrexate (Mtx) conjugates.[34] Our group has developed a novel strategy for site-specific covalent labeling of proteins in vivo by using native chemical ligation and intein-mediated protein splicing (Scheme 4).[19]

In our strategy (Scheme 4A), a protein of interest with an N-terminal cysteine group was expressed inside a live cell by using intein-mediated protein splicing.[35] Incubation of the cell with a thioester-containing, cell-permeable, small-molecule probe allowed the probe to efficiently penetrate through the cell membrane into the cell, where the chemoselective native chemical ligation reaction occurred between the thioester of the small molecule and the N-terminal cysteine of the protein, giving rise to the resulting labeled protein (Scheme 4B).[19] Few endogenous N-terminal cysteine-containing proteins are known in various bacterial and mammalian genome databases, making our labeling strategy feasible in future for different live-cell imaging experiments. Other endogenous molecules, such as cysteine and cystamine, are present in the cell and will also react with the probe. However, their reaction products are also small molecules in nature, and can be easily removed, together with any excessive unreacted probe, by extensive washing of the cells after labeling. Other methods for generating recombinant proteins possessing N-terminal cysteine residues, including the use of methionyl aminopeptidases,[36] the ubiquitin-fusion strategy,[27] or selective proteolysis with highly specific proteases like the TEV (tobacco etch virus) protease,[37] may potentially be used with our labeling strategy.

We selected the 17 kDa *Ssp* DnaB mini intein, whose splicing activity was genetically engineered to occur efficiently under physiological conditions, to generate an N-terminal cysteine in the target protein.[35] The intein-mediated approach requires no external factors (e.g., proteases), with the splicing activity affected primarily by the identity of amino acids at the splice junction. It provides a simple approach for site-specific labeling of proteins in live cells with minimal modifications to the target protein, apart from the introduction of a few extra amino acid residues at the N-terminus of the target protein. We have thus shown that the strategy may be readily applied to both bacterial and mammalian cells with a variety of thioester-containing small molecule probes (Scheme 4C).[19,38]



Scheme 4. A) Chemoselective native chemical ligation between an N-terminal cysteine in a protein and a thioester-containing probe, forming a stable amide bond. B) Strategy for site-specific covalent labeling of N-terminal cysteine proteins and thioester probes in live cells. C) Structures of cell-permeable, thioester probes used in the study.[19]

Future directions

Intein-mediated protein splicing has thus far only been observed with prokaryotic and unicellular eukaryotic organisms (e.g., yeasts).[10] Consequently, most of the intein-mediated protein engineering approaches developed to date are only suitable in bacteria and yeasts.[1d,10–15] Recent work by Muir et al. and our group, however, indicated that these approaches may work in mammalian cells as well.[16,17b,38] There are other known forms of protein autoprocessing, most notably ones in the family of hedgehog proteins found in animals.[39] For example, a typical hedgehog precursor protein contains a 20 kDa N-terminal domain (e.g., the signaling domain) and a 25 kDa C-terminal domain (e.g., the autoprocessing domain). It undergoes self-catalyzed cleavage of the C-terminal domain, by means of a mechanism similar to intein-mediated protein splicing, and results in the attachment of a cholesterol molecule to the N-terminal domain.[40] It is therefore reasonable to believe that hedgehog proteins, just like inteins for bacteria and yeasts, are ideal models to develop mammalian-based protein-engineering approaches that utilize the protein splicing mechanism. We are currently exploring the feasibility of using hedgehog-based protein splicing to site-specifically modify proteins in live mammalian cells.

Another area to develop efficient methods for in vivo site-specific modification of proteins may involve enzyme-based protein ligation reactions, of which very few approaches have been described to date.[41–45] For example, subtiligase is an engineered subtilisin mutant capable of catalyzing the joining of two peptide fragments in a reaction known as reverse proteolysis.[41] Some other natural proteases have also been found to undergo reverse proteolysis efficiently under certain reaction conditions.[42] We envisage that an extension of reverse proteolysis to highly specific proteases, for example, TEV protease,[43] could potentially allow for highly specific modifications of proteins inside live cells. Another enzyme-based in vivo protein modification technique may involve the use of sortase, a transpeptidase found in the cell envelope of many Gram-positive bacteria.[44] *S. aureus* sortase (SrtA) is known to undergo the so-called "transpeptidation" reaction, by cleaving between threonine and glycine at an LPXTG recognition motif in a protein and subsequently joining the carboxyl group of threonine to an amino group of pentaglycine on the cell-wall peptidoglycan. Recently, Pollock et al. successfully demonstrated the in vitro sortase-mediated ligation of proteins to both peptides and nonpeptides.[45] Current efforts in our group are directed toward extending this sortase-based approach to in vivo modifications of proteins.

Acknowledgement

Funding support was provided by the National University of Singapore (NUS) and the Agency for Science, Technology and Research (A*STAR) of Singapore. SQY is the recipient of the 2002 BMRC Young Investigator Award from the Biomedical Research Council of Singapore.

[1] a) G. A. Lemieux, C. R. Bertozzi, *Trends Biotechnol.* **1998**, *16*, 506–513; b) H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056–2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2041; c) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960. d) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249–289.

[2] a) T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, *Justus Liebigs Ann. Chem.* **1953**, *583*, 129–139; b) M. Brenner, J. P. Zimmermann, J. Wehrmuller, P. Quitt, A. Hartmann, W. Schneider, U. Berlinger, *Helv. Chim. Acta* **1957**, *40*, 1497–1505; c) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.

[3] a) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; b) L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533. c) T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073; d) G. S. Beligere, P. E. Dawson, *J. Am. Chem. Soc.* **1999**, *121*, 6332–6333; e) L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896; f) J. Offer, P. E. Dawson, *Org. Lett.* **2000**, *2*, 23–26.

[4] Y. G. Gololobov, L. F. Kasukhin, *Tetrahedron* **1992**, *48*, 1353–1405.

[5] a) E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010; b) E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141–2143; c) B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939–1941; d) B. L. Nilsson, R. J. Hondal, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 5268–5269.

[6] a) Z. Zhang, B. A. C. Smith, L. Wang, A. Brock, C. Cho, P. G. Schultz, *Biochemistry* **2003**, *42*, 6735–6746; b) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.

[7] L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125–1128.

[8] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.

[9] A. E. Speers, G. C. Adam, B. F. Cravatt, *J. Am. Chem. Soc.* **2003**, *125*, 4686–4687.

[10] H. Paulus, *Annu. Rev. Biochem.* **2000**, *69*, 447–496.

[11] a) K. Severinov, T. W. Muir, *J. Biol. Chem.* **1998**, *273*, 16205–16209; b) F. I. Valiyaveetil, R. MacKinnon, T. W. Muir, *J. Am. Chem. Soc.* **2002**, *124*, 9113–9120; c) F. I. Valiyaveetil, M. Sekedat, T. W. Muir, R. MacKinnon, *Angew. Chem.* **2004**, *116*, 2558–2561; *Angew. Chem. Int. Ed.* **2004**, *43*, 2504–2507; d) Z. Machova, R. von Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, *Angew. Chem.* **2003**, *115*, 5065–5067; *Angew. Chem. Int. Ed.* **2003**, *42*, 4916–4918.

[12] B. M. Lew, K. V. Mills, H. Paulus, *J. Biol. Chem.* **1998**, *273*, 15887–15890.

[13] T. Yamazaki, T. Otomo, N. Oda, Y. Kyogoku, K. Uegaki, N. Ito, Y. Ishino, H. Nakamura, *J. Am. Chem. Soc.* **1998**, *120*, 5591–5592.

[14] a) C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638–13643; b) T. C. Evans, D. Martin, R. Kolly, D. Panne, L. Sun, I. Ghosh, L. Chen, J. Benner, X. Q. Liu, M. Q. Xu, *J. Biol. Chem.* **2000**, *275*, 9091–9094; c) H. Iwai, A. Lingel, A. Pluckthun, *J. Biol. Chem.* **2001**, *276*, 16548–16554.

[15] a) T. Ozawa, S. Nogami, M. Sato, Y. Ohya, Y. Umezawa, *Anal. Chem.* **2000**, *72*, 5151–5157; b) T. Ozawa, Y. Umezawa, *Curr. Opin. Chem. Biol.* **2001**, *5*, 578–583.

[16] I. Giriat, T. W. Muir, *J. Am. Chem. Soc.* **2003**, *125*, 7180–7181.

[17] a) M. L. Lesaicherre, R. Y. P. Lue, G. Y. J. Chen, Q. Zhu, S. Q. Yao, *J. Am. Chem. Soc.* **2002**, *124*, 8768–8769; b) R. Y. P. Lue, G. Y. J. Chen, Y. Hu, Q. Zhu, S. Q. Yao, *J. Am. Chem. Soc.* **2004**, *126*, 1055–1062.

[18] a) R. Y. P. Lue, D. S. Y. Yeo, L. P. Tan, M. Uttamchandani, G. Y. J. Chen, S. Q. Yao, in *Protein Microarrays* (Ed.: M. Schena), Jones and Bartlett, Sudbury, MA, **2004**, Chapter 2; b) R. Y. P. Lue, D. S. Y. Yeo, L. P. Tan, G. Y. J. Chen, S. Q. Yao, in *Proteomics Handbook* (Ed.: J. Walker), Humana, Totowa, **2004**.

[19] D. S. Y. Yeo, R. Srinivasan, M. Uttamchandani, G. Y. J. Chen, Q. Zhu, S. Q. Yao, *Chem. Commun.* **2003**, 2870–2871.

[20] M. Wilchek, E. A. Bayer, *Methods Enzymol.* **1990**, *184*, 5–13.

[21] G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, San Diego, CA, **1996**.

[22] J. E. Cronan, K. E. Reed, *Methods Enzymol.* **2000**, *327*, 440–458.
[23] M. G. Cull, P. J. Schatz, *Methods Enzymol.* **2000**, *326*, 430–440.
[24] M. Y. He, M. J. Taussig, *Nucleic Acids Res.* **2001**, *29*, e73.
[25] a) G. MacBeath, S. L. Schreiber, *Science* **2000**, *289*, 1760–1763; b) G. Y. J. Chen, M. Uttamchandani, R. Y. P. Lue, M. L. Lesaicherre, S. Q. Yao, *Curr. Top. Med. Chem.* **2003**, *3*, 705–724.
[26] a) M. L. Lesaicherre, M. Uttamchandani, G. Y. J. Chen, S. Q. Yao, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2079–2083; b) M. Uttamchandani, E. W. S. Chan, G. Y. J. Chen, S. Q. Yao, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2997–3000; c) M. Uttanchandani, G. Y. J. Chen, M. L. Lesaicherre, S. Q. Yao, *Methods Mol. Biol.* **2004**, *278*, 191–204.
[27] R. T. Baker, *Curr. Opin. Biotechnol.* **1996**, *7*, 541–546.
[28] T. Tolbert, C. H. Wong, *J. Am. Chem. Soc.* **2000**, *122*, 5421–5428.
[29] a) A. Miyawaki, A. Sawano, T. Kogure, *Nat. Cell Biol.* **2003**, *Suppl.*, S1–S7; b) J. Zhang, R. E. Campbell, A. Y. Ting, R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906–918; c) R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509–544.
[30] a) B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269–272; b) G. Gaietta, T. J. Deerinck, S. R. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien, M. H. Ellisman, *Science* **2002**, *296*, 503–507.
[31] M. M. Wu, J. Llopis, S. Adams, J. M. McCaffery, M. S. Kulomaa, T. E. Machen, H. P. Moore, R. Y. Tsien, *Chem. Biol.* **2000**, *7*, 197–209.
[32] J. Farinas, A. S. Verkman, *J. Biol. Chem.* **1999**, *274*, 7603–7606.
[33] A. Keppler, S. Gendreizig, T. Gronemeyer, H, Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86–89.
[34] L. W. Miller, J. Sable, P. Goelet, M. P. Sheetz, V. W. Cornish, *Angew. Chem.* **2004**, *116*, 1704–1707; *Angew. Chem. Int. Ed.* **2004**, *43*, 1672–1675.
[35] M. Q. Xu, T. C. Evans, *Methods* **2001**, *24*, 257–277.
[36] P. H. Hirel, J. M. Schmitter, P. Dessen, G. Fayat, S. Blanquet, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8247–8351.
[37] T. J. Tolbert, C. H. Wong, *Angew. Chem.* **2002**, *114*, 2275–2278; *Angew. Chem. Int. Ed.* **2002**, *41*, 2171–2174.
[38] S. Y. D. Yeo, R. Srinivasan, G. Y. J. Chen, Q. Zhu, S. Q. Yao, unpublished results.
[39] a) J. J. Lee, S. C. Ekker, D. P. Kessler, J. A. Porter, B. I. Sun, P. A. Beachy, *Science* **1994**, *266*, 1528–1537; b) J. A. Porter, S. C. Ekker, W. J. Park, D. P. Kessler, K. E. Young, C. H. Chen, Y. Ma, A. S. Woods, R. J. Cotter, E. V. Koonin, P. A. Beachy, *Cell* **1996**, *86*, 21–34; c) J. A. Porter, , K. E. Young, P. A. Beachy, *Science* **1996**, *274*, 255–259.
[40] R. K. Mann, P. A. Beachy, *Biochim. Biophys. Acta* **2000**, *1529*, 188–202.
[41] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243–247.
[42] a) N. Wehofsky, N. Koglin, S. Thust, F. Bordusa, *J. Am. Chem. Soc.* **2003**, *125*, 6126–6133; b) F. Bordusa, *Chem. Rev.* **2002**, *102*, 4817–4867.
[43] J. Phan, A. Zdanov, A. G. Evdokimov, J. E. Tropea, H. K. Peters III, R. B. Kapust, M. Li, A. Wlodawer, D. S. Waugh, *J. Biol. Chem.* **2002**, *277*, 50564–50572.
[44] S. K. Mazmanian, G. Liu, H. Ton-That, O. Schneewind, *Science* **1999**, *285*, 760–763.
[45] H. Mao, S. A. Hart, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.

Published online: August 11, 2004

