Tetrahedron Letters xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

Digest Paper

Recent advances in bioorthogonal reactions for site-specific protein labeling and engineering

Yukang Gong, Lifeng Pan *

State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Lingling Road 345, Shanghai 200032, People's Republic of China

ARTICLE INFO

Article history:
Received 15 November 2014
Revised 11 March 2015
Accepted 12 March 2015
Available online xxxx

Keywords:
Bioorthogonal organic chemistry
Ligand-directed organic chemistry
Protein labeling and engineering

ABSTRACT

In the past two decades, with the rapid development of chemical biology, tremendous small-molecule based toolkits were created by organic chemists, and were widely used to study and manipulate proteins in order to dissect their complicated biological functions. This review summarizes some recent progresses of bioorthogonal reactions for site-specific protein labeling and engineering, and highlights the powers of using these methods to study the biological functions of some proteins.

Contents

Introduction

As one of the most abundant biomolecules, proteins are involved in most of the biological processes and perform a wide array of important functions within living organisms. Therefore, the study and manipulation of protein functions are not only of significant importance to fundamental scientific research, but also critical to the development of biomedical and biotechnological applications. Traditional genetics, molecular biology, biochemistry, cell biology, and allied methods have provided various tools to investigate the functions of proteins, and have led to tremendous achievements including visualization of a protein using fluorescent protein fusions and silence of a protein expression using RNA interferences. However, not all the proteins and related biological processes are within the easy reach of those conventional approaches. Fortunately, recent rapid progress in the chemical biology field provides abundant new technologies for the study

* Corresponding author.
E-mail address: panlf@sioc.ac.cn (L. Pan).

http://dx.doi.org/10.1016/j.tetlet.2015.03.065
0040-4039/

Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065

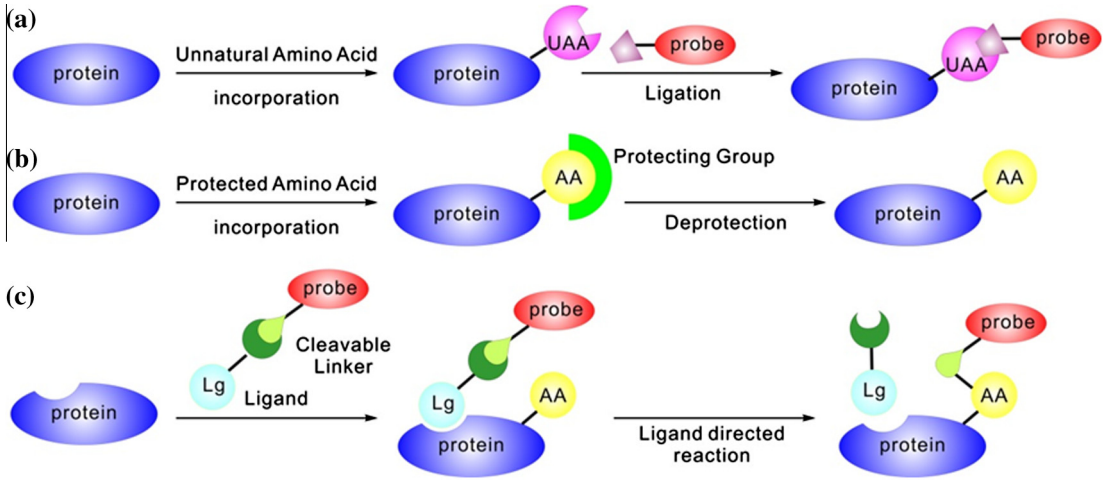


Figure 1. Three approaches to achieve the site-specific protein labeling and engineering. Schematic illustrations of site-specific incorporation of unnatural amino acids followed by bioorthogonal reactions including ligations (a) and deprotections (b). (c) Schematic illustration of the ligand-directed chemistry for site-specific protein labeling and engineering. UAA, unnatural amino acid; AA, natural amino acid; Lg, ligand.

of these challenging proteins and cellular processes. In particular, the modification of specific proteins with functional probes provides a powerful technique for the investigation of target proteins and their complex functions in detail. So far, there are a large variety of strategies developed by organic chemists to achieve site-specific labeling and engineering of target proteins with functional small molecules.[1] Due to space limitations, this digest only focuses on two strategies used for the site-specific protein labeling and engineering: (1) Bioorthogonal reactions with genetically encoded unnatural amino acids bearing functional groups that can be specifically ligated or deprotected (Fig. 1a and b); (2) Ligand-directed bioorthogonal reactions for site-specific modifications of target proteins (Fig. 1c).

The first strategy combines bioorthogonal reactions with genetically encoded unnatural amino acids bearing functional groups, such as aldehydes, ketones, azides, and alkenes, to facilitate the site-specific protein labeling and engineering (Fig. 1a and b).[2] Genetic code expansion and reprogramming enable the site-specific incorporations of diverse designed unnatural amino acids into proteins.[3] By evolving orthogonal ribosomes, developing mutually orthogonal synthetase/tRNA pairs and manipulating genomes, the efficiency of unnatural amino acids incorporations and the numbers of unnatural amino acids that can be site-specifically encoded are constantly increasing. Notablely, the development and application of the pyrrolysyl-transfer RNA (tRNA) synthetase/tRNA pair for unnatural amino acids incorporation have moved genetic code expansion from bacteria to eukaryotic cells and multicellular organisms.[4]

Although genetic code expansion is a powerful tool, it also has some limitations, for example, genetic modification and subsequent overexpression of proteins may perturb the physiological condition of cells. Thus, another strategy named ligand-directed chemistry for site-specific protein labeling that modifies selective endogenous proteins under their physiological conditions, was developed (Fig. 1c).[5] In this approach, a synthetic molecule containing three functional groups including a target protein binding ligand, a reactive linker and a functional probe is constructed, and firstly its ligand part specifically binds to its target protein, then driving a bioorthogonal reaction between the reactive linker group with an amino acid located at the vicinity of the ligand-binding site of the target protein facilitated by the proximity effect, finally end with the labeling of the target protein with the functional probe. Therefore, this method can satisfy the requirements of target selectivity and site specificity.

Since these two methods revolutionized our abilities to site-specifically label and manipulate intact proteins, these two areas are rapidly growing and many elegant applications have been recently reported. In this review, we summarize some recent developments in these two fields using bioorthogonal reactions for site-specific protein labeling and engineering.

Bioorthogonal reactions with genetically encoded unnatural amino acids for site-specific protein labeling and engineering

In general, the bioorthogonal reactions used for site-specific protein labeling and engineering based on genetically encoded unnatural amino acids can be mainly classified into two categories: ligation reactions (Fig. 1a) and deprotection reactions (Fig. 1b).

Bioorthogonal ligation reactions

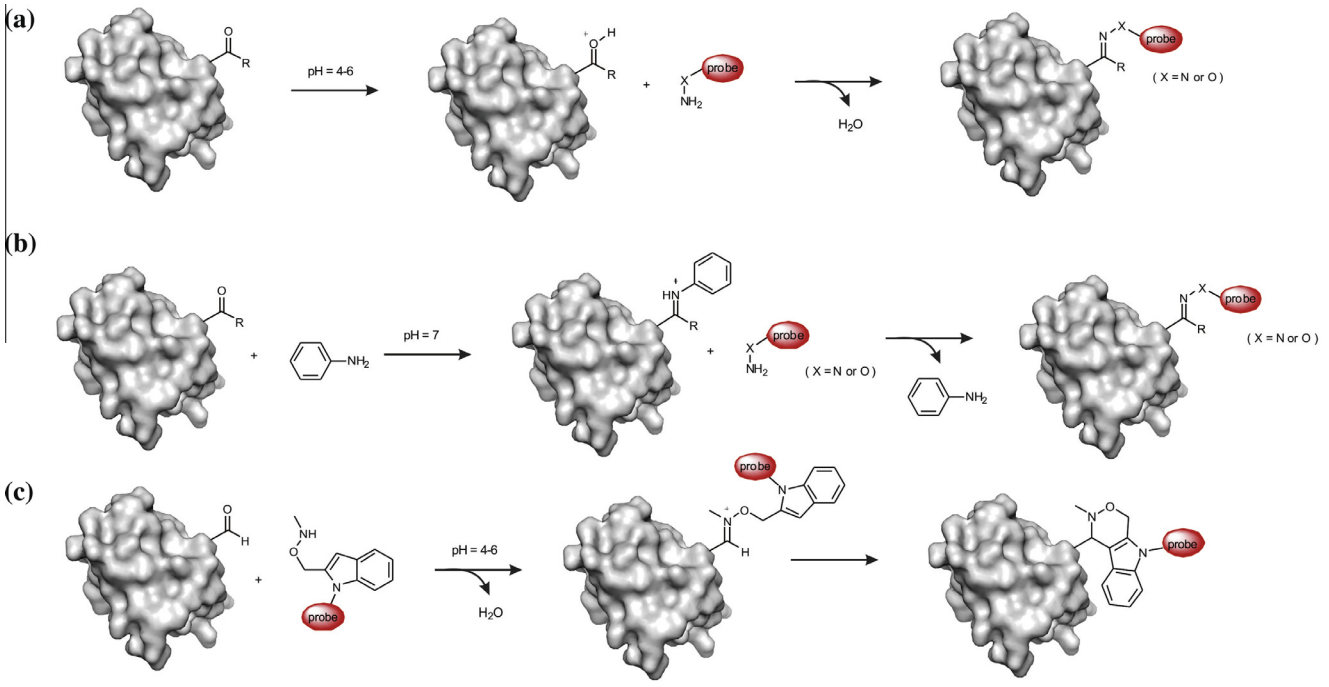
Ligation reactions through aldehydes and ketones

Genetically encoded aldehydes and ketones can specifically react with hydrazides and alkoxyamines to produce stable hydrazone and oxime, respectively (Scheme 1a),[6] and were successfully applied for the site-specific in vitro or cell surface protein labeling.[7] However the unfavorable acidic conditions and slow kinetics of these reactions prevent their applications in most intracellular settings.[7a,8] To overcome these shortcomings, aniline was identified and used as a nucleophilic catalyst for both specific cell surface and intracellular protein labeling (Scheme 1b).[9] Based on the classic Pictet–Spengler reaction between aldehydes and tryptamine nucleophiles, recently Bertozzi and co-workers reported the Pictet–Spengler ligation reaction (Scheme 1c).[10] In this reaction, aldehydes react with alkoxyamines to form intermediate oxyiminium ions, which then undergo intramolecular C–C bond formations with indole nucleophiles to form hydrolytically stable oxacarboline products. In conjunction with techniques for genetic incorporations of unnatural amino acids bearing aldehydes, the Pictet–Spengler ligation provides a unique tool to generate stable bioconjugates for biomedical applications.

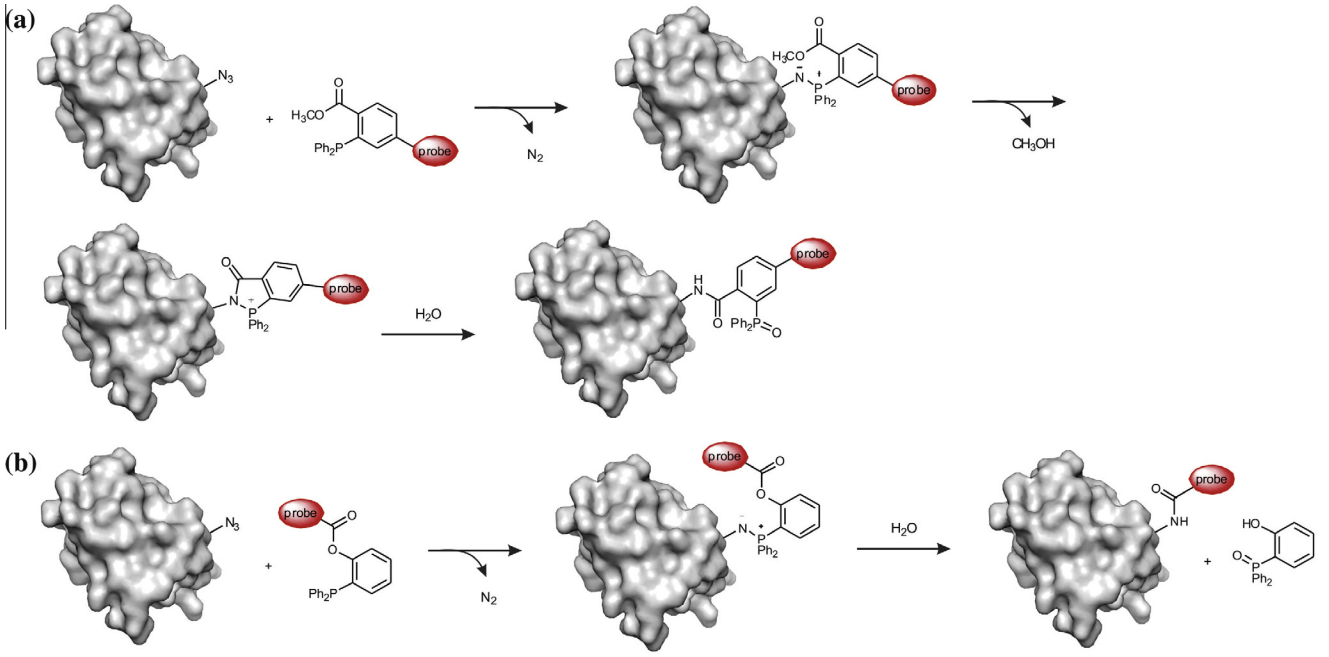
Ligation reactions through azides

Genetically encoded azide groups were firstly developed for the site-specific biomolecule labeling by the Bertozzi group through a process known as Staudinger ligation, a modification of the classic Staudinger reduction of azides with triphenylphosphine.[11] In this process, the proteins bearing alkyl azides undergo ligation reactions to form stable amide bonds with triarylphosphine derivatives

Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065



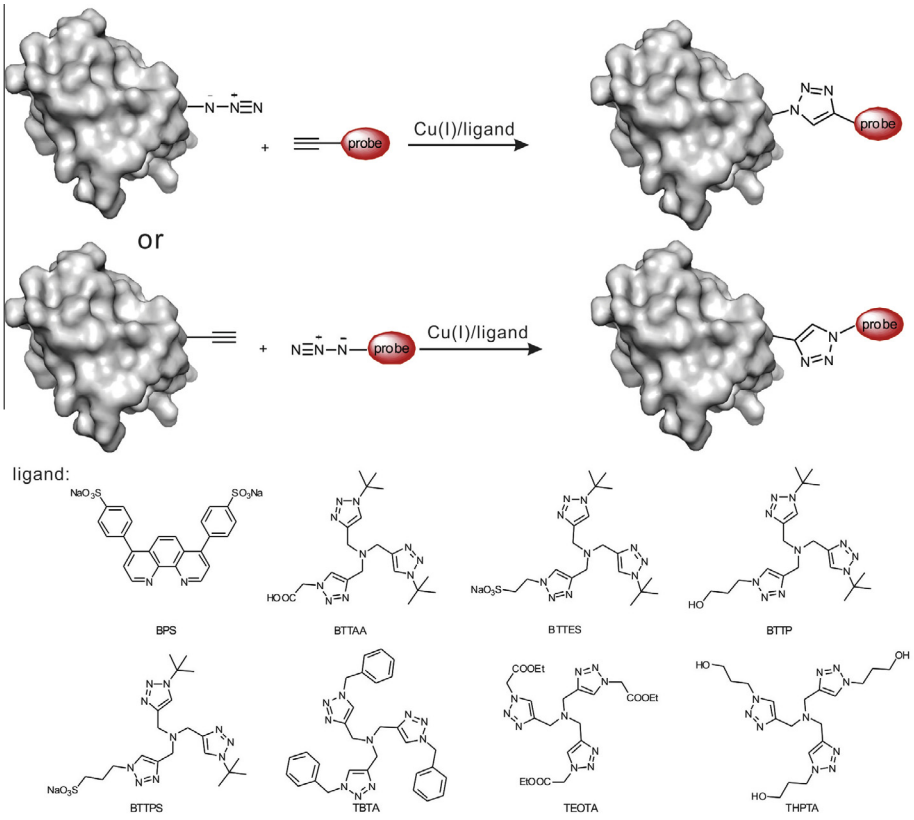
Scheme 1. Bioorthogonal reactions of genetically encoded ketones/aldehydes with hydrazines or alkoxyamines for site-specific protein labeling and engineering. (a) Acid-catalyzed reaction of genetically encoded aldehydes/ketones with amino nucleophiles. (b) Aniline-catalyzed reaction of genetically encoded aldehydes/ketones with hydrazines or alkoxyamines. (c) Pictet–Spengler ligation of genetically encoded aldehydes with tryptamine nucleophiles.



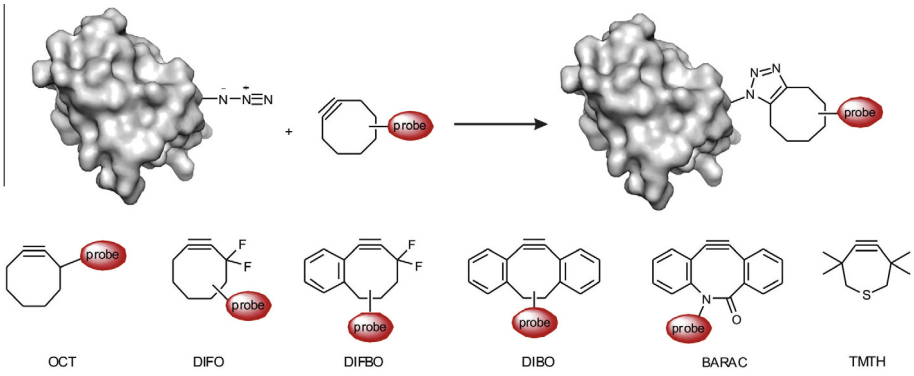
Scheme 2. Staudinger ligations of azides and triarylphosphines for site-specific protein labeling and engineering. (a) The Staudinger ligation between genetically encoded azides and triarylphosphines. (b) The traceless Staudinger ligation.

that have ester groups on their aromatic rings (Scheme 2a). Later, some phosphine reagents, in which the acyl group is attached via a cleavable linker to the phosphine group, were developed, and a variant of this reaction was reported, named as "traceless Staudinger ligation", where the phosphine oxide moiety is absent in the final bioconjugate (Scheme 2b).[12] These Staudinger ligations have been successfully used to site-specifically label protein in vitro and in many different cellular conditions.[13] Limitations of the Staudinger ligations are their slow kinetics and the oxidation sensitivities of phosphines, therefore the phosphine reagents have to be used at relatively high concentrations.[11b]

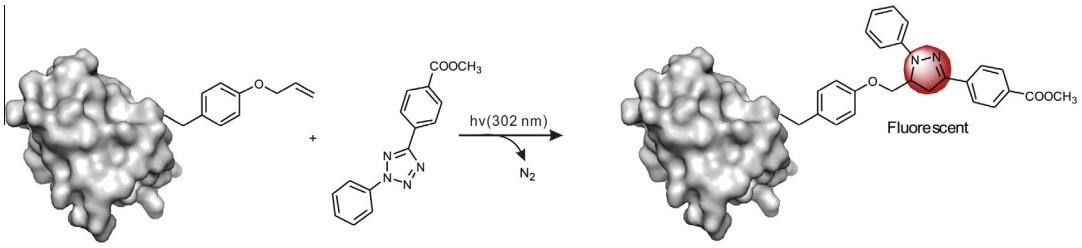
Besides Staudinger ligation, azides can also take part in [3+2] cycloadditions with alkynes to yield stable triazoles, but high temperature and pressure are normally required to form the triazole products in a reasonable yield.[14] There are generally two strategies to improve the reaction: catalyzing the reaction with



Scheme 3. Copper-catalyzed azide–alkyne cycloadditions (CuACC) for site-specific protein labeling and engineering together with ligands including BPS, TBTA, TEOTA, THPTA, BTTAA, BTTES, BTTP, and BTTPS for the coordination of Cu(I).



Scheme 4. Strain-promoted azide–alkyne 1,3-dipolar cycloadditions (SPAAC) for site-specific protein labeling and engineering together with strained cycloalkyne derivatives OCT, DIFO, DIFBO, DIBO, BARAC, and TMTH used in SPAAC.



Scheme 5. Photoclick chemistry of genetically encoded *O*-allyltyrosines with tetrazoles for site-specific protein labeling and engineering.

copper (I)[15] or introducing ring strain into the alkyne.[16] The former is termed as the copper catalyzed azide–alkyne cycloaddition (CuAAC), and the latter is named as the strain-promoted alkyne–azide cycloaddition (SPAAC).

CuAAC, taking advantage of the formation of a dinuclear copper intermediate to activate both terminal alkynes and azides,[17] is a representative of click reaction (Scheme 3).[18] Moreover, the Cu(I)-catalyzed cycloaddition can be further accelerated by the

Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065

use of specific ligands for Cu(I).[19] These ligands, such as BPS (Scheme 3) and natural amino acid histidine, coordinate Cu(I) to form activated copper catalysts that can promote the azide–alkyne cycloaddition at low concentrations of Cu(I). However, the main limitation of CuAAC is the toxicity of Cu(I),[20] which is caused by Cu(I)-mediated generation of reactive oxygen species (ROS) from O_2.[21] Recently, the use of water-soluble ligands including TBTA, TEOTA, THPTA, BTTAA, BTTES, BTTP, and BTTPS can reduce the apparent copper-associated cytotoxicity by serving as reductants to protect cells from ROS (Scheme 3).[22] By using these rate-accelerating ligands, the CuAAC have been used to site-specifically label proteins in living *Escherichia. coli* and mammalian cells.[23]

SPAAC is another approach to accelerate the azide–alkyne cycloaddition (Scheme 4). Cyclooctyne derivatives have been used to label proteins containing genetically encoded azides in living systems.[24] To improve the reaction, more reactive cyclooctyne compounds including difluorocyclooctyne (DIFO) derivatives, difluorobenzocyclooctyne (DIFBO), dibenzocyclooctynes (DIBO), and biarylazacyclooctynone compounds (BARAC) (Scheme 4) were developed and have been used to probe azide-containing proteins within complex biological systems including mammalian cells, *Caenorhabditis elegans* and zebrafish embryos.[25] Recently, thiacycloheptynes, such as TMTH (Scheme 4), have been developed as a new type of reagents for SPAAC.[26]

Ligation reactions through terminal alkenes and tetrazoles

Photoinduced organic reactions are also explored to enable the site-specific protein labeling processes in biological settings. Lin and co-workers developed the UV-light induced 1,3-dipolar cycloaddition reactions between tetrazole derivativess and terminal alkenes (Scheme 5), also named as photoclick chemistry.[27] By incorporations of tetrazole and alkene groups into proteins in the forms of unnatural amino acids, the photoclick chemistry has been used to site-specifically label and engineer proteins in vitro, and also to visualize proteins in living bacteria and mammalian cells.[27,28] There are several advantages of the photoclick chemistry: first and foremost, its inducibility by light makes it a powerful tool for spatiotemporal initiation of labeling reactions in living systems; second, the reaction is fluorogenic and only the resulting pyrazoline product is fluorescent, which is helpful for living cell imaging studies;[27a] third, the fast reaction kinetics, presenting a significantly faster bioorthogonal reaction than the Staudinger ligation and the SPAAC.[28a,29] Development of new tetrazole reagents, which are highly reactive and can be light-activated at wavelengths that are harmless to living cells, will make this reaction more attractive.

Ligation reactions through strained alkenes/alkynes and tetrazines

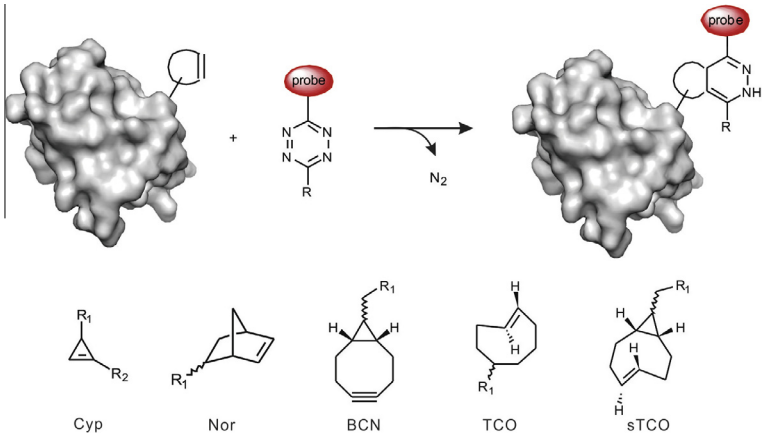
In 2008, Fox and co-worker firstly reported a protein bioconjugation strategy based on the inverse-electron-demand Diels–Alder reaction between a tetrazine compound and a strained *trans*-cyclooctene.[30] Later, various unnatural amino acids bearing strained alkene or alkyne groups, such as norbornenes (Nor), cyclopropenes (Cyp), bicyclononynes (BCN), and *trans*-cyclooctenes (TCO or sTCO), were synthesized and genetically incorporated into proteins for site-specific protein labeling and engineering with tetrazines in living bacteria and mammalian cells (Scheme 6).[31] These reactions are very specific, extraordinarily fast, and many of the red tetrazine fluorophores, whose fluorescences are initially quenched by the tetrazine, become strongly fluorescent upon cycloadditions with strained alkenes or alkynes, making them particularly useful for the labeling of proteins with low cellular abundances and for tracking fast protein-involved biological processes.

Ligation reactions through cross-metathesis

Another bioorthogonal reaction involving alkenes is olefin metathesis, one of the most powerful organic reactions for the construction of new carbon–carbon bonds. In 2008, Davis and co-workers reported the first utility of cross-metathesis in site-specifically labeling proteins containing an allyl sulfide group in vitro (Scheme 7).[32] They screened a small panel of alkenes and identified *S*-allylcysteine as the most efficient substrate for the cross-metathesis reaction with allyl alcohol using the Hoveyda–Grubbs second-generation catalyst.[33] Recently, they further developed a rapid and efficient cross-metathesis reaction using Se-allyl-selenocysteine for in vitro site-specific labeling and engineering of proteins.[34]

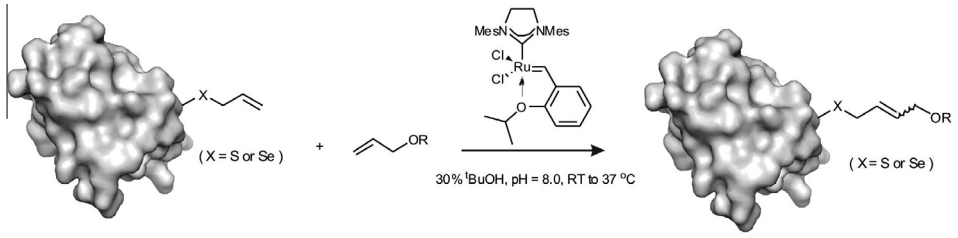
Ligation reactions through palladium-catalyzed cross-coupling

Palladium-catalyzed cross-coupling reactions have also been exploited to site-specifically label and engineer proteins. Initial reports using palladium-mediated cross-coupling reactions on proteins containing genetically encoded *p*-iodophenylalanine or *p*-boronophenylalanine suffered from very low reaction conversions or harsh reaction conditions.[35] Later, Davis and co-workers developed a water-soluble palladium catalyst, a sodium salt of 2-amino-4,6-dihydroxypyrimidine in complex with palladium [$Pd(OAc)_2(ADHP)_2$], for the Suzuki–Miyaura cross-coupling reaction between a genetically encoded *p*-iodobenzyl group and various aryl and alkenyl boronic acids (Scheme 8a).[36] The improved palladium-mediated Suzuki–Miyaura cross-coupling reactions typically reached completion within 1 h at 37 °C, and

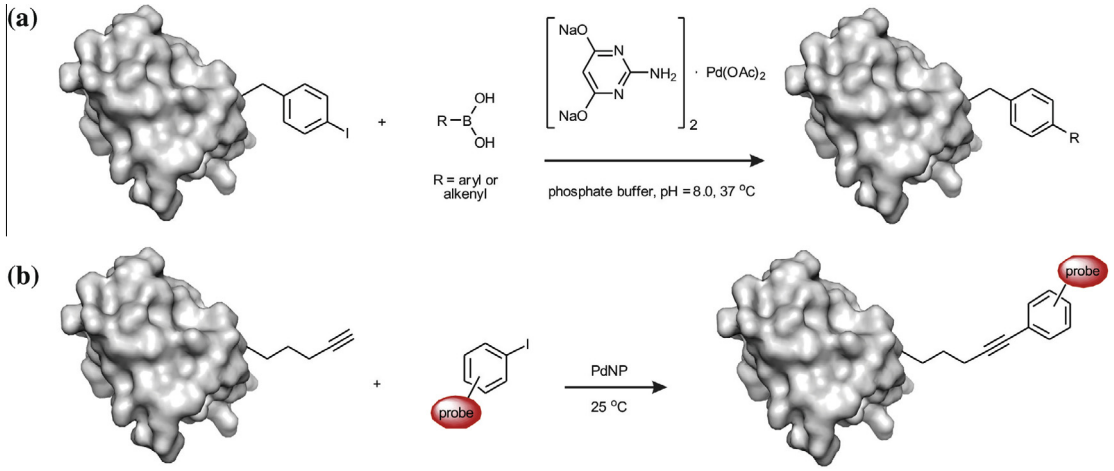


Scheme 6. Inverse-electron-demand Diels–Alder reactions between tetrazines and genetically encoded strained alkenes or alkynes including cyclopropenes (Cyp), alkynes norbornenes (Nor), bicyclononynes (BCN), and *trans*-cyclooctenes (TCO or sTCO) for site-specific protein labeling and engineering.

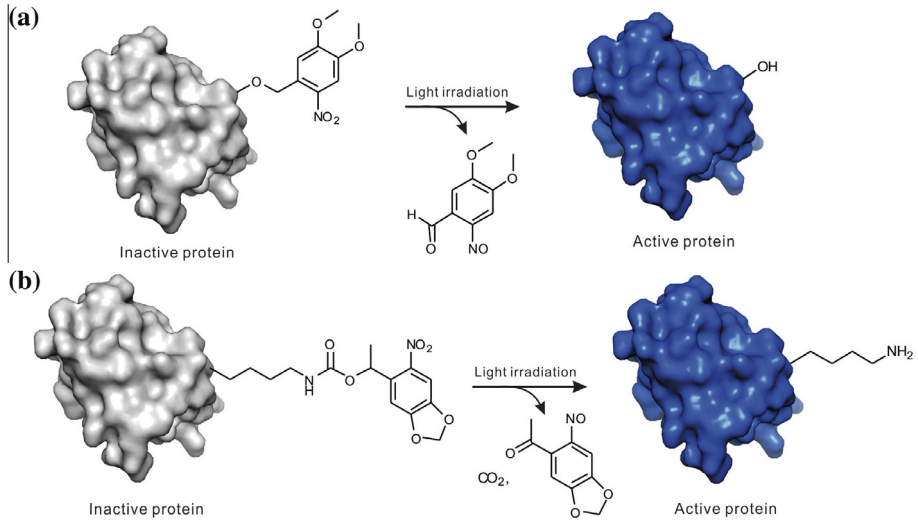
Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065



Scheme 7. Cross-metathesis of genetically encoded allyl sulfides or allyl selenides with allyl ethers using a Hoveyda–Grubbs second-generation catalyst for site-specific protein labeling and engineering.



Scheme 8. (a) Suzuki–Miyaura cross-coupling reactions between genetically encoded *p*-iodophenylalanines and aryl- or alkenyl-boronic acids mediated by the water-soluble palladium catalyst, $Pd(OAc)_2(ADHP)_2$. (b) Ligand free Sonogashira cross-coupling reactions between genetically encoded alkyne groups and iodophenyls mediated by the PdNP generated from $Pd(NO_3)_2$.



Scheme 9. Deprotection reactions of a genetically encoded photocaged serine (a) or a photocaged lysine (b) using the photo-uncaging method.

were used to site-specifically label the membrane protein OmpC on *E. coli* surface. Another famous cross-coupling reaction, Sonogashira reaction, has also been developed as a bioorthogonal reaction to label alkyne-encoded proteins. Qing Li and co-workers firstly reported a bioorthogonal Pd-mediated Sonogashira cross-coupling reaction using a robust aminopyrimidine–palladium(II) complex, which selectively modify a homopropargylglycine (HPG)-encoded ubiquitin in aqueous medium as well as in *E. coli*.[37] Recently, Chen and co-workers developed a ligand-free Pd-mediated Sonogashira cross-coupling reactions system (Scheme 8b), and proved that PdNPs generated from water soluble $Pd(NO_3)_2$ can be an efficient and biocompatible catalyst for site-specific labeling of alkyne-modified proteins inside living *E. coli* and other Gram-negative bacterial pathogens such as *Shigella*.[38]

Bioorthogonal deprotection reactions

Deprotection reactions through photo-uncaging

The activities of some proteins can be controlled through site-specific installations of caging groups on side chains of key

residues that are essential for protein function. Since light irradiation is a relatively noninvasive method, direct caging of proteins via a genetically encoded unnatural amino acid bearing a photocleavable group allows photochemically control the functionality or localization of the target proteins in complex cellular conditions.[39] In conventional photocaging strategy, a photocaged serine or lysine residue bearing a photocleavable group was designed and genetically encoded into target proteins (Scheme 9). This photocaged serine or lysine amino acid was converted to wild type amino acid by irradiation with relatively low-energy light, and was used to photochemically control the function of the target protein. In 2011, Chin and co-workers used this photocaging strategy to control a photocaged MEK1 kinase, and demonstrated a receptor independent activation of an artificial subnetwork within the Raf/MEK/ERK pathway, which provided new insight into adaptive feedback and the kinetics of single steps in the MAP kinase signaling cascades.[40] More recently, Deiters and co-workers used this strategy to control a photocaged T7 RNA polymerase, and demonstrated the photocaging of a synthetic gene network in mammalian cells.[41] The main advantages of this photocaging strategy are that the unnatural amino acids are site-specifically incorporated and the modified proteins are directly generated inside the cell, eliminating the requirement for additional transfection or injection.

Deprotection reactions through palladium-mediated uncaging

Even using non-phototoxic light, the poor penetration ability of light hinders the further utilizations of these photo-uncaging methods in deep tissues or intact animals. To avoid these limitations, Chen and co-workers recently reported the development of a palladium-mediated chemical uncaging method to control lysine-dependent activation of intracellular proteins (Scheme 10).[42] In this method, they firstly used a genetically encoded lysine analogue bearing a propargyl carbamate group to protect a key lysine residue of the target protein, then used biocompatible and efficient palladium catalysts, such as allyl$_2$Pd$_2$Cl$_2$, to cleave the propargyl carbamate group of the protected lysine analogue to generate a free lysine residue. This palladium-mediated deprotection strategy was further proved to work with a range of different proteins and cell lines, and was successfully used to reveal the detail virulence mechanism of a bacterial Type III effector protein.[42]

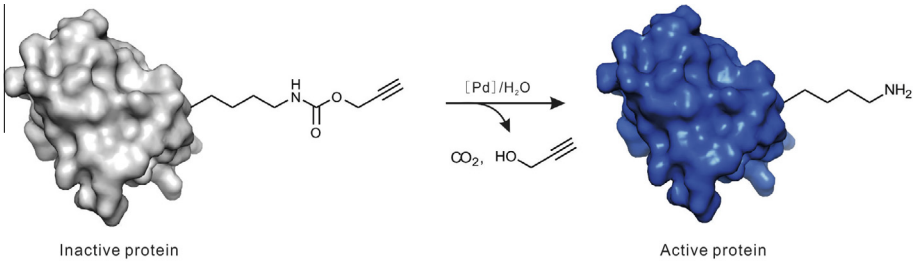
Ligand-directed reactions for site-specific protein labeling and engineering

Unlike the incorporations of unnatural amino acids by gene manipulation, which made the target proteins no longer endogenous, the ligand-directed protein labeling method can achieve the site-specific labeling of an endogenous target protein.

Recently, a ligand-directed tosyl (LDT) chemistry employing a unique LDT reagent, which consists of a target protein binding ligand, a reactive tosyl linker and a functional probe, has been developed to label natural proteins (Scheme 11a). The tosyl linker of the LDT reagent behaves not only as a linker between the ligand and the probe, but also as a reactive group.[5,43] When the LDT reagent binds to the target protein, a S_N2-type reaction is triggered due to the proximity effect between the tosyl moiety and a nucleophilic amino acid located at the vicinity of the ligand-binding pocket of target proteins, resulting in the site-specific labeling of the target protein and release of the ligand moiety. This labeling method can be applied for the site-specific modification of a variety of endogenous proteins under intracellular environments.[5,44] The power of LDT chemistry for characterizations of both the target protein and the ligand-binding sites has been well demonstrated.[45] Moreover, combining with fluorescent protein tag technology and FRET imaging, the LDT chemistry has been demonstrated to construct a natural protein as a fluorescent reporter, which can be used to detect its molecular interactions in vitro and in living cells.[46]

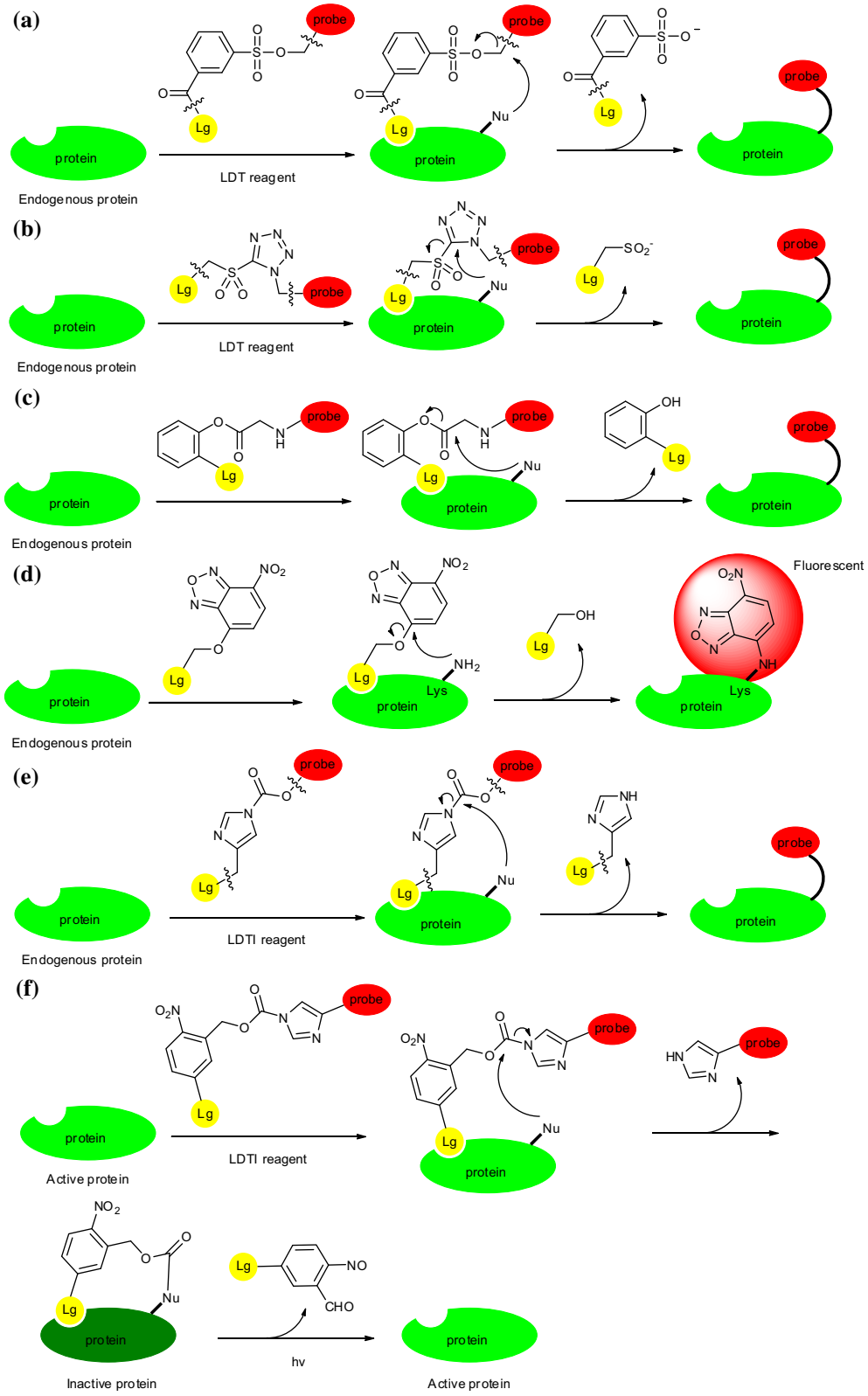
Meanwhile, additional ligand-directed chemoselective reactions based on different reactive linkers have been developed. Hideaki Kakeya and co-workers reported the use of 5-sulfonyl tetrazole as the reactive linker to develop a ligand-directed reaction for site-specific protein modifications (Scheme 11b). They employed this method to achieve the chemical labeling of the cellular receptor of the natural product, cyclosporine A.[47] Using an acyl phenol moiety as the reactive linker, Fenical group demonstrated another ligand-directed chemoselective reaction for site-specific protein modifications (Scheme 11c), which was successfully applied to identify the target protein of the anticancer natural product, marrinopyrrole A.[48] Based on a bifunctional O-NBD unit (NBD: nitrobenzoxadiazole), Sodeoka group reported a simple ligand-directed chemoselective probe for site-specific protein labeling (Scheme 11d). The O-NBD unit is non-fluorescent, but can be converted into a fluorogenic amino NBD group when reacted with a Lys residue. After using the O-NBD unit attached *N*,*N*-dialkyl-2-phenylindol-3-ylglyoxylamides (PIGAs), which are ligands of translocator protein (TSPO), they were able to visualize the mitochondria expressing TSPO in living cells.[49]

More recently, a new type of ligand-directed chemistry, known as ligand-directed acyl imidazole (LDAI) chemistry, has been developed (Scheme 11e).[50] In this LDAI labeling method, a triple functional LDAI reagent is used, which contains a moderately reactive alkyloxyacyl imidazole linker in addition to a target protein binding group and a functional probe. Similar to the tosyl linker in LDT chemistry, the alkyloxyacyl imidazole linker can react with an accessible nucleophilic residue of target protein assisted by ligand binding. The power of LDAI chemistry has been proved by its broad applications to selective chemical labeling of various types of membrane proteins under living cell conditions for functional studies.[51] In addition to introducing the functional probe, the LDAI chemistry can also be applied to construct caged proteins in a rational one-step manner, which was demonstrated by Itaru Hamachi and his co-workers (Scheme 11f). They showed that the activity of the caged carbonic anhydrase I was almost fully suppressed and absolutely recovered by light irradiation under in vitro conditions.[52] Since LDAI reagents can react with Ser, Tyr,



Scheme 10. Deprotection reactions of a genetically encoded caged lysine with the propargyl group using a palladium-mediated chemical uncaging method.

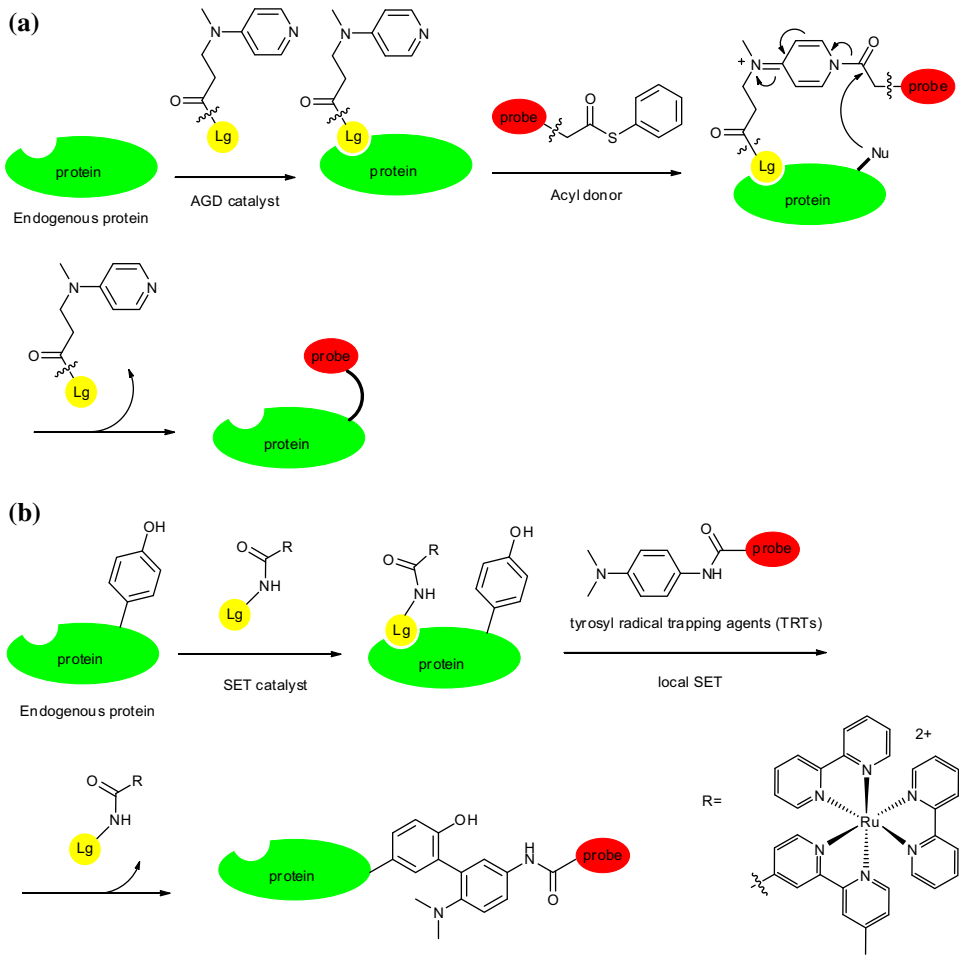
Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065



Scheme 11. Ligand-directed chemoselective reactions for site-specific protein labeling and engineering. Schematic illustrations of the ligand-directed tosyl (LDT) chemistry (a), the ligand-directed 5-sulfonyl tetrazole chemistry (b), the ligand-directed acyl phenol chemistry (c), the ligand-directed O-NBD chemistry (d), the ligand-directed acyl imidazole (LDAI) chemistry (e), and one-step construction of a caged protein by LDAI chemistry as well as the following photo-uncaging (f).

and Lys residues, while LDT reagents are reactive with His, Tyr and Glu residues, therefore these two methods are mutually complementary.[52]

The application of catalysts has also been introduced in ligand-directed chemistry. Hamachi group pioneered the development of catalyst-mediated ligand-directed chemoselective reactions for



Scheme 12. Catalytic ligand-directed reactions for site-specific protein labeling and engineering. Schematic illustrations of the AGD catalyst-mediated (a) and the SET catalyst-mediated (b) ligand-directed site-specific protein modification methods.

site-specific protein modification using the affinity-guided *N*,*N*-dimethylaminopyridine (AGD) catalysts, which transfer acyl donor probes to nucleophilic residues near the ligand-binding site of the target protein through an S_N2-type reaction (Scheme 12a).[53] Recently, Hiroyuki Nakamura and co-workers developed a ligand-directed site-specific protein modification method based on local single-electron transfer (SET) catalysis (Scheme 12b). In this catalyst-mediated method, a single-electron transfer between the ruthenium tris(2,2′-bipyridyl) complex ($[Ru(bpy)_3]^{2+}$) and a nearby Tyr residue of the targeted protein generates the tyrosyl radical that can react with tyrosyl radical trapping agents containing an *N*′-acyl-*N*,*N*-dimethyl-1,4-phenylenediamine unit through a catalytic oxidative radical addition reaction.[54]

Conclusions and perspectives

Recent progress of chemical biology has provided significant advances in the field of selective protein modifications with functional small molecules. Two important strategies based on the bioorthogonal reactions for site-specific protein labeling and engineering are summarized in this review. The development of the bioorthogonal reactions coupled with the genetic-code expansion method not only revolutionized our abilities to site-specifically label and manipulate intact proteins, but also will have important implications for future practical applications. Meanwhile, with the development of new reagents, the ligand-directed reactions will become increasingly powerful for site-specific labeling and engineering of endogenous proteins. However, it is worthwhile to point out that the use of organic chemistry for labeling and manipulating endogenous proteins in their native conditions is currently still in its infancy and none of these bioorthogonal reactions mentioned in this review are fully perfect inside living cells. Thus, future efforts are expected to improve the reaction biocompatibility and specificity, and to enhance the reaction efficiency as well as to invent new organic chemistry based tools that can efficiently and site-specifically label and engineer natural proteins in living systems.

Acknowledgements

Financial support from the National Natural Science Foundation of China (No. 31470749), a ‘Thousand Talents Program’ young investigator award of China, a Shanghai Rising Star Scholar award (No. 13QA1404300), the start-up fund from State Key Laboratory of Bioorganic and Natural Products Chemistry and Chinese Academy of Sciences are gratefully appreciated.

References and notes

1. (a) Boyce, M.; Bertozzi, C. R. *Nat. Methods* **2011**, *8*, 638; (b) Kim, C. H.; Axup, J. Y.; Schultz, P. G. *Curr. Opin. Chem. Biol.* **2013**, *17*, 412; (c) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13.
2. (a) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem.* **2009**, *48*, 6974; (b) Lim, R. K.; Lin, Q. *Chem. Commun. (Camb.)* **2010**, 1589; (c) Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16; (d) Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764.
3. Chin, J. W. *Annu. Rev. Biochem.* **2014**, *83*, 379.

4. Greiss, S.; Chin, J. W. *J. Am. Chem. Soc.* **2011**, *133*, 14196.
5. Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. *Nat. Chem. Biol.* **2009**, *5*, 341.
6. (a) Gaertner, F. H.; Rose, K.; Cotton, R.; Timms, D.; Camble, R.; Offord, E. R. *Bioconjugate Chem.* **1992**, *3*, 262; (b) Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30; (c) Canne, L. E.; Ferre-D'Amare, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998.
7. (a) Mahal, L. K. *Science* **1997**, *276*, 1125; (b) Zhang, Z.; Smith, B. A.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G. *Biochemistry* **2003**, *42*, 6735; (c) Sadamoto, R.; Niikura, K.; Ueda, T.; Monde, K.; Fukuhara, N.; Nishimura, S. *J. Am. Chem. Soc.* **2004**, *126*, 3755; (d) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99.
8. (a) Rideout, D. *Science* **1986**, *233*, 561; (b) Nauman, D. A.; Bertozzi, C. R. *Biochim. Biophys. Acta Bioenerg.* **2001**, *1568*, 147.
9. (a) Dirksen, A.; Dawson, P. E. *Bioconjugate Chem.* **2008**, *19*, 2543; (b) Zeng, Y.; Ramya, T. N.; Dirksen, A.; Dawson, P. E.; Paulson, J. C. *Nat. Methods* **2009**, *6*, 207; (c) Rayo, J.; Amara, N.; Krief, P.; Meijler, M. M. *J. Am. Chem. Soc.* **2011**, *133*, 7469.
10. (a) Stockigt, J.; Antonchick, A. P.; Wu, F.; Waldmann, H. *Angew. Chem.* **2011**, *50*, 8538; (b) Agarwal, P.; van der Weijden, J.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 46.
11. (a) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939; (b) Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 2686.
12. Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. *Org. Lett.* **2000**, *2*, 2141.
13. (a) Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14846; (b) Ovaa, H.; van Swieten, P. F.; Kessler, B. M.; Leeuwenburgh, M. A.; Fiebiger, E.; van den Nieuwendijk, A. M.; Galardy, P. J.; van der Marel, G. A.; Ploegh, H. L.; Overkleeft, H. S. *Angew. Chem.* **2003**, *42*, 3626; (c) Vocadlo, D. J.; Hang, H. C.; Kim, E. J.; Hanover, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9116; (d) Luchansky, S. J.; Argade, S.; Hayes, B. K.; Bertozzi, C. R. *Biochemistry* **2004**, *43*, 12358; (e) Watzke, A.; Kohn, M.; Gutierrez-Rodriguez, M.; Wacker, R.; Schroder, H.; Breinbauer, R.; Kuhlmann, J.; Alexandrov, K.; Niemeyer, C. M.; Goody, R. S.; Waldmann, H. *Angew. Chem.* **2006**, *45*, 1408; (f) Hangauer, M. J.; Bertozzi, C. R. *Angew. Chem.* **2008**, *47*, 2394; (g) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. *J. Am. Chem. Soc.* **2009**, *131*, 4967.
14. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. *J. Am. Chem. Soc.* **2005**, *127*, 210.
15. (a) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.* **2002**, *41*, 2596; (b) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
16. (a) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046; (b) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644.
17. Worrell, B. T.; Malik, J. A.; Fokin, V. V. *Science* **2013**, *340*, 457.
18. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem.* **2001**, *40*, 2004.
19. (a) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.* **2004**, *6*, 2853; (b) Rodionov, V. O.; Presolski, S. I.; Diaz, D. D.; Fokin, V. V.; Finn, M. G. *J. Am. Chem. Soc.* **2007**, *129*, 12705.
20. Wolbers, F.; ter Braak, P.; Le Gac, S.; Luttge, R.; Andersson, H.; Vermes, I.; van den Berg, A. *Electrophoresis* **2006**, *27*, 5073.
21. Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2010**, *21*, 1912.
22. Yang, M.; Li, J.; Chen, P. R. *Chem. Soc. Rev.* **2014**, *43*, 6511.
23. (a) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem.* **2011**, *50*, 8051; (b) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. *Angew. Chem.* **2009**, *48*, 9879; (c) Kennedy, D. C.; McKay, C. S.; Legault, M. C.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. *J. Am. Chem. Soc.* **2011**, *133*, 17993; (d) Soriano Del Amo, D.; Wang, W.; Jiang, H.; Besanceney, C.; Yan, A. C.; Levy, M.; Liu, Y.; Marlow, F. L.; Wu, P. *J. Am. Chem. Soc.* **2010**, *132*, 16893.
24. (a) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793; (b) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664.
25. (a) Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 11486; (b) Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G. J. *Angew. Chem.* **2008**, *47*, 2253; (c) Sletten, E. M.; Nakamura, H.; Jewett, J. C.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 11799; (d) Gordon, C. G.; Mackey, J. L.; Jewett, J. C.; Sletten, E. M.; Houk, K. N.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9199.
26. de Almeida, G.; Sletten, E. M.; Nakamura, H.; Palaniappan, K. K.; Bertozzi, C. R. *Angew. Chem.* **2012**, *51*, 2443.
27. (a) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. *Angew. Chem.* **2008**, *47*, 2832; (b) Song, W.; Wang, Y.; Yu, Z.; Vera, C. I.; Qu, J.; Lin, Q. *ACS Chem. Biol.* **2010**, *5*, 875; (c) Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q. *Angew. Chem.* **2012**, *51*, 10600.
28. (a) Song, W.; Wang, Y.; Qu, J.; Lin, Q. *J. Am. Chem. Soc.* **2008**, *130*, 9654; (b) Yu, Z.; Ohulchanskyy, T. Y.; An, P.; Prasad, P. N.; Lin, Q. *J. Am. Chem. Soc.* **2013**, *135*, 16766.
29. Wang, Y.; Song, W.; Hu, W. J.; Lin, Q. *Angew. Chem.* **2009**, *48*, 5330.
30. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518.
31. Boger, D. L. *Chem. Rev.* **1986**, *86*, 781.
32. Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642.
33. Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 8168.
34. Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 12156.
35. (a) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. *Chembiochem* **2007**, *8*, 232; (b) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem.* **2008**, *47*, 8220; (c) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. *Chembiochem* **2006**, *7*, 134.
36. Chalker, J. M.; Wood, C. S.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346.
37. Li, N.; Lim, R. K.; Edwardraja, S.; Lin, Q. *J. Am. Chem. Soc.* **2011**, *133*, 15316.
38. Li, J.; Lin, S.; Wang, J.; Jia, S.; Yang, M.; Hao, Z.; Zhang, X.; Chen, P. R. *J. Am. Chem. Soc.* **2013**, *135*, 7330.
39. Marriott, G.; Ottl, J. *Methods Enzymol.* **1998**, *291*, 155.
40. Gautier, A.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2011**, *133*, 2124.
41. Hemphill, J.; Chou, C.; Chin, J. W.; Deiters, A. *J. Am. Chem. Soc.* **2013**, *135*, 13433.
42. Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R. *Nat. Chem.* **2014**, *6*, 352.
43. Tsukiji, S.; Wang, H.; Miyagawa, M.; Tamura, T.; Takaoka, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2009**, *131*, 9046.
44. (a) Tamura, T.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2012**, *134*, 2216; (b) Takaoka, Y.; Sun, Y.; Tsukiji, S.; Hamachi, I. *Chem. Sci.* **2011**, *2*, 511.
45. (a) Cho, H. J.; Gee, H. Y.; Baek, K. H.; Ko, S. K.; Park, J. M.; Lee, H.; Kim, N. D.; Lee, M. G.; Shin, I. *J. Am. Chem. Soc.* **2011**, *133*, 20267; (b) Takahashi, M.; Kawamura, A.; Kato, N.; Nishi, T.; Hamachi, I.; Ohkanda, J. *Angew. Chem.* **2012**, *51*, 509.
46. Tamura, T.; Kioi, Y.; Miki, T.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2013**, *135*, 6782.
47. Otsuki, S.; Nishimura, S.; Takabatake, H.; Nakajima, K.; Takasu, Y.; Yagura, T.; Sakai, Y.; Hattori, A.; Kakeya, H. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1608.
48. Hughes, C. C.; Yang, Y. L.; Liu, W. T.; Dorrestein, P. C.; La Clair, J. J.; Fenical, W. *J. Am. Chem. Soc.* **2009**, *131*, 12094.
49. Yamaguchi, T.; Asanuma, M.; Nakanishi, S.; Saito, Y.; Okazaki, M.; Dodo, K.; Sodeoka, M. *Chem. Sci.* **2014**, *5*, 1021.
50. Fujishima, S. H.; Yasui, R.; Miki, T.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2012**, *134*, 3961.
51. Miki, T.; Fujishima, S. H.; Komatsu, K.; Kuwata, K.; Kiyonaka, S.; Hamachi, I. *Chem. Biol.* **2014**, *21*, 1013.
52. Matsuo, K.; Kioi, Y.; Yasui, R.; Takaoka, Y.; Miki, T.; Fujishima, S.-H.; Hamachi, I. *Chem. Sci.* **2013**, *4*, 2573.
53. (a) Wang, H.; Koshi, Y.; Minato, D.; Nonaka, H.; Kiyonaka, S.; Mori, Y.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2011**, *133*, 12220; (b) Koshi, Y.; Nakata, E.; Miyagawa, M.; Tsukiji, S.; Ogawa, T.; Hamachi, I. *J. Am. Chem. Soc.* **2008**, *130*, 245.
54. Sato, S.; Nakamura, H. *Angew. Chem.* **2013**, *52*, 8681.

Please cite this article in press as: Gong, Y.; Pan, L. *Tetrahedron Lett.* (2015), http://dx.doi.org/10.1016/j.tetlet.2015.03.065