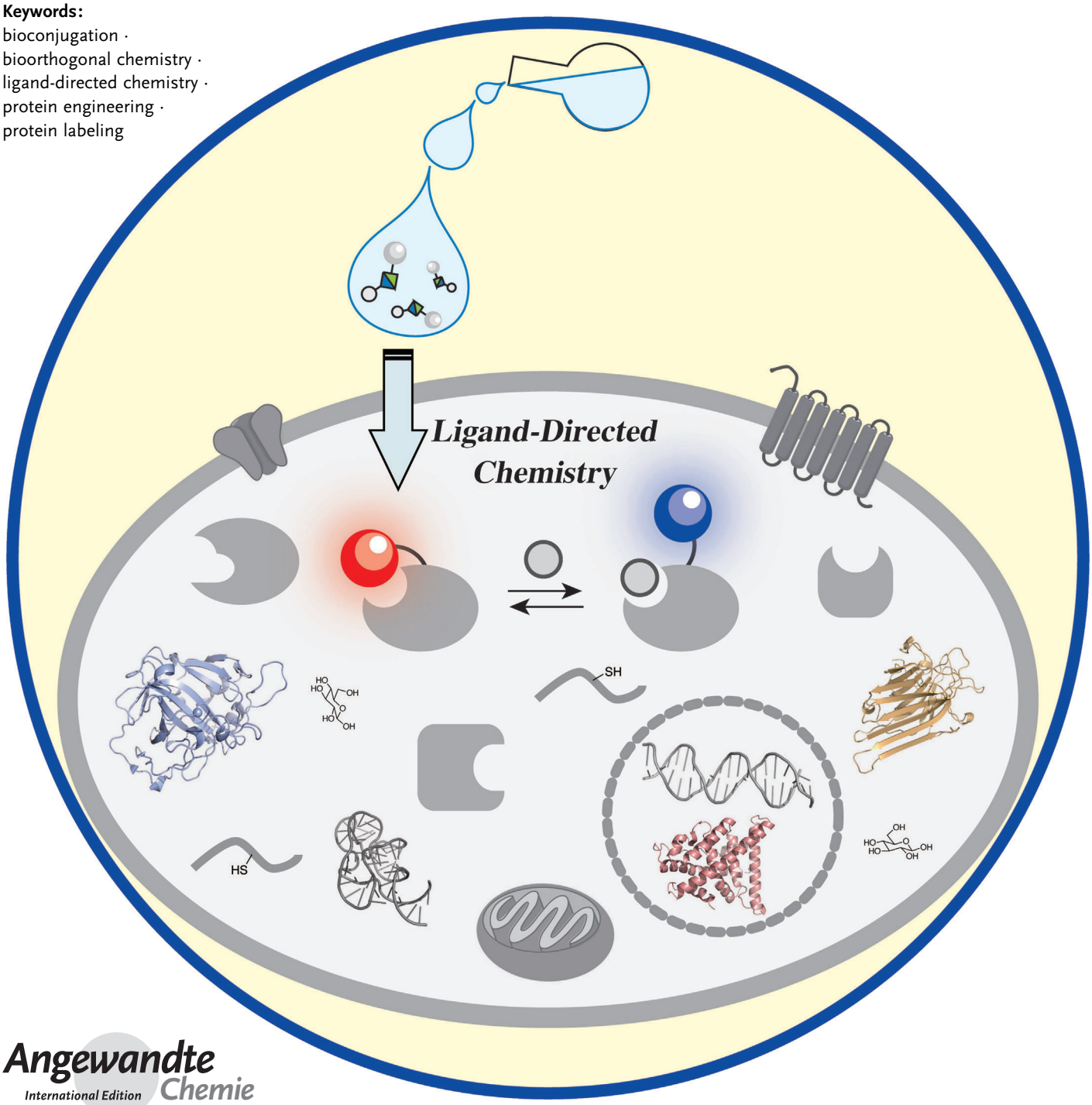
DOI: 10.1002/anie.201207089

Protein Modification

Protein Organic Chemistry and Applications for Labeling and Engineering in Live-Cell Systems

*Yousuke Takaoka, Akio Ojida, and Itaru Hamachi**

Keywords:
bioconjugation · bioorthogonal chemistry · ligand-directed chemistry · protein engineering · protein labeling



Angewandte International Edition Chemie

The modification of proteins with synthetic probes is a powerful means of elucidating and engineering the functions of proteins both in vitro and in live cells or in vivo. Herein we review recent progress in chemistry-based protein modification methods and their application in protein engineering, with particular emphasis on the following four strategies: 1) the bioconjugation reactions of amino acids on the surfaces of natural proteins, mainly applied in test-tube settings; 2) the bioorthogonal reactions of proteins with non-natural functional groups; 3) the coupling of recognition and reactive sites using an enzyme or short peptide tag–probe pair for labeling natural amino acids; and 4) ligand-directed labeling chemistries for the selective labeling of endogenous proteins in living systems. Overall, these techniques represent a useful set of tools for application in chemical biology, with the methods 2–4 in particular being applicable to crude (living) habitats. Although still in its infancy, the use of organic chemistry for the manipulation of endogenous proteins, with subsequent applications in living systems, represents a worthy challenge for many chemists.

From the Contents

1. Introduction

The study of protein structure and function is not only of significant importance to fundamental scientific research, but also critical to the development of biomedical and biotechnological applications, because proteins are implicated in so many biological systems. The modification of proteins with molecular probes provides a powerful technique for the elucidation of protein functions.[1,2] For fluorescence bioimaging studies in live cells, for example, the proteins of interest should be labeled with a fluorescent marker, which allows for the selective real-time detection of their localization, trafficking, and activities. Protein modification with synthetic molecules is also a valuable technique for the engineering of protein functions to create new biocatalysts and bio-analytical tools. In a recent trend in chemical biology research, there has been an aspiration to conduct protein modification under more biological (crude) conditions, such as in live cells or tissues/bodies, rather than in pure test tubes (in vitro). Several issues need to be addressed for the construction of homogenously modified proteins with synthetic molecules, including 1) protein selectivity, 2) labeling-site selectivity, 3) control of the labeling-site numbers, and 4) expansion of the labeled amino acids. Recently developed high-precision protein modification methods can be roughly divided into two categories, with the first involving the use of a genetic modification system exploiting the expanded genetic code, and the second being based on the labeling of expressed proteins (post-translational protein modification). The first technique was pioneered by Peter G. Schultz, and involves the direct introduction of a non-natural amino acid possessing a fluorophore or other functionalities into a target protein scaffold at a desirable position.[3–6] The application of this technique, however, can sometimes be frustrating because of the limited functionality of the probes and the restricted number of usable cell lines, as well as the insufficient yield of the genetically engineered proteins. In contrast, the technique of post-translational protein modification is better suited to fully engage the variety of different functionalities present in synthetic molecules. This method of protein modification, which is driven predominantly by chemistry, allows for the flexible and efficient modification of proteins at any appropriate time and position in the protein sequence. The proteins themselves can be modified with a variety of functional molecules, including luminescent dyes, spin-active probes, NMR-active probes, photo-responsive (caged) molecules, polymers, provided the issues mentioned above can be overcome. The main focus of the current Review is post-translational protein modification, and a description of recent progress in this research area toward expanding the utility of these chemical tools for fundamental biological research and applications has been provided (Scheme 1).

[*] Dr. Y. Takaoka, Prof. Dr. I. Hamachi
Department of Synthetic Chemistry and Biological Chemistry
Kyoto University
Katsura, Nishikyo-Ku, Kyoto 615-8510 (Japan)
E-mail: ihamachi@sbchem.kyoto-u.ac.jp
Prof. Dr. A. Ojida
Graduate School of Pharmaceutical Sciences, Kyushu University
3-1-1 Maidashi, Higashi-Ku, Fukuoka 812-8582 (Japan)
Prof. Dr. I. Hamachi
Japan Science and Technology Agency (JST), CREST
5 Sanbancho, Chiyoda-ku, Tokyo 102-0075 (Japan)

2. Bioconjugation of Natural Amino Acids for Protein Modification and Engineering

2.1. Bioconjugation Reactions of Natural Amino Acids on a Protein Surface

2.1.1. Classical Bioconjugation Methods for Proteins

From the chemical perspective, all proteins can be regarded as folded polymers which have many nucleophilic functional groups on their surfaces. Of the various naturally occurring nucleophilic amino acid residues present in proteins, the thiol group of cysteine is often used for the site-specific modification of proteins because of its higher nucleophilicity and relatively low level of occurrence on the surfaces of many proteins (only 2.3% genome-wide). If a reactive cysteine does not exist on the protein surface then one can be introduced by a genetic point mutation, which is an effective technique for site-specific protein modification. The thiol group of cysteine can readily undergo alkylation by reaction with α-haloketones or Michael acceptors, such as maleimide derivatives (Scheme 2a). Although considerably more prevalent than cysteine, the ε-amino group of lysine is a popular target for protein modification because of the abundance of methods available for the selective modification of primary amines.[7] Lysine can react with activated esters, sulfonyl chlorides, isocyanates, and isothiocyanates to afford the corresponding amides, sulfonamides, ureas, and thioureas, respectively (Scheme 2b). It is worthy of note that these reagents can also modify the N termini of proteins.

Dr. Yousuke Takaoka was born in 1982 in Shimane, Japan. He received his PhD (with Prof. I. Hamachi) from Kyoto University in 2010 and carried out post-doctoral research (Prof. K. Hirose) at the University of Tokyo. He joined the faculty of Kyoto University as an assistant professor in 2011. His research interests include supramolecular chemistry, protein engineering, neurobiology, and chemical biology.

Prof. Akio Ojida was born in 1968 in Fukuoka, Japan. He received his PhD (with Prof. K. Kanemitsu) from Kyushu University in 1995 and carried out post-doctoral research (Prof. M. Shionoya) at the Institute for Molecular Science. He worked for Takeda Chemical Industries (1997–2001) then became an assistant professor at Kyushu University. In 2003 he became an assistant professor at Kyoto University, and a lecturer in 2007. In 2010 he was made a full professor as Kyushu University. His interests focus on chemical biology, especially on design of small molecular probes for elucidating biological functions.

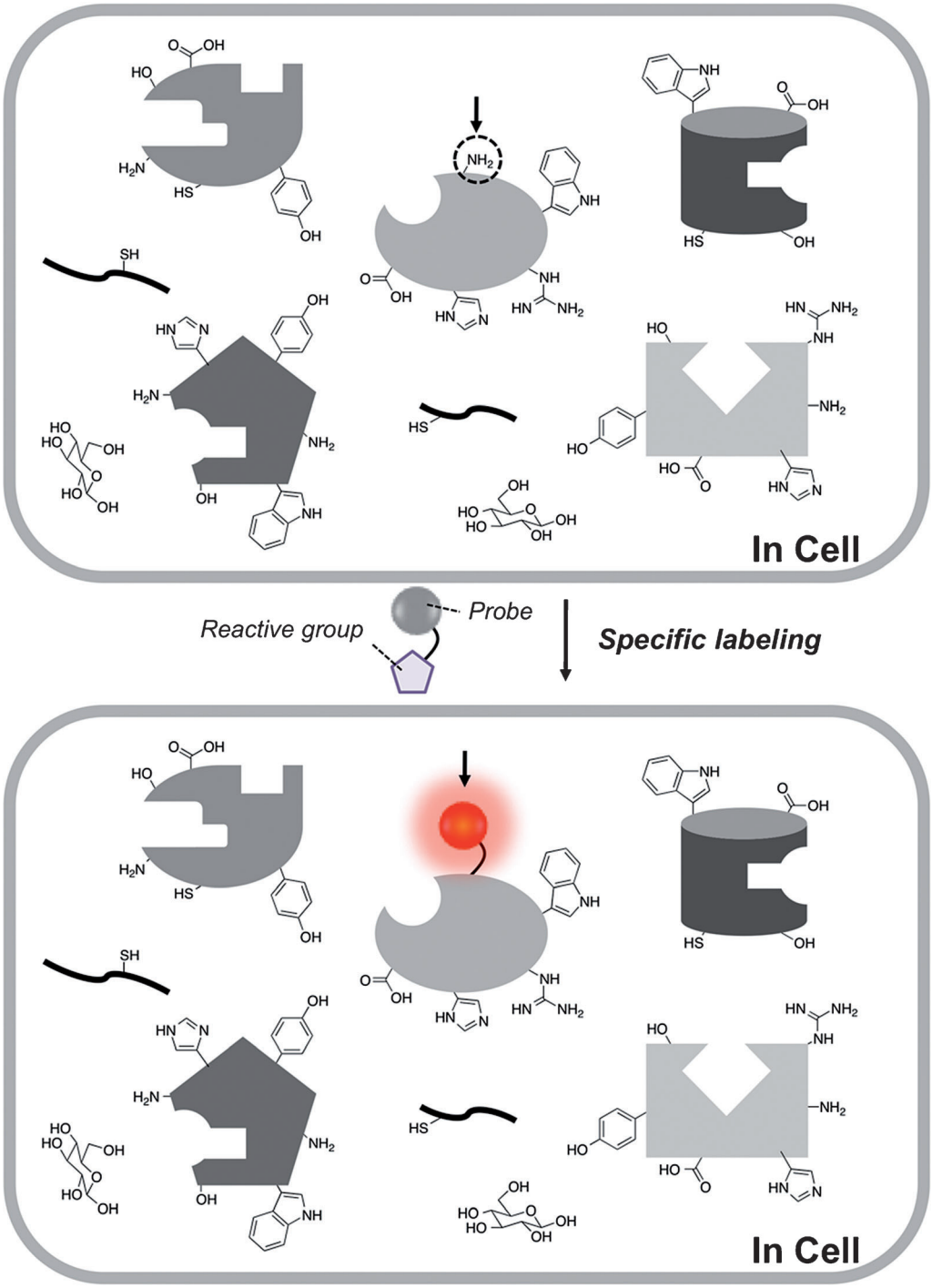
Prof. Itaru Hamachi was born in 1960 in Fukuoka, Japan. He received his PhD (Prof. I. Tabushi) from Kyoto University in 1988. He joined the faculty of Kyushu University as an assistant professor in 1988, and became an associate professor at Kyushu University in 1992. In 2001, he became a full professor at Kyushu University and then moved to Kyoto University in 2005. He is now CREST investigator of JST and a RSC fellow. His research interests include biomolecular chemistry, nanobiochemistry and engineering, supramolecular chemistry, and chemical biology.

2.1.2. New Methods for the Modification of Lysine and Cysteine Residues

Several new methods for the selective modification of amines and thiols have recently been developed and optimized. For example, Francis and McFarland reported a lysine-specific reductive alkylation reaction that proceeded by an iridium-catalyzed transfer hydrogenation (Scheme 3a).[8] In contrast to the classical technique of reductive alkylation using sodium cyanoborohydride under acidic conditions, the iridium-mediated process provided high yields of the desired products under neutral pH conditions. Francis's group also reported a unique biomimetic transamination of the N terminus of a protein. This method involved the condensation of the N-terminal amine of the protein with pyridoxal-5-phosphate, followed by hydrolysis to provide the corresponding pyruvamide (Scheme 3b).[9] The resulting protein was then further modified by the reaction of the newly generated ketone moiety of the pyruvamide with aminooxy reagents. Fukase et al. reported a new lysine-based labeling method for proteins on the basis of a 6π-aza-electrocyclization (Scheme 3c).[10] Using this techniques, a short-lived PET (positron emission tomography) probes can be incorporated into the target proteins or hormones and utilized for the in vivo imaging of glycans and glycoconjugates because this reaction is extremely rapid (within 30 min under neutral pH conditions).[10,11] Davis et al. recently developed a two-step cysteine modification method (Scheme 3d), with the first step involving the transformation of cysteine to dehydroalanine by treatment with *O*-mesitylenesulfonylhydroxylamine. The dehydroalanine residues were then used as "reactive handles" for further modifications, such as a Michael addition with thiol reagents[12] or olefin cross-metathesis reactions with allyl sulfides catalyzed by the Hoveyda–Grubbs ruthenium catalyst.[13]

2.1.3. Bioconjugation for Tyrosine and Tryptophane

In contrast to Cys and Lys, the remaining canonical 18 amino acid residues have been only minimally explored as tools for selective modification. Novel methods have been developed for the modification of tyrosine and tryptophane involving transition-metal-mediated processes. In their pioneering work, Kodadek et al. developed a method for the oxidative coupling of the phenolic groups of two Tyr residues for cross-linking two proteins using a ruthenium(II) catalyst and a co-oxidant.[14,15] Francis and co-workers explored the modification of tyrosine residues through a three-component Mannich reaction with aldehydes and anilines (Scheme 4a).[16,17] They also succeeded in modifying Tyr residues

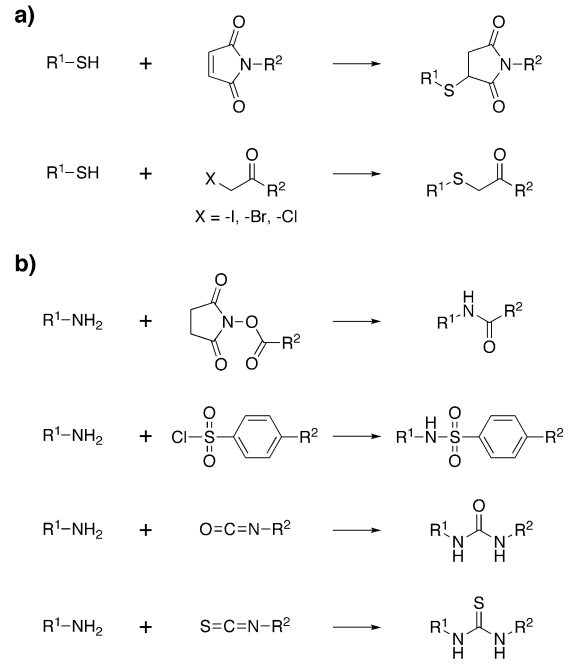


Scheme 1. Schematic illustration of the chemical reactions used for specific endogenously expressed protein modifications in miscellaneous conditions including various reactive species (such as proteins, sugars, and peptides).

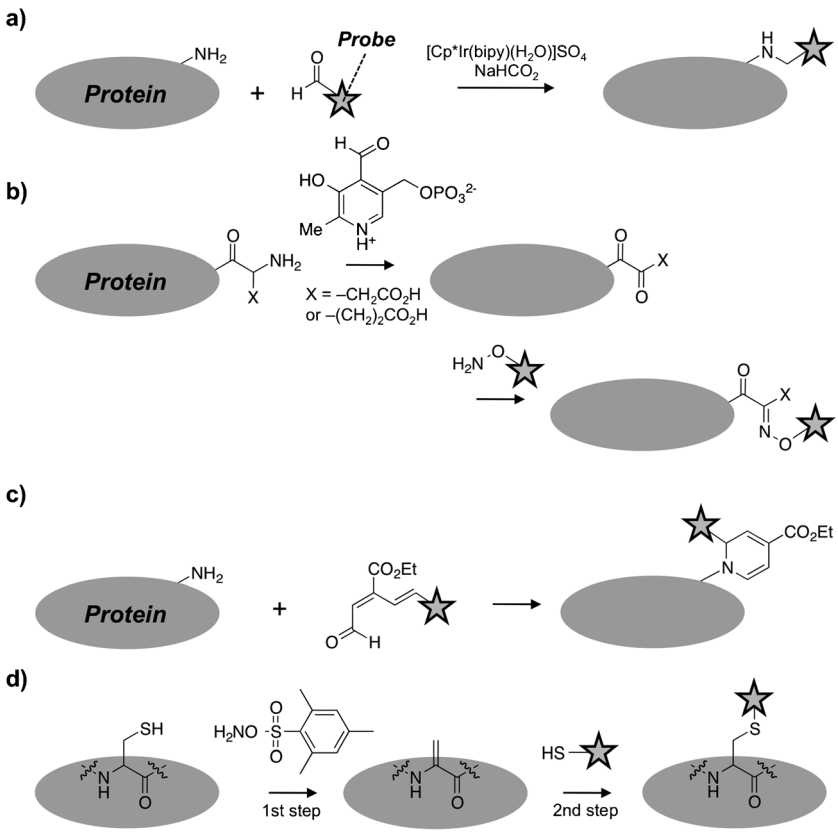
using π-allyl palladium complexes (Scheme 4b).[18] Francis et al. reported an interesting bioconjugation reaction for tryptophane, which is the rarest amino acid in proteins, involving a rhodium carbenoid that was generated in situ from rhodium acetate and a diazo compound (Scheme 4c).[19,20]

2.2. Protein Engineering Using the Bioconjugation Method

The majority of the methods of bioconjugation for protein engineering are conducted in test-tube settings, which provides a variety of different opportunities for manipulating, improving, or mimicking the performances of proteins. The modification of proteins with polyethylene glycol (PEG) groups, otherwise known as PEGylation, is one of the most representative applications in protein bioconjugation.[21,22]



Scheme 2. Classical bioconjugation reactions for the modification of a) cysteine and b) lysine.

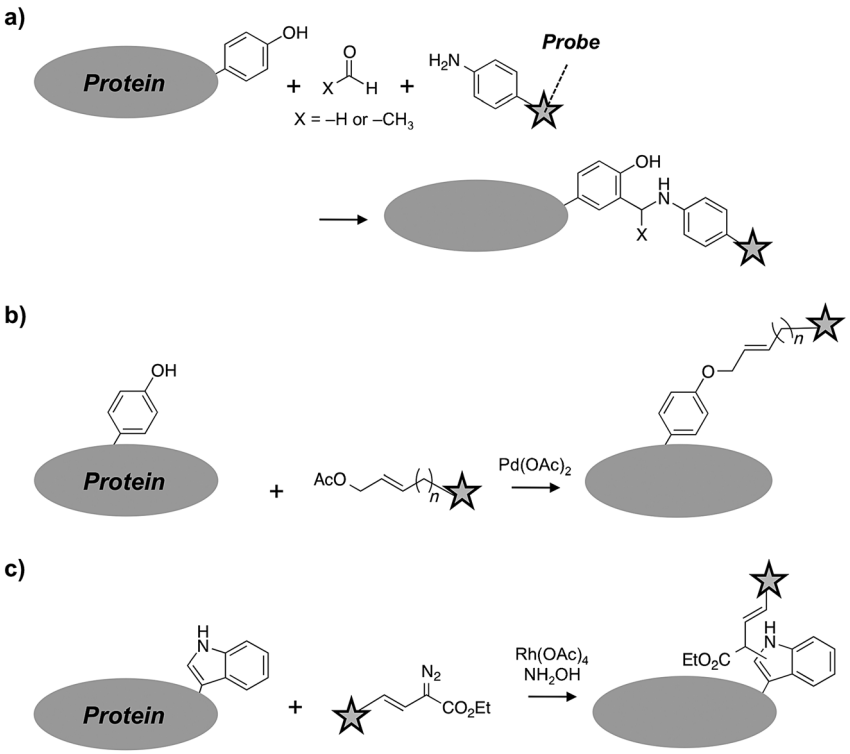


Scheme 3. Modern bioconjugation reactions for protein modification using lysine and cysteine. a) Iridium-catalyzed transfer hydrogenation for a lysine-specific reductive alkylation reaction (Cp* = C_5Me_5, bipy = bipyridyl). b) Biomimetic transmamination, specific for the N-terminus, with pyridoxal-5-phosphate and labeling of the newly generated ketone with aminooxy reagents. c) 6π-Aza-electrocyclization reaction for a lysine specific modification. d) Transformation of cysteine into dehydroalanine with *O*-mesitylenesulfonylhydroxylamine and the Michael addition of thiol reagents to give cysteine-specific modifications.

PEG–protein conjugates are typically constructed using the methods highlighted above (e.g. cysteine–maleimide, disulfide bond formation, and lysine-activated esters).[23,24] These conjugates offer several advantages relative to the parent proteins, including increased solubility, prolonged residence in the body, decreased degradation by metabolic enzymes and reduced immunogenicity, and these properties can play an important roles in the delivery of protein drugs[25] and the development of improved therapeutic agents.[26]

2.2.1. *Protein-Based Biosensors for Detection of Biologically Active Agents*

Protein-based semisynthetic biosensors, which consist of a ligand-binding protein and an artificial probe for the output signal, provide an indispensable approach to detect and monitor not only a range of important substances, such as saccharides and various chemicals, but also many biological events in live cells, such as protein–protein interactions (Scheme 5).[27,28] The site-specific replacement of a less-reactive amino acid residue with cysteine by using conventional site-directed mutagenesis techniques is the most common way to covalently attach a synthetic probe to a ligand-binding protein. In their pioneering work, Hellinga et al. converted bacterial periplasmic saccharide binding proteins (bPBP) into fluorescent biosensors for the specific detection of saccharides (Scheme 5b).[29] They succeeded in developing a variety of bPBP-based fluorescent biosensors on the basis of the site-specific conjugation of eleven different bPBP scaffolds with different microenvironmentally sensitive fluorophores (e.g. IANBD, Scheme 5c).[30] Morii et al. reported a novel inositol-1,4,5-trisphosphate (IP_3) sensor by the conjugation of the pleckstrin homology (PH) domain from phospholipase C-δ with a microenvironmentally sensitive fluorophore (e.g. acrylodan, Scheme 5c) in the proximity of the IP_3-binding pocket.[31] The IP_3 biosensor was successfully translocated into live cells by using arginine-tagged peptide sequences, which enabled the real-time quantitative monitoring of physiological IP_3 production.[31,32] Hahn et al. have developed a fluorescent biosensor that visualizes the activation of endogenous Cdc42 in live cells by ratiometric sensing.[33] A domain from the Wiskott–Aldrich syndrome protein (WASP), which is a Cdc42 effector protein, was covalently labeled with a solvent-sensitive fluorescent dye (e.g. JPW4039 or I-SO, Scheme 5c)[34] using an in vitro cysteine modification. The labeled WASP domain was fused to an enhanced green fluorescent protein (EGFP), which provided a fluorescence signal that was insensitive to Cdc42 binding and was used as an internal standard. By



Scheme 4. Modern bioconjugation reactions for the protein modification of tyrosine and tryptophane residues. a) Three-component tyrosine-specific Mannich reaction with aldehydes and anilines. b) Tyrosine specific reactions via palladium π-allyl species. c) A tryptophan specific reaction with rhodium acetate and a diazo compound.

injection into cells, the ratiometric biosensor was successfully used to image the activation of endogenous Cdc42 under physiological conditions.[33] More recently, Rauh et al. have developed a new screening assay system for kinase inhibitors using fluorophore-appended kinases.[35,36] Their semisynthetic biosensors were useful to find a drug-like lead compound, which binds to an extra site of the ATP binding pocket.

Bayley et al. developed several stochastic sensors for the detection of organic molecules, DNA/RNA and proteins using an engineered transmembrane protein nanopore of nanometer dimensions in an insulating membrane and by measuring voltage-driven ionic transport through the pore in the presence of analytes.[37] In their earlier works, β-cyclodextrin (β-CD) adapters immobilized noncovalently within the lumen of the nanopore formed by α-hemolysin (α-HL), providing a substantial but incomplete blockage of the channel.[38] To eliminate the background signals caused by the dissociation of the adapters, Bayley's team attached β-CD covalently within the α-HL pore in stable orientations by disulfide bond formation with the cysteine residues of α-HL (Scheme 5d). Thereafter, oligo-DNA and DNA aptamers were conjugated to α-HL for bioanalysis at a single-molecule level.

2.2.2. *Other Examples of Protein Engineering by Using Bioconjugation*

Francis and Stephanopolous have reported the unique applications of several engineered proteins that were constructed using the bioconjugation method. A representative example is the engineered virus capsid.[39] The "tobacco-mosaic-virus coat protein monomer" building blocks were modified by attaching various maleimide-tethered functional molecules, including PEGs, drugs, fluorophores, and porphyrins to cysteine residues (Scheme 6a). Modified virus capsids of this type are attractive nanomaterials for the targeted delivery of therapeutic and imaging agents, light-harvesting systems, and singlet oxygen generation. Furthermore, the Francis group developed protein–polymer hybrid materials for water remediation,[40,41] by using metallothionein scaffolds, which are cysteine-rich proteins capable of binding toxic metal ions. The resulting bioconjugates retained their selectivity and effectively removed trace amounts of toxic metal ions from ocean water with concurrent polymer gel shrinking.

The bioconjugation reaction is also powerful for the construction of semi-synthetic protein mimics involving post-translational modifications. Davis et al. created semi-synthetic glycosylated and sulfonated protein mimics by using two orthogonal bioconjugation strategies, including azide–alkyne click chemistry and disulfide formation on a cysteine residue (Scheme 6b).[42] Their reports clearly showed that the introduction of appropriate chemical modifications with suitable spatial arrangements could mimic the function of a native protein precisely.

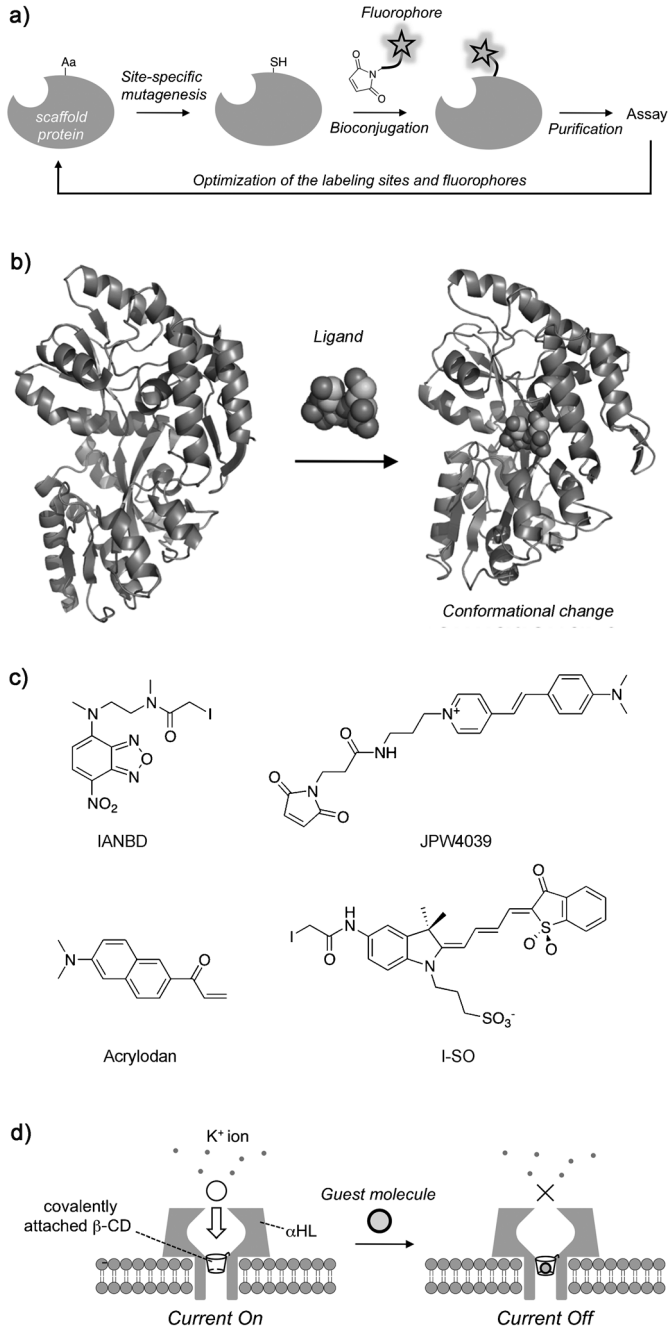
Trauner and co-workers reported a unique approach for the photo-switching of channel-protein activity on live cell surfaces using an azobenzene optical switch conjugated to a cysteine residue by a Michael addition (Scheme 6c). Their method could potentially be applied to a variety of different channel proteins, including the ionotropic glutamate receptor[43] and the potassium channel.[44,45] Furthermore, it could be used to manipulate the neuronal activity with light even in living zebrafish, as well as in cultured cells.

3. Specific Protein Labeling in Cells and In Vivo

3.1. Bioorthogonal Organic Reactions for Protein Labeling

3.1.1. *Bioorthogonal Reactions with Ketones and Aldehydes*

Although the bioconjugation methods outlined in Section 2 for natural amino acids are indeed very useful for protein engineering in vitro, the modification of a specific target protein in living cells containing a variety of other non-target proteins and reactive biomolecules represents a significant challenge. In essence, for the chemical modification of a protein in a crude biological system with a high degree of

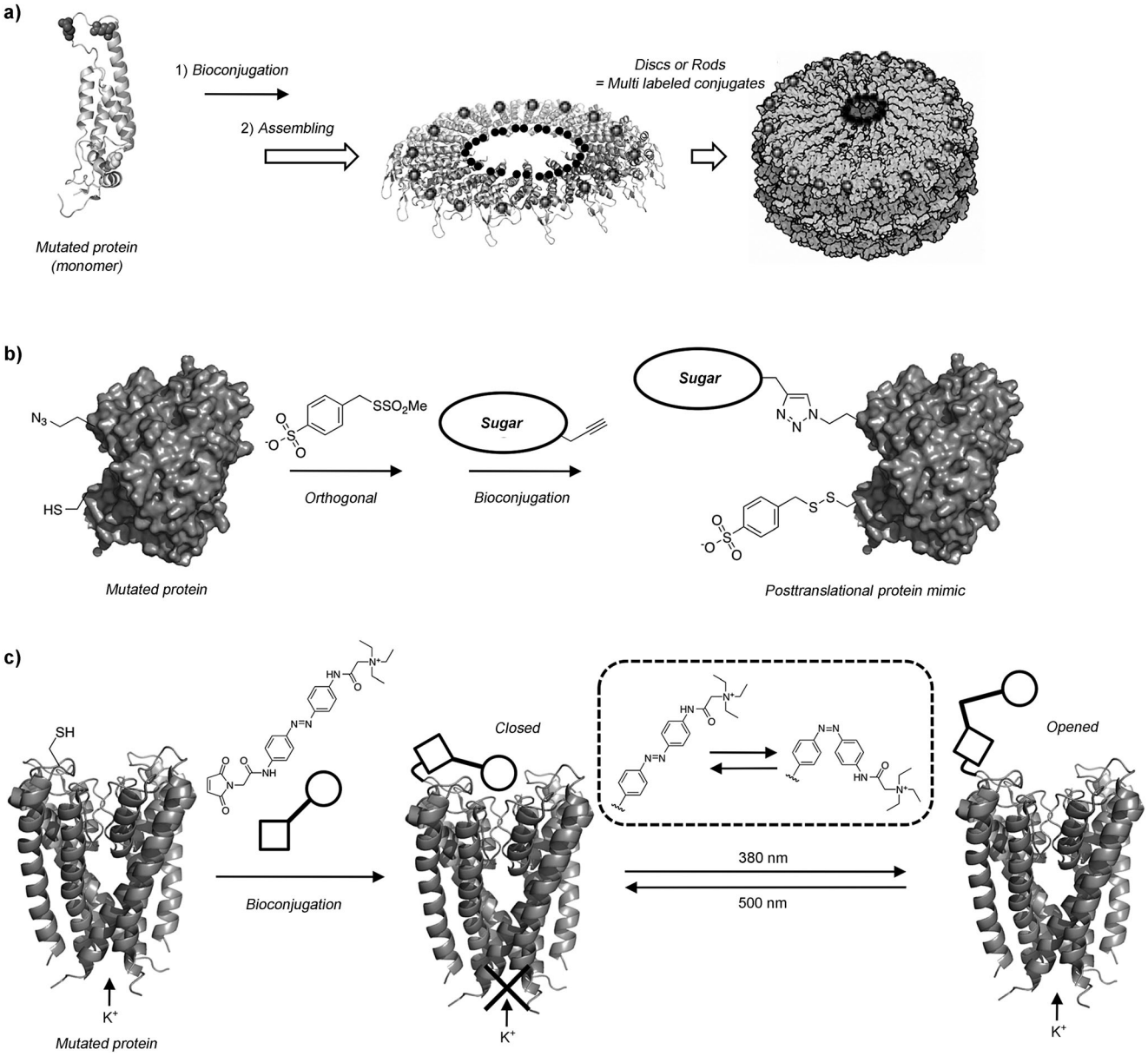


Scheme 5. a) Schematic illustration of the typical strategy for constructing a semisynthetic protein biosensor using the bioconjugation method. b) Conformational changes of the maltose binding protein (MBP, Protein data bank (PDB)code: 1OMP and 1ANF) upon ligand binding. c) Chemical structures of microenvironmentally sensitive and reactive fluorophores for the construction of semisynthetic biosensors. d) Stochastic sensors for the detection of single organic molecules based on β-CD-conjugated α-hemolysin, reported by Bayley and Cremer.[37]

specificity, a pair of biologically rare reactive groups that could react exclusively with each other would represent a useful scenario. These types of bioorthogonal reactions between non-natural reactive groups have been actively reported in recent years.[2] Ketone and aldehyde groups incorporated into a protein can serve as selective reactive handles because they do not usually exist in proteins (Scheme 7a). To date, several methods have been developed for the incorporation of ketone and aldehyde groups into the proteins of interest.[46–48] For instance, Meijler et al. reported the successful installation of a ketone group on a bacterial receptor in a living cell system using the site-specific reaction of a reactive ligand mimic.[49] Bertozzi et al. reported an enzymatic method for the generation of an aldehyde group on a genetically encoded short peptide tag using a formylglycine generating enzyme—the so called aldehyde tag.[50] They also demonstrated the utility of this method in the specific modification of recombinant proteins produced in mammalian cells. It is well established that ketone and aldehyde groups can bio-orthogonally form oximes or hydrazones by reaction with the corresponding hydroxylamine or hydrazine derivatives, respectively, allowing for the introduction of a wide variety of functional molecules into the protein scaffolds which have been so modified.

3.1.2. Bioorthogonal Reactions with Azides and Alkynes

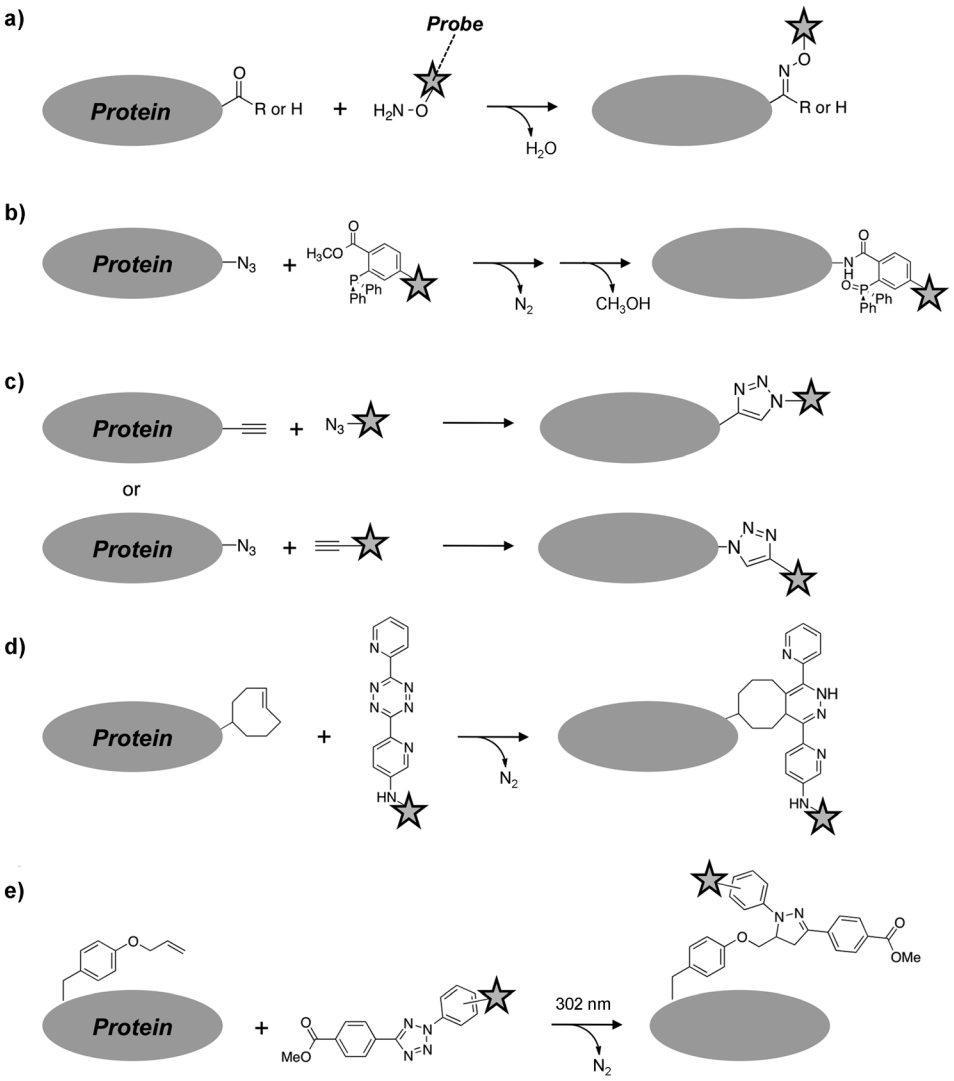
Living organisms do not generally employ any azide compounds. In addition, the azide group has a high biocompatibility because of its low reactivity with biological substances and low toxicity. Conjugate reactions with an azide therefore represent an ideal method of protein labeling, providing a high level of bioorthogonality. For example, in their pioneering work, Bertozzi et al. applied a classical reduction reaction of alkyl azide, catalyzed by triarylphosphine (the Staudinger reaction), for protein modification in a cellular environment (Scheme 7b).[51] They found that proteins bearing an alkyl azide group readily undergo a ligation reaction with triarylphosphine derivatives which have an ester group on their aromatic rings. This process is known as the Staudinger–Bertozzi ligation reaction and has been widely applied for probing biomolecules, especially glycoproteins, in a variety of different cellular environments,[52] as well as in living mice.[53] Another versatile reaction well suited to protein modification and that harnesses the high bioorthogonality of the azide group is the azide–alkyne [3+2] cycloaddition which is otherwise known as the Huisgen cycloaddition (Scheme 7c). The utility of this reaction was first demonstrated by Sharpless and Finn et al. as part of their 'click chemistry' research.[54,55] Meldal et al. separately developed this reaction for site-specific peptide modification.[56] The reaction can be conducted at room temperature to afford a triazole compound by the catalytic action of a Cu^{I} salt. Bertozzi et al. have recently reported a copper-free Huisgen cycloaddition reaction using strained cyclooctyne reagents,[57,58] thus avoiding the use of toxic Cu^{I}. This improvement significantly expanded the utility of this ligation reaction in complicated biological systems, including cultured cells,[59] zebrafish embryos,[60] as well as living mice.[61]



Scheme 6. Typical examples of applications using the bioconjugation methods. a) Self-assembly of bioconjugated tobacco-mosaic-virus (TMV) coat proteins (PDB code: 1EI7 and 3 KML) into disk and rod structures, reported by Francis et al.[39–41] b) Construction of posttranslational modified protein-mimics by using orthogonal bioconjugation reactions, reported by Davis et al.[42] (PDB code: 1GOW). c) Photoswitchable affinity label (PAL) for imparting light sensitivity onto native K^+ channels, reported by Trauner et al.[43] (PDB code: 2A9H).

3.1.3. Other Reactive Groups for Bioorthogonal Reactions

Following on from the successes outlined above, a variety of other elaborated bioorthogonal reactions using non-natural reactive handles have been actively devised by many groups.[62–64] Fox et al. reported a protein bioconjugation based on the inverse-electron-demand Diels–Alder reaction between a strained *trans*-cyclooctene and a tetrazine compound.[65] This reaction proceeded at an unusually fast rate (typically within 5 min at room temperature) to quantitatively afford a chemically modified protein (Scheme 7 d). Most recently, Bertozzi et al. reported the application of another type of cycloaddition reaction for protein modification, known as the quadricyclane ligation.[66,67] They showed that the [2+2+2] cycloaddition reaction between the highly strained quadricyclane and Ni bis(dithiolene) proceeded rapidly in aqueous environments at a rate comparable to that of the cyclooctyne–azide cycloaddition. They also demonstrated that the quadricyclane ligation was compatible with and orthogonal to the strain-promoted azide–alkyne cycloaddition and oxime ligation reactions.[67] Palladium-catalyzed cross-coupling reactions, such as the Suzuki–Miyaura coupling represent another useful tool for protein modification. Davis et al. contributed to the development of palladium chemistry on protein surfaces and demonstrated the use of 2-amino-4,6-dihydroxypyrimidine as a convenient



Scheme 7. Bioorthogonal reactions for selective protein labeling. a) Reactions of aldehydes and ketones with hydroxyamines. b) Staudinger ligation of azides and triarylphosphines. c) Click reactions of alkynes and azides. d) Inverse-electron-demand Diels–Alder reaction of *trans*-cyclooxtenes and dipyridyl tetrazines. e) Photoclick reactions of *O*-ally-tyrosine and diaryltetrazoles.

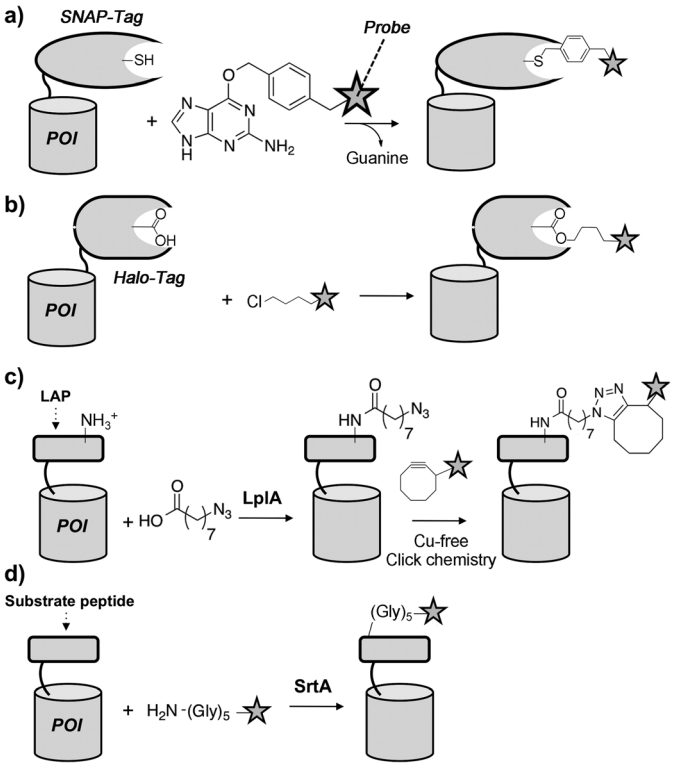
catalyst for the Suzuki–Miyaura coupling reaction of 4-iodophenylalanine displayed on protein surface.[68,69] Most recently, Geierstanger et al. reported a unique site-specific protein-modification method using the non-natural amino acid pyrroline-carboxy-lysine (Pcl).[70] They showed that Pcl can be genetically incorporated into a broad range of target proteins with a high degree of efficiency using the amber codon (TAG). The reversible bioorthogonal conjugation reactions of Pcl with aminobenzaldehyde and 2-aminoacetophenone reagents proceeded efficiently at neutral pH to afford the corresponding functionally modified proteins.

Photoinduced organic reactions provide a unique tool for enabling the spatial and temporal control of protein labeling processes, especially in cell systems. Lin et al. reported the photoinduced 1,3-dipolar cycloaddition reactions of alkenes with tetrazole derivatives, otherwise known as "photoclick chemistry" (Scheme 7e).[71] The application of UV-light irradiation induced the formation of a nitrile imine from the tetrazoles, which subsequently underwent a cycloaddition reaction to afford a pyrazole cycloadduct. They also reported the successfully application of this reaction to a protein bearing unnatural amino acids such as *O*-allyltyrosine inside *E. coli* cells, as well as in buffer solutions. Recently, Popik et al. reported a photo-triggered click reaction between an azide and strained cycloalkyne.[72]

3.2. Selective Protein Labeling by Using an Enzyme or Peptide Tag Binding to a Small Molecular Probe

3.2.1. Enzyme Tags for Protein Modification

Most of the bioorthogonal reactions described above invariably require special pre-treatments to decorate the target protein with biologically unique non-natural functional groups, such as azide, alkyne, and ketone groups. These modifications are based on metabolic labeling, genetic incorporation, and site-selective bioconjugation. The application of such laborious and cumbersome pretreatments can sometimes reduce the protein selectivity and restrict the applicable cell lines. With this in mind, an alternative strategy capable of directly modifying proteins consisting of natural amino acids would be desirable. A number of enzyme-catalyzed protein-labeling methods have been developed recently for this purpose on the basis of the strict enzyme–substrate recognition, with Johnsson et al. providing some pioneering work in this area. This method employs a mutant of human O^6-alkylguanine–DNA alkyltransferase (hAGT) as an enzyme tag, which specifically reacts with an O^6-benzylguanine derivative to afford a covalently labeled protein (Scheme 8a).[73] This labeling method is known as SNAP-tag technology and is currently widely used for a variety of biological applications.[73–75] Another practically useful enzyme tag, known as HaloTag technology, was developed by Promega (Scheme 8b).[76] This technology has also been widely applied not only for specific protein labeling in a cellular context, but for protein purification,[77] immobilization,[78] and incorporation of degradation tags into target proteins,[79] demonstrating its versatility as a biological research tool. In contrast to these techniques, the substrate-conjugation strategy is another type of enzyme-catalyzed labeling method, which uses a short peptide or small protein domain fused to a target protein as an enzyme substrate. Ting et al. reported the probe incorporation mediated by enzymes (PRIME) protein-labeling system



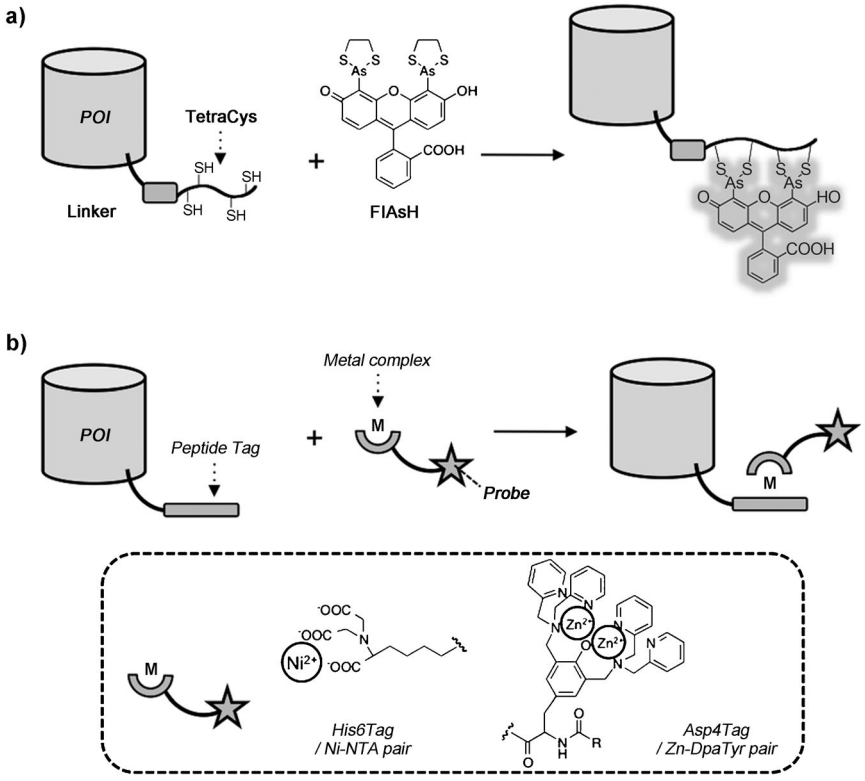
Scheme 8. a,b) Schematic illustration of the enzyme tags for recombinant protein labeling. a) SNAP-tag with benzyl guanine probe. b) Halo-tag with alkyl chloride probe. c,d) Schematic illustration of the enzymatic protein labeling by fusion with a peptide tag. c) Lipoic acid ligase (LplA) with the acceptor peptide (LAP). d) Sortase A (SrtA) with the substrate peptide. POI = protein of interest.

catalyzed by lipoic acid ligase (LplA, Scheme 8c).[80,81] They created a mutant LplA capable of recognizing a coumarin-type substrate and catalyzing its covalent conjugation to a 13 amino acid peptide LplA acceptor peptide (LAP) tag. They successfully demonstrated that this fluorescence labeling process occurred in living cells within 10 min and was highly specific for the LAP tag-fused protein over all other endogenous mammalian proteins. Several other enzyme-catalyzed labeling methods have been developed (Scheme 8d) and their applications demonstrated in live-cell protein labeling as well as in vitro.[82–84]

3.2.2. *Short Peptide Tags for Protein Modification*

Unfortunately, there are potential drawbacks to the enzymatic labeling methods in that they can have an adverse impact on protein functions because of the large molecular weights of the enzyme units (hAGT, 20 kDa; HaloTag dehydrogenase, 33 kDa) which are comparable to those of green fluorescent proteins (GFPs; 27 kDa). There have been several reports describing the disruption of correct protein assembly that was induced by the fused GFPs in a bioimaging study.[85] To avoid this problem, another labeling strategy has been proposed using a complementary recognition pair of a short peptide tag and a small molecular probe. In a pioneering piece of work conducted in this field in 1998, Tsien et al. demonstrated that probes containing two arsenic atoms (a biarsenical probe), such as FlAsH selectively binds to a tetracysteine motif (CCXXCC) with a remarkably strong binding affinity ($100\ \text{pM} > K_d$).[86] Furthermore, the binding induced large fluorescent enhancement of the probe (Scheme 9a). This labeling method was also successfully applied to the fluorescence bio-imaging study of proteins inside and outside living cells,[87] including the determination of protein lifetime by multicolor fluorescence pulse-chase[88] and chromophore-assisted light inactivation (CALI) for the spatiotemporal inactivation of proteins.[89] More recently, the same group have developed a biarsenical calcium indicator, known as Calcium Green FlAsH for imaging regional calcium concentration changes in the proximity of the tag-fused proteins in cells.[90] In addition to these bioimaging applications, the FlAsH method has also been used in other studies, including the investigation of protein–protein interactions, monitoring of protein folding, and enzyme activity control in test-tube systems.[91–93]

The oligo-histidine tag (His-tag), originally developed by F. Hoffmann-La Roche as an affinity tag for protein purification,[94] has also been applied for selective protein labeling. Vogel et al. and Piehler et al. independently reported a protein-labeling system using a conventional binding pair of



Scheme 9. Schematic illustration of peptide tag and probe pair systems for recombinant protein labeling. a) Tetra-cysteine-tag with a FlAsH probe, b) His6-tag with a Ni-NTA probe and Asp4-tag with a Zn-DpaTyr probe.

a His-tag and a Ni^{II} nitrilotriacetic acid (NTA) probe (Scheme 9b). The Vogel group demonstrated the feasibility of their approach by binding Ni^{II}-NTA-chromophore conjugates to a ligand-gated ion channel and a G-protein coupled receptor (GPCR), each containing a His-tag sequence. The ionotropic 5-hydroxytryptamine (5HT) receptor was selectively labeled with Ni^{II}-NTA probes for fluorescence measurements to characterize the structure and plasma membrane distribution of the receptor on the surface of live cells.[95] The Piehler group developed an efficient method for the selective and stable noncovalent labeling of His-tag proteins using multivalent tris-Ni^{II}-NTA probes.[96] Fluorescence labeling with tris-Ni^{II}-NTA-fluorophore conjugates was applied for the real-time sensing of multi-protein complex formation in solution and on surfaces. In a separate piece of research, Tsien et al. developed HisZnFiT as a new fluorescent Zn^{II} complex that bound strongly to a His6 tag ($K_d \approx 40$ nM), which was designed to overcome the strong fluorescence quenching effect and cell toxicity caused by the nickel ion of the Ni^{II}-NTA probes.[97] Very recently, Hamachi et al. also reported the use of multinuclear zinc complexes as a selective fluorescent probes for the bio-imaging of His-tag fused GPCR on the surfaces of live cell.[98,99]

In general, the de novo design of complementary recognition systems comprising small molecules and peptides with a strong and selective binding affinity represent a significant challenge. In spite of the issues associated with this strategy, several new tag–probe pairs have been reported in recent years, including a pair comprising a tetraserine motif and boronic acid probe,[100] and a pair comprising a lanthanoid binding tag and Tb^{III} ion.[101] A distinct tag–probe pair based on coordination chemistry was developed by Hamachi et al. The pair was composed of an oligo(aspartate) peptide (D4 tag; DDDD) and a binuclear Zn^{II} complex (Zn^{II}-DpaTyr; Scheme 9b).[102] This pair was successfully applied to the fluorescent bio-imaging study of a (GPCR) protein fused to a triply repeated D4-tag [$(D4)_3$] expressed on the cell surface. Following on from this work, Hamachi et al. further advanced their covalent protein-labeling system based on this tag–probe pair, developing a tag and a probe that were elaborately designed to undergo a specific nucleophilic reaction in the binding complex between the cysteine residue incorporated into the D4 tag and the α-chloroacetyl group of the Zn^{II}-DpaTyr probe.[103] The rapid labeling (10–30 min) of GPCRs such as the bradykinine B2 receptor (B2R) and the acetyl choline receptor (M1R), proceeded selectively with Zn^{II}-DpaTyr probes bearing a variety of different functional groups,[104] and enabled the fluorescence visualization of the GPCRs on surface of HEK293 cells and any shifts in their localization.

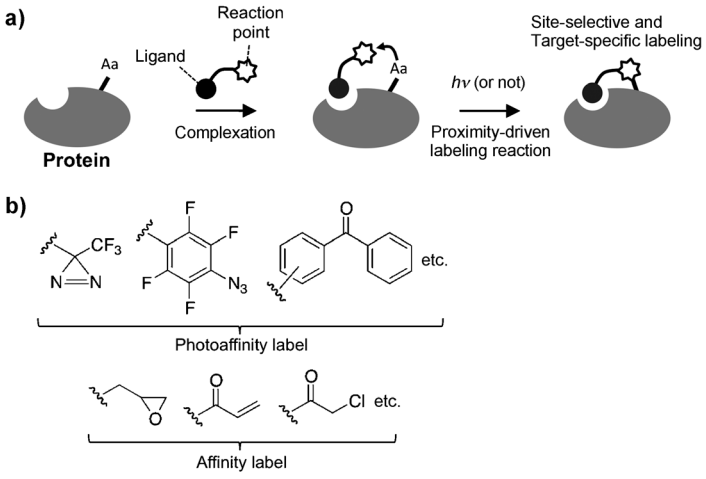
These non-enzymatic labeling methods have several potential advantages over the enzyme-catalyzed methods in that: 1) there is no need to incorporate a large enzyme or protein domain into the target proteins; 2) a wide range of labeling reagents are available without the structural restrictions that they must be suitable as enzyme substrates; and 3) the labeling conditions are independent of the enzyme properties.

4. Selective Endogenous Protein Labeling in Live-Cell Systems

By using the bioorthogonal reactions or the enzyme/short-peptide tags mentioned above, specific protein labeling events can be achieved even under the crude conditions experienced in cells and in vivo. The incorporation of a reactive handle (bioorthogonal non-natural reactive handle or enzyme/peptide tag) into the target proteins using genetic manipulation, however, is a necessary pre-requisite for the application of this strategy. Consequently, these target proteins are no longer endogenous, but over expressed. The specific labeling of a natural protein that is endogenously expressed in the typical context of a living system represents an ideal scenario for precise and quantitative intracellular protein analysis. The affinity labeling method provided us with an important clue as to how to achieve the specific labeling of an endogenous target protein. Several chemistry-based methods have recently been proposed in this particular context.[105] Such methods may come to be used in selective molecular therapy in future medicinal applications.

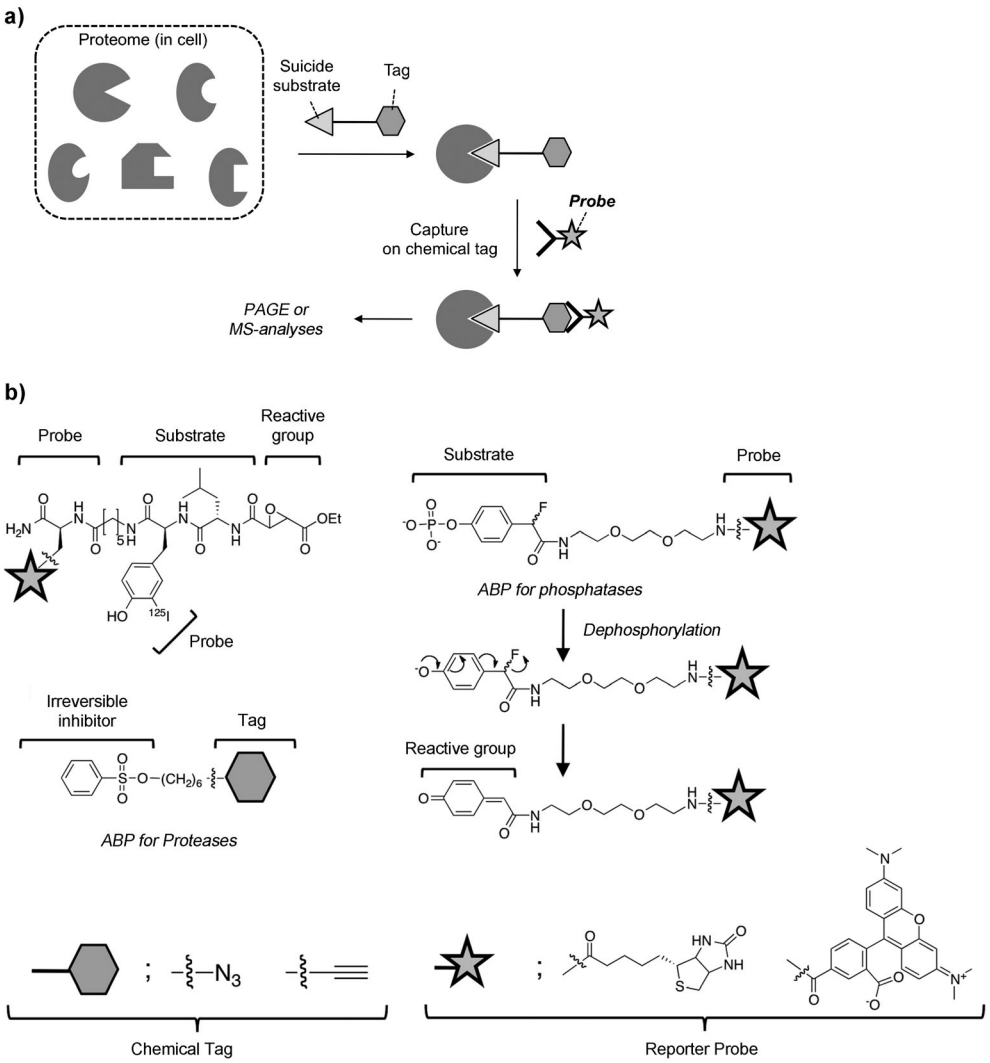
4.1. Affinity-Based Labeling Method for Selective Endogenous Protein Labeling

Affinity labeling, in which the protein of interest is labeled with a reactive group attached to a ligand, has been widely used for the identification and characterization of endogenously expressed proteins (Scheme 10).[106–109] By coupling this technology with modern mass-spectrometry techniques, additional information can be acquired regarding the structure and function of target proteins, such as the identification of the exact amino acid residues in the active sites and the examination of the stoichiometry and site-specificity of biomolecular interactions. Since the first report of affinity labeling in the early 1960s, significant efforts have been invested in the fine tuning of the chemical structures used in



Scheme 10. a) Schematic illustration of the (photo)affinity labeling method for target-specific protein labeling. b) Chemical structures of the reactive groups widely used for (photo)affinity labeling.

reactive-group-enabled protein labeling, culminating in excellent levels of target selectivity in the whole proteomes of living cells. The most commonly used affinity labeling methods are based on several photoactivatable groups such as diazirine, phenylazide, and benzophenone. Although these photo-triggered reactions can occur in all of the natural amino acids present on the surfaces of proteins, labeling efficiency is generally low because of the short lifespan of the reactive carbene or nitrene species in aqueous solution. Instead of photo-affinity labeling methods, recent progress has been made using thermally driven chemical reactions that require relatively moderate levels of reactivity that are suitable for effectively providing high levels of selectivity to the protein-labeling process. Sames et al. reported a new affinity labeling reagent based on an epoxide group that provide highly efficient and site-selective labeling of purified human carbonic anhydrase II (hCAII) both in vitro and in cell lysates.[110] Systematic screening revealed that the epoxide functionality had the special combination of stability and reactivity necessary to make it stable towards proteins in solution but reactive in the proximity of the active site.



Scheme 11. a) Schematic illustration of activity-based proteomic profiling (ABPP). b) Chemical structures of typical activity-based probes (ABP).

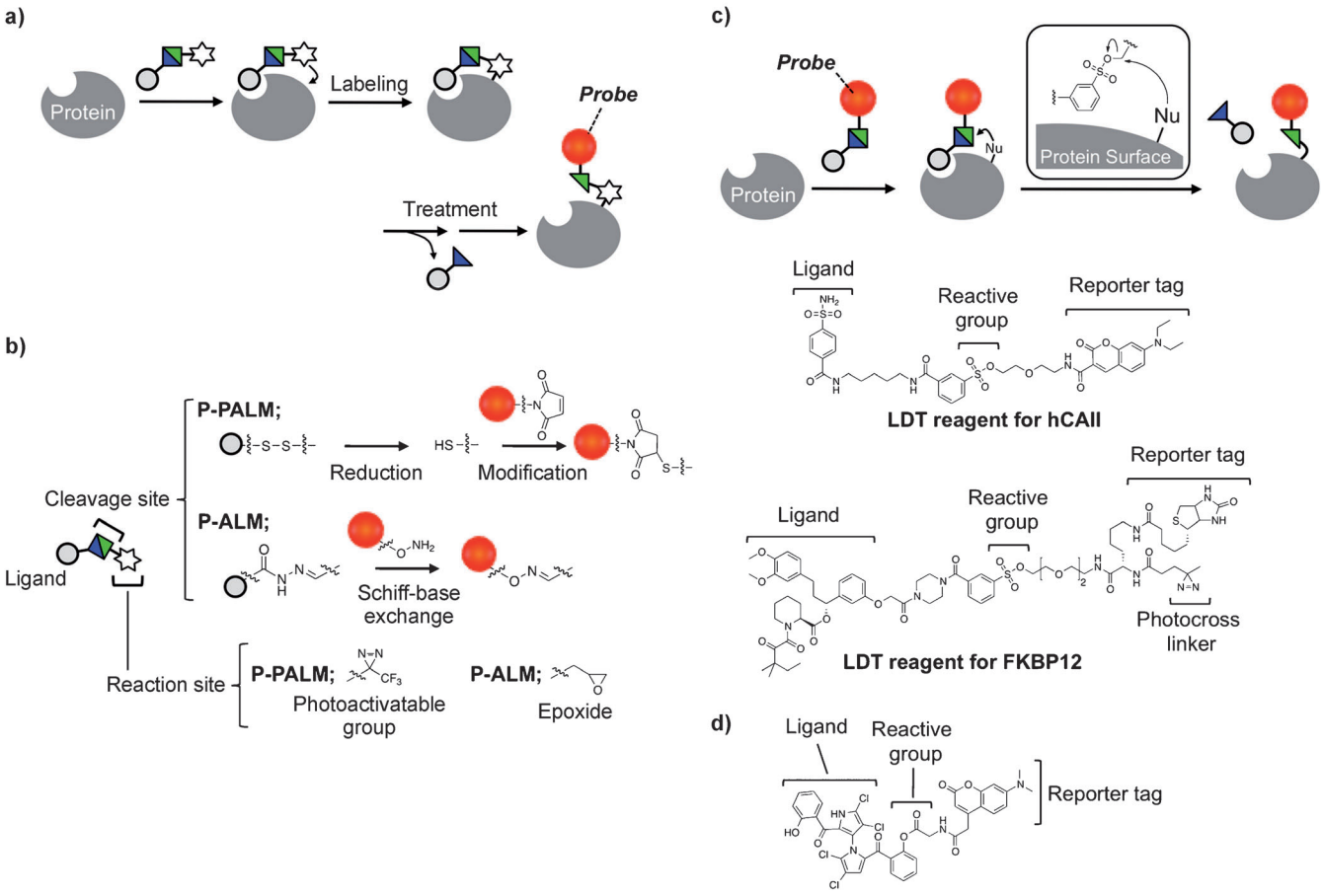
4.2. Activity-Based Protein Labeling for Proteomic Profiling

Based on the affinity labeling method, activity-based probes (ABP) were developed using suicide substrates for the corresponding enzymes.[111,112] By greatly improving the ABP, Cravatt et al. and Bogyo et al. independently demonstrated activity-based protein profiling (ABPP) as a powerful proteomic strategy for the direct characterization of the "active form" of an enzyme under native biological habitats (Scheme 11).[108,113,114] ABPP reagents are generally designed on the basis of them behaving as suicide substrates for the target enzymes, and generally consist of three basic elements: 1) a reactive group, which binds and covalently modifies the active sites of the enzymes; 2) a linker region; and 3) a reporter tag, such as a fluorophore or biotin molecule, for the detection and isolation of probe-labeled enzymes from proteomes. Subsequently, Cravatt et al. reported the application of a bioorthogonal chemical tag for ABPP which eliminated the need for the detection tag that usually restricted the incorporation efficiency into cells and intracellular distribution of the reagents. They initially used an azide-functionalized phenyl sulfonate that targeted serine hydrolases.[115] The azide probes were introduced into live cells, where the active enzymes were covalently modified. Cell lysates were generated and subsequently treated with an alkyne–rhodamine tag using copper-catalyzed click reactions, allowing for several enzyme classes including glutathione *S*-transferase, enoyl CoA hydratase and aldehyde dehydrogenase to be analyzed and clarified.[116–118] Bogyo et al. also explored various ABPP probes for imaging of cysteine proteases in live-cell systems.[119] During the past decade, several research groups designed a number of ABPP probes as active-site-directed covalent probes allowing for the

discovery of many different enzyme classes in complex proteomes, including mettalloproteases,[120] kinases,[121] phosphatases,[122] glycosidases,[123] and oxidoreductases.[124] More recently, Cravatt et al. also reported the use of 1,2,3-triazole urea as a versatile chemical species for the selective inhibition of serine hydrolases.[125] These 1,2,3-triazole urea derivatives exhibited their inhibitory activities through the covalent carbamoylation of the reactive serine residues of the hydrolase enzymes. They clearly demonstrated that rapid lead optimization and competitive activity-based profiling were carried out by ABPP for identifying selective inhibitors for a diverse series of serine hydrolases, including peptidases, lipases, and uncharacterized hydrolases, with high potency in cells and mice. This work represents an intriguing example of the exploitation of an endogenous protein-labeling reaction to discover a mechanism-based irreversible inhibitor. In an attempt to extend the scope of the ABPP strategy, Yao et al. modified the method by introducing a photo-crosslinkable moiety to the ABPP reagents.[126,127]

4.3. Ligand-Directed and Traceless Chemical Labeling for Protein Engineering in Live Cells

Although the affinity labeling methods including the ABPP approaches are powerful for the specific labeling of proteins and proteomic analysis under miscellaneous conditions, the associated probes usually inactivate the target proteins, limiting the value of any potential applications, such as the real-time monitoring of protein activities in cells. For any subsequent functional studies of the labeled proteins, the restoration of the original functions of the protein is essential even after labeling. Hamachi et al. developed several new approaches for achieving such traceless protein labeling.[128] In their earlier work, disulfide or hydrazone linkers were used as cleavage sites for the removal of the ligand following protein labeling (Scheme 12a,b). The disulfide linkers could be cleaved by dithiothreitol treatment (termed post-photoaffinity labeling method, P-PALM),[129,130] whereas the hydrazone linkers could be converted into a non-obstructive oxime using an excess of hydroxylamine (termed post-affinity labeling method, P-ALM).[131,132] These improvements in affinity labeling, however, are still insufficient for the application of the technique in living systems because of their low bioorthogonality. Recently, Hamachi et al. developed a new labeling strategy, known as ligand-directed "tosyl" (LDT) chemistry, in which the phenylsulfonate (sometimes termed "tosyl") linker behaves both as a linker between the ligand and the probe and as a reactive group (Scheme 12c).[133,134] A typical S_N2-type reaction between the phenylsulfonate moiety and a (natural) nucleophilic amino acid residue resulted in the release of the ligand moiety during



Scheme 12. a) Schematic illustration of the post-(photo)affinity labeling method (P-(P)ALM). b) Molecular design of the P-PALM and P-ALM reagents. c) Schematic illustration of the ligand-directed "tosyl" (LDT) chemistry and chemical structures of LDT reagents for hCAII labeling and for FKBP12 engineering. d) Chemical structure of acyl phenol-type marrinopyrrole-tethered labeling reagent.[140]

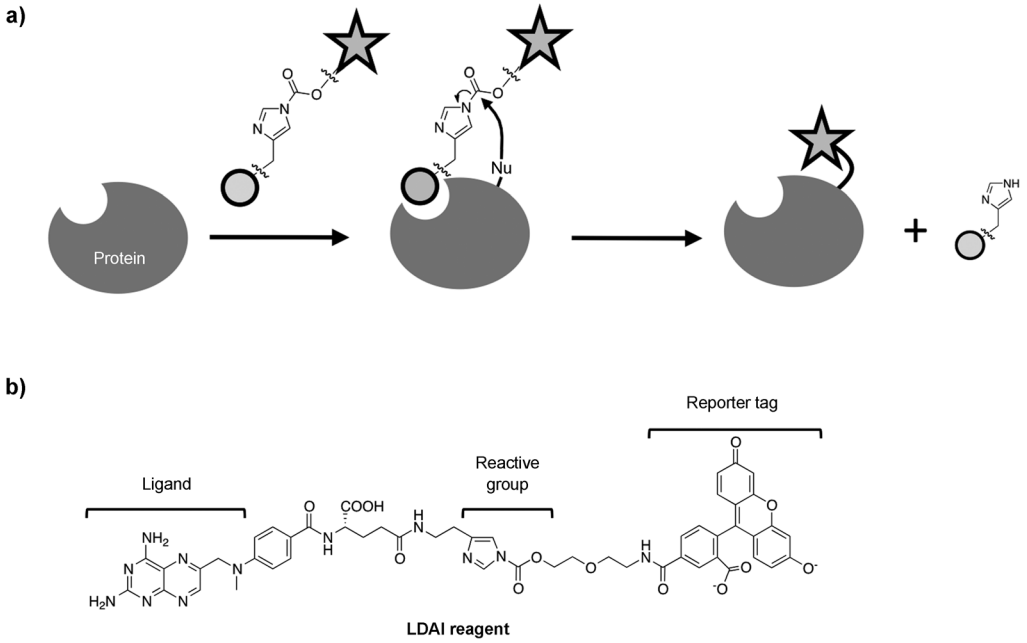
the labeling reaction and the retention of the protein activity. This method allowed for the execution of efficient and extremely specific protein-labeling reactions in live cells, on tissue, and even in live mice, as well as in vitro. Taking advantage of LDT chemistry, endogenous proteins can be converted into protein-based biosensors inside cells. In a proof-of-principle experiment, an in-cell ^{19}F NMR spectroscopy biosensor was constructed in human red blood cells (RBCs) on the basis of an endogenously expressed carbonic anhydrase I (CAI),[133,135] which can sense the ligand binding to CAI through a change in the chemical shift of a ^{19}F signal. More recently, the ligand-bound and unbound structures of ^{19}F-labeled CAI were fully characterized by X-ray crystallography, and an unprecedented quantitative comparison of the protein's dynamics was conducted in RBCs and in vitro using ^{19}F exchange (EXSY) NMR spectroscopy.[136]

Given the benefits of the modular nature of the LDT reagents, it is envisaged that this labeling method will be applicable to the specific modification of a variety of different endogenous proteins under many cellular environments. For example, the FK506-binding protein 12 (FKBP12) in mammalian cells was successfully labeled using optimized LDT reagents containing a photoreactive probe, and the UV-induced covalent cross-linking of FKBP12 with its interacting proteins (FRB or calcineurin) was achieved in living cells (Scheme 12c).[137] Recently, other researchers have utilized LDT chemistry for selective protein labeling and characterization. Shin et al. recently used LDT chemistry to demonstrate that the small-molecule apoptozole (Az) was bound to the ATP-binding pocket of the molecular chaperon heat shock cognate 70 (Hsc70).[138] Ohkanda et al. also reported the phosphopeptide-dependent site-selective fluorescent labeling of 14-3-3δ protein using an LDT reagent tethering a diterpene fusicoccin A (FC) ligand.[139] These two results clearly demonstrated that the power of LDT chemistry for characterization of both the target protein and ligand-binding site, with the success reflecting the strict site-selectivity of the technique.

Other methods for chemically labeling endogenous proteins in a cellular context have also been reported using distinct reactive groups with carefully tuned levels of reactivity. For example, Fenical et al. reported the use of a so-called acyl phenol (that is, a 2-hydroxybenzoyl) moiety as a reactive group for identifying the target protein of marrinopyrrole A, which is a natural product with anticancer properties in live cells (Scheme 12d).[140] Hamachi et al. very recently reported a new type of ligand-directed chemistry, known as ligand-directed acyl imidazole (LDAI) chemistry (Scheme 13).[141] A moderately reactive alkyloxyacyl imidazole (AI) assisted by ligand binding enabled the selective modification of natural proteins, such as the endogenous folate receptor (FR) on the surface of live cells, which could not be efficiently labeled using the LDT chemistry. Interestingly, the fluorescein-labeled FR was shown to work as a fluorescent biosensor on the live-cell surface, which allowed for the unprecedented in situ ligand-binding kinetic analysis of the FR.

4.4. Ligand-Tethered Catalysts for Traceless Protein Labeling on the Cell Surface

Another strategy for endogenous protein labeling was recently reported by the Hamachi group, they used a catalyst decorated in such a way as to achieve specific affinity for a target protein.[142] 4-Dimethylaminopyridine (DMAP) has been widely used as an organocatalyst in a variety of acyl-transfer reactions in organic synthesis. The Hamachi group



Scheme 13. a) Schematic illustration of the ligand-directed acyl imidazole (LDAI) chemistry; Nu = nucleophile. b) Chemical structure of the LDAI reagent for FR labeling.

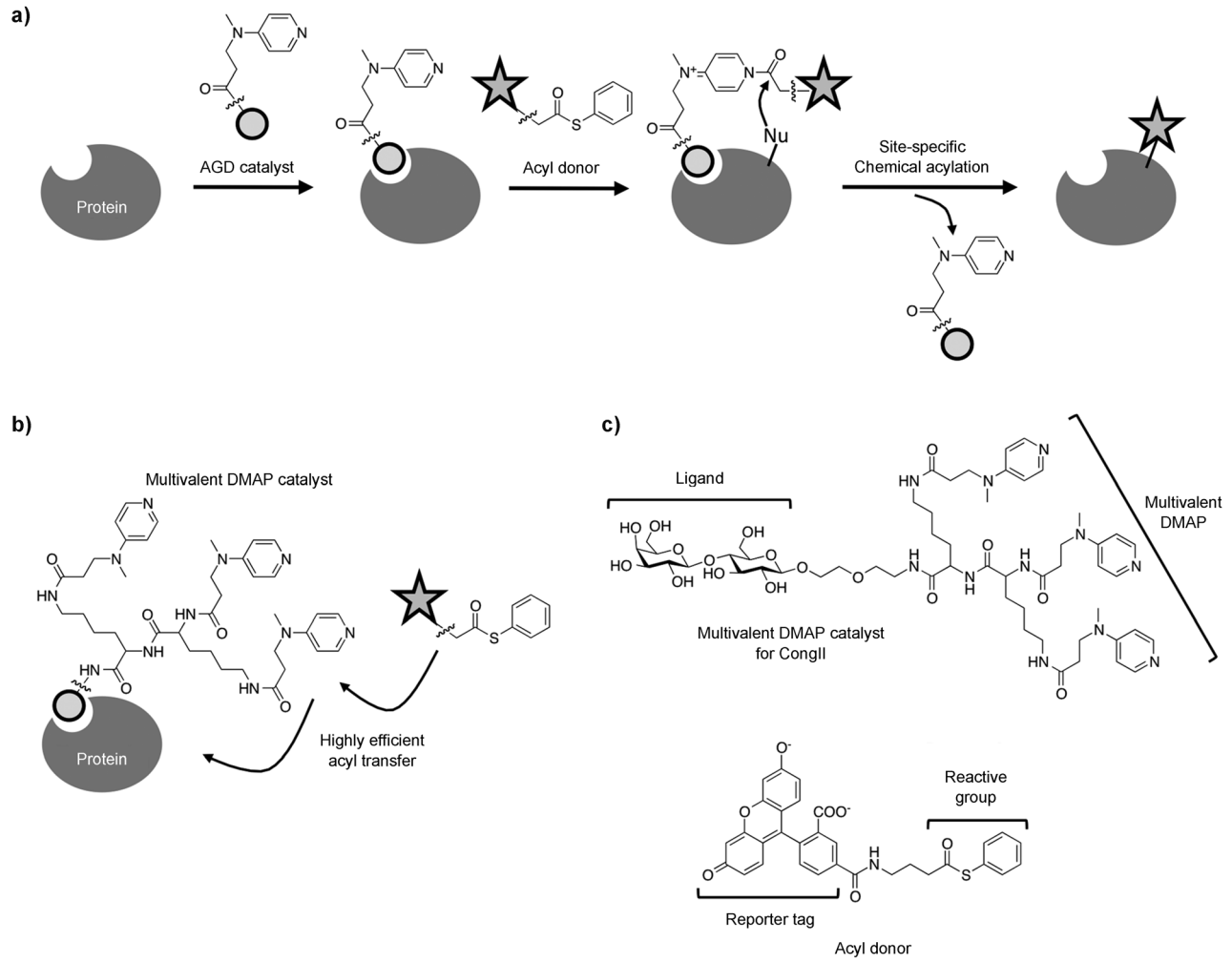
designed affinity-guided DMAP (AGD) catalysts consisting of a DMAP group and an affinity ligand (Scheme 14a,b). Selective binding of the AGD catalyst to a target protein facilitated the acyl transfer reaction of a thioester probe in the proximity of the protein active site. The validity of this chemistry was demonstrated using lectins, sugar-binding proteins, such as Congerin II (CongII), Concanavarin A (ConA), and wheat germ agglutinin (WGA), and the ligand-binding proteins, such as the SH2-domain and FKBP12.[142–144] More recently, the realization was made that the labeling reaction could be significantly accelerated by increasing the number of DMAP groups in the ligand–catalyst complex.[144] Given the high efficiency and selectivity of multivalent AGD-based protein labeling, the labeling of B2R was conducted on the surface of a live cell (85% yield within 30 min). Based on the fluorescently labeled B2R, the ligand-binding assay was performed using a bimolecular fluorescence quenching and recovery (BFQR) system on the live-cell surfaces. To date, approximately 50% of all modern drugs are targeted to GPCRs. With this in mind, the development of an efficient and robust drug-screening platform for GPCRs, is of broad significance to pharmaceutical chemistry and drug discovery.

Other types of affinity-guided catalyst have been prepared, Kunishima reported the use of a biotin-tethered chlorotriazine for avidin labeling (Scheme 15a).[145] Ball et al. recently reported a new labeling method using rhodium(II) metallopeptides and functionalized diazo compounds (Scheme 15b).[146,147] Site-specific, proximity-driven protein modification can be carried out using a unique combination of peptide-based molecular recognition and a rhodium catalyst capable of modifying a wide range of amino acid side chains in cell lysates as well as in vitro.

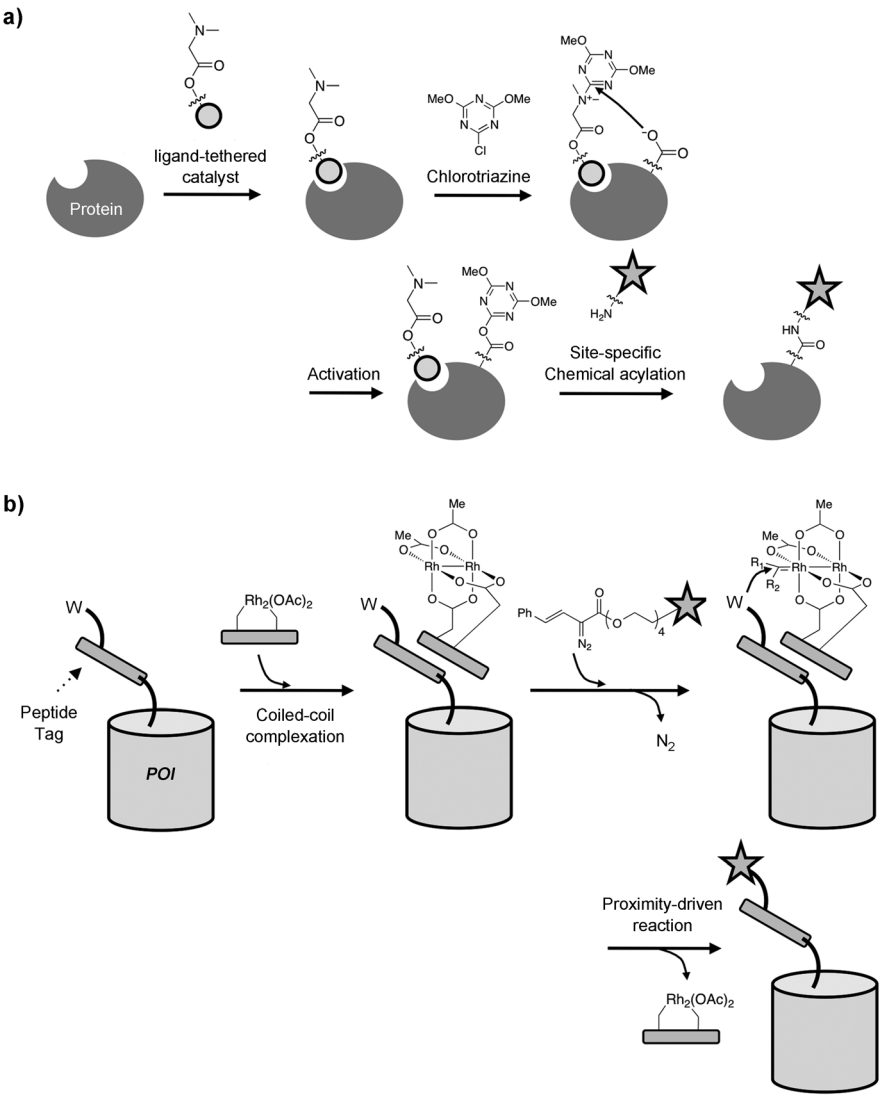
5. Summary and Outlook

In summary, recent efforts have provided important and significant advances in the field of selective protein modification with synthetic molecules. Four main strategies have been reviewed in this article:

1) the bioconjugation of natural amino acids (predominantly in vitro)
2) the bioorthogonal reaction of non-natural amino acids



Scheme 14. a,b) AGD catalyst-mediated selective chemical protein labeling. Schematic illustration of a) the basic strategy and b) the multivalent DMAP system for efficient labeling. c) Chemical structures of the lactose-type AGD catalyst for CongII labeling and fluorescein-type acyl donor.



Scheme 15. a) Schematic illustration of the ligand-tethered chlorothiazine for protein labeling.[145] b) Schematic illustration of the proximity-driven protein modification using rhodium(II) metallopeptides and diazo compound.[146,147]

3) the coupling of selective recognition and reactivity using an enzyme or short-peptide tag and a probe label (tag–probe pair)
4) the use of ligand-directed labeling reactions for endogenous proteins

These techniques represent a useful set of tools for use in chemical biology,[148–151] with methods 2–4 in particular been applicable to crude (i.e. living) habitats. The complementary or combined use of these methods should further expand opportunities for highlighting the power of chemistry-based approaches as essential tools in biological investigations. To potentially address some of the currently unsolved and future issues in cellular biology, pharmaceutical research, and the medical sciences, the development of new protein-modification methods in living systems are undoubtedly required. Any future developments should be able to facilitate the analysis of huge protein complexes, identify target proteins for new medicines, and allow for the rational design of irreversible inhibitors, and as well as providing cancer detection and diagnosis in vivo. Research into the use of organic chemistry for manipulating and altering endogenous natural proteins in their live habitats is currently at an early stage in its development (Table 1), and the development of new chemical techniques that can be successfully applied in living systems

Table 1: Ligand-directed reactions for specific protein labeling.

Method	Reactive group	Modified protein	Labeled amino acid	Conditions	Ref.
P-PALM	trifluoromethyldiazirine	ConA	Tyr	Test tube	[129,130]
P-ALM	epoxide	hCAII	His	Test tube	[131]
LDT	phenylsulfonate	hCAI, hCAII, SH2 domain, FKBP12, CongII, 14-3-3δ, Hsc70	His, Tyr, Glu	Test tube Live cell In vivo	[133–139]
LDAI	acyl imidazole	DHFR, FR	Lys	Test tube Live cell	[141]
Acyl dye transfer	acylphenol	actin	Lys	Test tube Live cell	[140]
AGD-DMAP	DMAP /thioester	CongII, ConA, WGA, SH2 domain, FKBP12, FR, B2R	Lys, Tyr	Test tube Cell lysates Live cell	[142–144]
AGD-chlorothiazine	trimethoxychlorothiazine	Avidin	Glu	Test tube	[145]
Rh-diazo	rhodium(II)/diazo	Coiled coils, SH3 domain	Trp, Tyr, Phe, Asn, Gln, Asp, Glu, Arg, Cys	Test tube Cell lysates	[146,147]

represents a worthy challenge to organic chemistry and those who practice the art.

Received: September 1, 2012
Published online: February 20, 2013

[1] A. N. Glazer, *Annu. Rev. Biochem.* **1970**, *39*, 101 – 130.
[2] E. M. Sletten, C. R. Bertozzi, *Angew. Chem.* **2009**, *121*, 7108 – 7133; *Angew. Chem. Int. Ed.* **2009**, *48*, 6974 – 6998.
[3] A. M. Wu, P. D. Senter, *Nat. Biotechnol.* **2005**, *23*, 1137 – 1146.
[4] J. M. Antos, M. B. Francis, *Curr. Opin. Chem. Biol.* **2006**, *10*, 253 – 262.
[5] L. Wang, P. G. Schultz, *Angew. Chem.* **2005**, *117*, 34 – 68; *Angew. Chem. Int. Ed.* **2005**, *44*, 34 – 66.
[6] A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603 – 609.
[7] I. Johnsson, M. T. Z. Spence, *The Molecular Probes Handbook*, 11th ed., **2010**, Part I.
[8] J. M. McFarland, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 13490 – 13491.
[9] R. A. Scheck, M. B. Francis, *ACS Chem. Biol.* **2007**, *2*, 247 – 251.
[10] K. Tanaka, T. Masuyama, K. Hasegawa, T. Tahara, H. Mizuma, Y. Wada, Y. Watanabe, K. Fukase, *Angew. Chem.* **2008**, *120*, 108 – 111; *Angew. Chem. Int. Ed.* **2008**, *47*, 102 – 105.
[11] K. Tanaka, M. Kitadani, K. Fukase, *Org. Biomol. Chem.* **2011**, *9*, 5346 – 5349.
[12] G. J. L. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 5052 – 5053.
[13] Y. A. Lin, J. M. Chalker, B. G. Davis, *J. Am. Chem. Soc.* **2010**, *132*, 16805 – 16811.
[14] C. Denison, T. Kodadek, *J. Proteome Res.* **2004**, *3*, 417 – 425.
[15] T. Kodadek, I. Duroux-Richard, J. C. Bonnafous, *Trends Pharmacol. Sci.* **2005**, *26*, 210 – 217.
[16] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942 – 15943.
[17] T. L. Schlick, Z. Ding, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 3718 – 3723.
[18] S. D. Tilley, M. B. Francis, *J. Am. Chem. Soc.* **2006**, *128*, 1080 – 1081.
[19] J. M. Antos, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 10256 – 10257.
[20] J. M. Antos, J. M. McFarland, A. T. Lavarone, M. B. Francis, *J. Am. Chem. Soc.* **2009**, *131*, 6301 – 6308.
[21] A. Abuchowski, T. V. Es, N. C. Palczuk, F. F. Davis, *J. Biol. Chem.* **1977**, *252*, 3578 – 3581.
[22] A. Abuchowski, J. R. McCoy, N. C. Palczuk, T. V. Es, F. F. Davis, *J. Biol. Chem.* **1977**, *252*, 3582 – 3586.
[23] J. M. Harris, R. B. Chess, *Nat. Rev. Drug Discovery* **2003**, *2*, 214 – 221.
[24] M. J. Roberts, M. D. Bentley, J. M. Harris, *Adv. Drug Delivery Rev.* **2002**, *54*, 459 – 476.
[25] F. M. Veronese, G. Pasut, *Drug Discovery Today* **2005**, *10*, 1451 – 1458.
[26] S. R. Adams, A. T. Harootunian, Y. J. Buechler, S. S. Taylor, R. Y. Tsien, *Nature* **1991**, *349*, 694 – 697.
[27] H. Wang, E. Nakata, I. Hamachi, *ChemBioChem* **2009**, *10*, 2560 – 2577.
[28] M. A. Dwyer, H. W. Hellinga, *Curr. Opin. Struct. Biol.* **2004**, *14*, 495 – 504.
[29] J. S. Marvin, E. E. Corcoran, N. A. Hattangadi, J. V. Xhang, S. A. Gere, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4366 – 4371.
[30] R. M. de Lorimier, J. J. Smith, M. A. Dwyer, L. L. Looger, K. M. Sali, C. D. Paavola, S. S. Rizk, S. Sadigov, D. W. Conrad, L. Loew, H. W. Hellinga, *Protein Sci.* **2002**, *11*, 2655 – 2675.
[31] K. Sugimoto, M. Nishida, M. Otsuka, K. Makino, K. Ohokubo, Y. Mori, T. Morii, *Chem. Biol.* **2004**, *11*, 475 – 485.
[32] R. Sakaguchi, K. Tainaka, N. Shimada, S. Nakano, M. Inoue, S. Kiyonaka, Y. Mori, T. Morii, *Angew. Chem.* **2010**, *122*, 2196 – 2199; *Angew. Chem. Int. Ed.* **2010**, *49*, 2150 – 2153.
[33] P. Nalbant, L. Hodgson, V. Kraynov, A. Toutchkine, K. M. Hahn, *Science* **2004**, *305*, 1615 – 1619.
[34] A. Toutchkine, V. Kraynov, K. Hahn, *J. Am. Chem. Soc.* **2003**, *125*, 4132 – 4145.
[35] J. R. Simard, S. Klüter, C. Grütter, M. Getlik, M. Rabiller, H. B. Rode, D. Rauh, *Nat. Chem. Biol.* **2009**, *5*, 394 – 396.
[36] J. R. Simard, M. Getlik, C. Grütter, V. Pawar, S. Wulfert, M. Rabiller, D. Rauh, *J. Am. Chem. Soc.* **2009**, *131*, 13286 – 13296.
[37] H. Bayley, P. S. Cremer, *Nature* **2001**, *413*, 226 – 230.
[38] L. Movileanu, S. Howorka, O. Braha, H. Bayley, *Nat. Biotechnol.* **2000**, *18*, 1091 – 1095.
[39] N. Stephanopoulos, M. B. Francis, *Nat. Chem. Biol.* **2011**, *7*, 876 – 884.
[40] T. L. Schlick, Z. Ding, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 3718 – 3723.
[41] R. A. Miller, A. D. Presley, M. D. Francis, *J. Am. Chem. Soc.* **2007**, *129*, 3104 – 3109.
[42] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105 – 1109.
[43] M. Volgraf, P. Gorostiza, R. Numano, R. H. Kramer, E. Y. Isacoff, D. Trauner, *Nat. Chem. Biol.* **2006**, *2*, 47 – 52.
[44] D. L. Fortin et al., *Nat. Methods* **2008**, *5*, 331 – 338.
[45] C. Wyart, F. D. Bene, E. Warp, E. K. Scott, D. Trauner, H. Baier, E. Y. Isacoff, *Nature* **2009**, *461*, 407 – 411.
[46] Y. Zeng, T. N. C. Ramya, A. Dirksen, P. E. Dawson, J. C. Paulson, *Nat. Methods* **2009**, *6*, 207 – 209.
[47] A. Banerjee, T. D. Panosian, K. Mukherjee, R. Ravindra, S. Gal, D. L. Sackett, S. Bane, *ACS Chem. Biol.* **2010**, *5*, 777 – 785.
[48] See Ref. [46].
[49] J. Rayo, N. Amara, P. Krief, M. M. Meijler, *J. Am. Chem. Soc.* **2011**, *133*, 7469 – 7475.
[50] I. S. Carrico, B. L. Carlson, C. R. Bertozzi, *Nat. Chem. Biol.* **2007**, *3*, 321 – 322.
[51] Staudinger Reaction: E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007 – 2010.
[52] S. T. Laughlin, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12 – 17.
[53] J. A. Prescher, D. H. Dube, C. R. Bertozzi, *Nature* **2004**, *430*, 873 – 877.
[54] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056 – 2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004 – 2021.
[55] Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192 – 3193.
[56] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057 – 3064.
[57] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046 – 15047.
[58] J. A. Codelli, J. M. Baskin, N. J. Agard, C. R. Bertozzi, *J. Am. Chem. Soc.* **2008**, *130*, 11486 – 11493.
[59] K. E. Beatty, J. Szychowski, J. D. Fisk, D. A. Tirrell, *ChemBioChem* **2011**, *12*, 2137 – 2139.
[60] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664 – 667.
[61] P. V. Chang, J. A. Prescher, E. M. Sletten, J. M. Baskin, I. A. Miller, N. J. Agard, A. Lo, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1821 – 1826.
[62] K. Kodama, S. Fukuzawa, H. Nakayama, T. Kigawa, K. Sakamoto, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, K. Tachibana, S. Yokoyama, *ChemBioChem* **2006**, *7*, 134 – 139. Corrigendum: *ChemBioChem* **2007**, *8*, 159.

[63] A. Ojida, H. Tsutsumi, N. Kasagi, I. Hamachi, *Tetrahedron Lett.* **2005**, *46*, 3301 – 3305.
[64] T. Sasaki, K. Kodama, H. Suzuki, S. Fukuzawa, K. Tachibana, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4550 – 4553.
[65] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518 – 13519.
[66] M. T. Taylor, M. L. Blackman, O. Dmitrenko, J. M. Fox, *J. Am. Chem. Soc.* **2011**, *133*, 9646 – 9649.
[67] E. M. Sletten, C. R. Bertozzi, *J. Am. Chem. Soc.* **2011**, *133*, 17570 – 17573.
[68] See Ref. [13]
[69] J. M. Chalker, C. S. C. Wood, B. G. Davis, *J. Am. Chem. Soc.* **2009**, *131*, 16346 – 16347.
[70] W. Ou, T. Uno, H.-P. Chiu, J. Grünewald, S. E. Cellitti, T. Crossgrove, X. Hao, Q. Fan, L. L. Quinn, P. Patterson, L. Okach, D. H. Jones, S. A. Lesley, A. Brock, B. H. Geierstanger, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10437 – 10442.
[71] Photoclick Reactions: W. Song, Y. Wang, J. Qu, Q. Lin, *J. Am. Chem. Soc.* **2008**, *130*, 9654 – 9655.
[72] A. A. Poloukhtine, N. E. Mbua, M. A. Wolfert, G.-J. Boons, V. V. Popik, *J. Am. Chem. Soc.* **2009**, *131*, 15769 – 15776.
[73] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86 – 89.
[74] A. Gautier, A. Juillerat, C. Heinis, I. R. Corrêa, M. Kindermann, F. Beaufils, K. Johnsson, *Chem. Biol.* **2008**, *15*, 128 – 136.
[75] D. Maurel, L. C. Agrar, C. Brock, M.-L. Rives, E. Bourrier, M. A. Ayoub, H. Bazin, N. Tinel, T. Durroux, L. Prezeau, E. Trinquet, J.-P. Pin, *Nat. Methods* **2008**, *5*, 561 – 567.
[76] G. V. Los et al., *ACS Chem. Biol.* **2008**, *3*, 373 – 382.
[77] R. F. Ohana, L. P. Encell, K. Zhao, D. Simpson, M. R. Slater, M. Urh, K. V. Wood, *Protein Expression Purif.* **2009**, *68*, 110 – 120.
[78] M. Urh, D. Hartzell, J. Mendez, D. H. Klaubert, K. Wood, *Methods Mol. Biol.* **2008**, *421*, 191 – 210.
[79] T. K. Neklesa, H. S. Tae, A. R. Schneekloth, M. J. Stulberg, T. W. Corson, T. B. Sundberg, K. Raina, S. A. Holley, C. M. Crews, *Nat. Chem. Biol.* **2011**, *7*, 538 – 543.
[80] M. Fernández-Suárez, H. Baruah, L. Martínez-Hernández, K. T. Xie, J. M. Baskin, C. R. Bertozzi, A. Y. Ting, *Nat. Biotechnol.* **2007**, *25*, 1483 – 1487.
[81] C. Uttamapinant, K. A. White, H. Baruah, S. Thompson, M. Fernández-Suárez, S. Puthenveetil, A. Y. Ting, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10914 – 10919.
[82] I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, *2*, 99 – 104.
[83] M. W. Popp, J. M. Antos, G. M. Grotenbreg, E. Spooner, H. L. Ploegh, *Nat. Chem. Biol.* **2007**, *3*, 707 – 708.
[84] T. Tanaka, T. Yamamoto, S. Tsukiji, T. Nagamune, *ChemBioChem* **2008**, *9*, 802 – 807.
[85] C. S. Lisenbee, S. K. Karnik, R. N. Trelease, *Traffic* **2003**, *4*, 491 – 501.
[86] B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269 – 272.
[87] S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis, R. Y. Tsien, *J. Am. Chem. Soc.* **2002**, *124*, 6063 – 6076.
[88] G. Gaietta, T. J. Deerinck, S. A. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien, M. H. Ellisman, *Science* **2002**, *296*, 503 – 507.
[89] O. Tour, R. M. Meijer, D. A. Zacharias, S. R. Adams, R. Y. Tsien, *Nat. Biotechnol.* **2003**, *21*, 1505 – 1508.
[90] O. Tour, S. R. Adams, R. A. Kerr, R. M. Meijer, T. J. Sejnowski, R. W. Tsien, R. Y. Tsien, *Nat. Chem. Biol.* **2007**, *3*, 423 – 431.
[91] B. Liu, C. T. Archer, L. Burdine, T. G. Gillette, T. Kodadek, *J. Am. Chem. Soc.* **2007**, *129*, 12348 – 12349.
[92] Z. Ignatova, B. Krishnan, J. P. Bombardier, A. M. C. Marcelino, J. Hong, L. M. Gierasch, *Pept. Sci.* **2007**, *2*, 157 – 163.
[93] K. W. Marek, G. W. Davis, *Neuron* **2002**, *36*, 805 – 813.
[94] E. Hochuli, H. Dobeli, A. Schacher, *J. Chromatography* **1987**, *411*, 177 – 184.
[95] E. G. Guignet, R. Hovius, H. Vogel, *Nat. Biotechnol.* **2004**, *22*, 440 – 444.
[96] S. Lata, M. Gavutis, R. Tampé, J. Piehler, *J. Am. Chem. Soc.* **2006**, *128*, 2365 – 2372.
[97] C. T. Hauser, R. Y. Tsien, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3693 – 3697.
[98] S. Fujishima, H. Nonaka, S. Uchinomiya, Y. A. Kawase, A. Ojida, I. Hamachi, *Chem. Commun.* **2012**, *48*, 594 – 596.
[99] S. Uchinomiya, H. Nonaka, S. Fujishima, S. Tsukiji, A. Ojida, I. Hamachi, *Chem. Commun.* **2009**, *39*, 5880 – 5882.
[100] T. L. Halo, J. Appelbaum, E. M. Hobert, D. M. Balkin, A. Schepartz, *J. Am. Chem. Soc.* **2009**, *131*, 438 – 439.
[101] B. R. Sculimbrene, B. Imperiali, *J. Am. Chem. Soc.* **2006**, *128*, 7346 – 7352.
[102] A. Ojida, K. Honda, D. Shinmi, S. Kiyonaka, Y. Mori, I. Hamachi, *J. Am. Chem. Soc.* **2006**, *128*, 10452 – 10459.
[103] H. Nonaka, S. Tsukiji, A. Ojida, I. Hamachi, *J. Am. Chem. Soc.* **2007**, *129*, 15777 – 15779.
[104] H. Nonaka, S. Fujishima, S. Uchinomiya, A. Ojida, I. Hamachi, *J. Am. Chem. Soc.* **2010**, *132*, 9301 – 9309.
[105] H. Kiefer, J. Lindstrom, E. S. Lennox, S. J. Singer, *Proc. Natl. Acad. Sci. USA* **1970**, *67*, 1688 – 1694.
[106] V. Chowdhry, F. H. Westheimer, *Annu. Rev. Biochem.* **1979**, *48*, 293 – 325.
[107] J. Das, *Chem. Rev.* **2011**, *111*, 4405 – 4417.
[108] M. J. Evans, B. F. Cravatt, *Chem. Rev.* **2006**, *106*, 3279 – 3301.
[109] A. M. Sadaghiani, S. H. L. Verhelst, M. Bogryo, *Curr. Opin. Chem. Biol.* **2007**, *11*, 20 – 28.
[110] G. Chen, A. Heim, D. Riether, D. Yee, Y. Milgrom, M. A. Gawinowicz, D. Sames, *J. Am. Chem. Soc.* **2003**, *125*, 8130 – 8133.
[111] C.-M. Kam, A. S. Abuelyaman, Z. Li, D. Hudig, J. C. Powers, *Bioconjugate Chem.* **1993**, *4*, 560 – 567.
[112] A. S. Abuelyaman, D. Hudig, S. L. Woodard, J. C. Powers, *Bioconjugate Chem.* **1994**, *5*, 400 – 405.
[113] Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694 – 14699.
[114] D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569 – 581.
[115] A. E. Speers, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 535 – 546.
[116] K. Tsuboi, D. A. Bachovchin, A. E. Speers, T. P. Spicer, V. Fernandez-Vega, P. Hodder, H. Rosen, B. F. Cravatt, *J. Am. Chem. Soc.* **2011**, *133*, 16605 – 16616.
[117] D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nat. Biotechnol.* **2003**, *21*, 687 – 691.
[118] N. Jessani, M. Humphrey, W. H. McDonald, S. Niessen, K. Masuda, B. Gangadharan, J. R. Yates III, B. M. Mueller, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13756 – 13761.
[119] G. Blum, S. R. Mullins, K. Keren, M. Fonovic, C. Jedeszko, M. J. Rice, B. F. Sloane, M. Bogyo, *Nat. Chem. Biol.* **2005**, *1*, 203 – 209.
[120] A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10000 – 10005.
[121] M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner, J. W. Kozarich, *Biochemistry* **2007**, *46*, 350 – 358.
[122] L.-C. Lo, T.-L. Pang, C.-H. Kuo, Y.-L. Chiang, H.-Y. Wang, J.-J. Lin, *J. Proteome Res.* **2002**, *1*, 35 – 40.
[123] C.-S. Tsai, Y.-K. Li, L.-C. Lo, *Org. Lett.* **2002**, *4*, 3607 – 3610.
[124] G. C. Adam, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2002**, *20*, 805 – 809.
[125] A. Adibekian, B. R. Martin, C. Wang, K.-L. Hsu, D. A. Bachovchin, S. Niessen, H. Hoover, B. F. Cravatt, *Nat. Chem. Biol.* **2011**, *7*, 469 – 478.

[126] E. W. S. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang, S. Q. Yao, *J. Am. Chem. Soc.* **2004**, *126*, 14435 – 14446.
[127] H. Shi, X. Cheng, S. K. Sze, S. Q. Yao, *Chem. Commun.* **2011**, *47*, 11306 – 11308.
[128] T. Hayashi, I. Hamachi, *Acc. Chem. Res.* **2012**, *45*, 1460 – 1469.
[129] I. Hamachi, T. Nagase, S. Shinkai, *J. Am. Chem. Soc.* **2000**, *122*, 12065 – 12066.
[130] E. Nakata, Y. Koshi, E. Koga, Y. Katayama, I. Hamachi, *J. Am. Chem. Soc.* **2005**, *127*, 13253 – 13261.
[131] Y. Takaoka, H. Tsutsumi, N. Kasagi, E. Nakata, I. Hamachi, *J. Am. Chem. Soc.* **2006**, *128*, 3273 – 3280.
[132] H. Wakabayashi, M. Miyagawa, Y. Koshi, Y. Takaoka, S. Tsukiji, I. Hamachi, *Chem. Asian J.* **2008**, *3*, 1134 – 1139.
[133] S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.* **2009**, *5*, 341 – 343.
[134] S. Tsukiji, H. Wang, M. Miyagawa, T. Tamura, Y. Takaoka, I. Hamachi, *J. Am. Chem. Soc.* **2009**, *131*, 9046 – 9054.
[135] Y. Takaoka, Y. Sun, S. Tsukiji, I. Hamachi, *Chem. Sci.* **2011**, *2*, 511 – 520.
[136] Y. Takaoka et al. unpublished results.
[137] T. Tamura, S. Tsukiji, I. Hamachi, *J. Am. Chem. Soc.* **2012**, *134*, 2216 – 2226.
[138] H. J. Cho, H. Y. Gee, K.-H. Baek, S.-K. Ko, J.-M. Park, H. Lee, N.-D. Kim, M. G. Lee, I. Shin, *J. Am. Chem. Soc.* **2011**, *133*, 20267 – 20276.
[139] M. Takahashi, A. Kawamura, N. Kato, T. Nishi, I. Hamachi, J. Ohkanda, *Angew. Chem.* **2012**, *124*, 524 – 527; *Angew. Chem. Int. Ed.* **2012**, *51*, 509 – 512.
[140] C. C. Hughes, Y.-L. Yang, W.-T. Liu, P. C. Dorrestein, J. L. Clair, W. Fenical, *J. Am. Chem. Soc.* **2009**, *131*, 12094 – 12096.
[141] S. Fujishima, R. Yasui, T. Miki, A. Ojida, I. Hamachi, *J. Am. Chem. Soc.* **2012**, *134*, 3961 – 3964.
[142] Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa, I. Hamachi, *J. Am. Chem. Soc.* **2008**, *130*, 245 – 251.
[143] Y. Sun, Y. Takaoka, S. Tsukiji, M. Narazaki, T. Matsuda, I. Hamachi, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4393 – 4396.
[144] H. Wang, Y. Koshi, D. Minato, H. Nonaka, S. Kiyonaka, Y. Mori, S. Tsukiji, I. Hamachi, *J. Am. Chem. Soc.* **2011**, *133*, 12220 – 12228.
[145] M. Kunishima, S. Nakanishi, J. Nishida, H. Tanaka, D. Morisaki, K. Hioki, H. Nomoto, *Chem. Commun.* **2009**, *37*, 5597 – 5599.
[146] Z. Chen, B. V. Popp, C. L. Bovet, Z. T. Ball, *ACS Chem. Biol.* **2011**, *6*, 920 – 925.
[147] Z. Chen, F. Vohidov, J. M. Coughlin, L. J. Stagg, S. T. Arold, J. E. Ladbury, Z. T. Ball, *J. Am. Chem. Soc.* **2012**, *134*, 10138 – 10145.
[148] S. W. Hell, *Science* **2007**, *316*, 1153 – 1158.
[149] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. L. Schwartz, H. F. Hess, *Science* **2006**, *313*, 1642 – 1645.
[150] V. N. Belov, C. A. Wurm, V. P. Boyarskiy, S. W. Hell, *Angew. Chem.* **2010**, *122*, 3598 – 3602; *Angew. Chem. Int. Ed.* **2010**, *49*, 3520 – 3523.
[151] K. Kolmakov, V. N. Belov, J. Bierwagen, C. Ringemann, V. Muller, C. Eggeling, S. W. Hell, *Chem. Eur. J.* **2010**, *16*, 158 – 166.