3

Copper-free Click Chemistry

Jeremy M. Baskin and Carolyn R. Bertozzi

3.1 Introduction

Biological systems exhibit astounding chemical complexity. A single mammalian cell contains thousands of proteins, DNA and RNA, glycans, lipids, myriad small-molecule metabolites, and metal ions, all in an aqueous environment. Each of these components displays many chemical functionalities, including nucleophiles, electrophiles, oxidants, and reductants. Within this environment, enzymes choreograph the innumerable chemical transformations that together constitute the life of the cell: catabolic conversion of metabolites into energy, biosynthesis and posttranslational modification of proteins, replication of DNA, and many other biochemical processes.

Chemical biologists have striven to study the molecular intricacies of living systems by labeling individual components or groups of components – within the complexity of the living system – with probes such as fluorophores and affinity tags. This approach permits both the tracking of biomolecules within the living cell by imaging and also the determination of their exact molecular identities and compositions after purification from a cell lysate. A critical aspect of any strategy for labeling a target biomolecule inside a living cell or organism using a chemical reaction is that the reaction must be exquisitely chemoselective.

The term 'bio-orthogonal' – defined as noninteracting with biology – encompasses all of the characteristics of such a chemical reaction. A bio-orthogonal ligation is thus a chemical reaction in which two functional groups selectively react with one another to form a covalent linkage in the presence of all of the functionality in biological systems (Figure 3.1).[1] In order to maximize labeling efficiency, the reaction should display the properties of a 'click' reaction as outlined by Kolb, Finn and Sharpless: rapid kinetics, high

Click Chemistry for Biotechnology and Materials Science Edited by Joerg Lahann
 ISBN: 978-0-470-69970-6

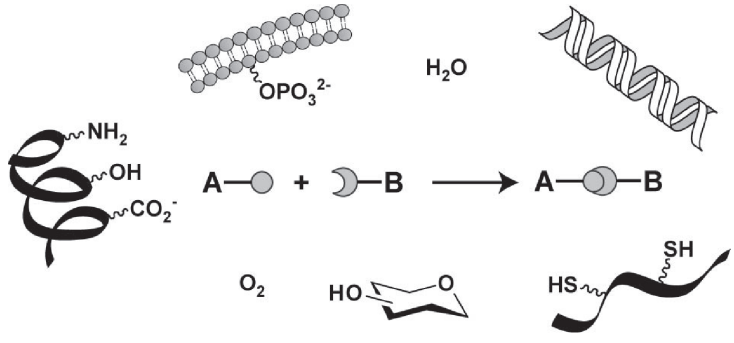


Figure 3.1 *Schematic of a bio-orthogonal reaction. Two molecules (A and B) are ligated together in the presence of diverse biological functionality by virtue of two complementary bio-orthogonal functionalities (gray circle and crescent) that selectively react only with each other to form a covalent bond.*

yield, lack of unwanted by-products, broad solvent compatibility (including water), and readily accessible starting materials.[2]

Arguably the most widely used click reaction is the Cu-catalyzed azide–alkyne cycloaddition reported simultaneously in 2002 by Sharpless and coworkers and Meldal and coworkers.[3,4] This reaction, which utilizes a Cu(I) catalyst to effect the 1,3-dipolar cycloaddition of azides and terminal alkynes to form 1,4-disubstituted 1,2,3-triazoles, displays the characteristics of an ideal click reaction mentioned above (Figure 3.2). It has thus been widely employed throughout medicinal chemistry, materials science and chemical biology,[5–7] applications that are summarized in other chapters in this book. However, the strict requirement of a Cu(I) catalyst precludes the use of CuAAC for labeling within living systems, due to the cytotoxicity of copper.

Many copper-free, and hence nontoxic, alternatives to CuAAC exist. This chapter will begin by describing these copper-free click chemistries, which include condensations of ketones with hydrazide and aminooxy reagents, Staudinger ligation of phosphines and azides, strain-promoted [3 + 2] cycloaddition of cyclooctynes with azides, and various bio-orthogonal ligations of alkenes. After outlining the chemistries, we will present selected applications of copper-free click chemistries in chemical biology in the context of live cells and whole animals.

3.2 Bio-orthogonal Ligations

Amidst the plethora of organic transformations that form a covalent bond between two reactants, very few fit the criteria of bio-orthogonality.[8] Reactions must take place in aqueous

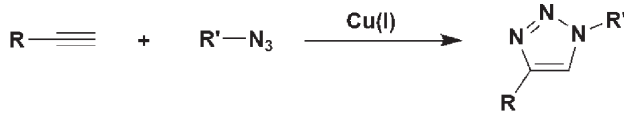


Figure 3.2 *Cu-catalyzed azide–alkyne cycloaddition. Azides and terminal alkynes react in the presence of a Cu(I) catalyst to form 1,4-disubstituted 1,2,3-triazole products.*

solvent, involve no toxic catalysts or reagents, and – most importantly – involve two functional groups that are not present in and which do not cross-react with any functionality present in biological systems. Classic transformations such as bimolecular substitution reactions of alkyl halides and various nucleophiles, as well as the coupling of amines and carboxylic acids to form amides, are not bioorthogonal, as the reagents would have significant cross-reactivity with many amino acid side chains, other cellular nucleophiles and water. Neither are more modern metal-promoted cross-coupling reactions, as they involve toxic catalysts. Even standard protein bioconjugation reactions (e.g. Michael addition of thiols to maleimides) do not qualify as bio-orthogonal, as thiols are ubiquitous within a cellular environment.

An important principle that has emerged in roughly a decade of bio-orthogonal reaction development is that these reactions do not fit a uniform mold. Early work focused on highly selective nucleophile/electrophile coupling reactions, first the condensation of ketones with heteroatom-bound amine reagents (Section 2.1) and later a modified Staudinger reaction of phosphines and azides (Section 2.2). Much work in recent years has shifted to pericyclic reactions, notably 1,3-dipolar cycloadditions of azides and alkynes (Section 2.3). Finally, an emerging area is the development of bio-orthogonal ligations of alkenes (Section 2.4).

3.2.1 Condensations of Ketones and Aldehydes with Heteroatom-bound Amines

Historically, the first bio-orthogonal ligations involved ketone–aldehyde condensation reactions. While ketones and aldehydes can form reversible imine adducts with many amines found in biological systems, this process is thermodynamically unfavorable in water. The use of hydrazides and aminooxy reagents, often called 'α-effect amines' because the heteroatom-bound amine is much more nucleophilic than simple amines, shifts the equilibrium dramatically to the hydrazone and oxime products, respectively (Figure 3.3). For example, the equilibrium constant for the condensation of acetone and hydroxylamine in water is 1×10^6.[9] This reaction is not an optimal bio-orthogonal ligation for live cell applications, however, because it proceeds most efficiently at pH values of 3–6, well below the physiological level (pH 7.4).[9] The kinetics of these condensation reactions can be improved considerably both at acidic and neutral pH by the addition of aniline-based nucleophilic

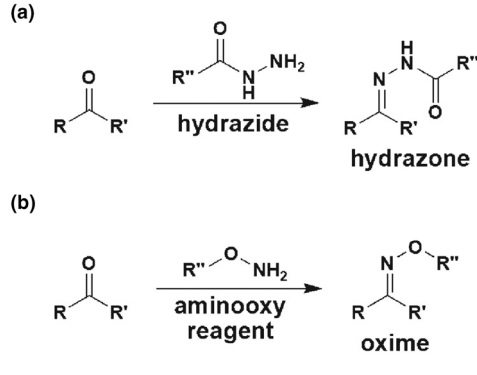


Figure 3.3 *Condensation reactions of ketones and aldehydes with (a) hydrazides to form hydrazones or (b) aminooxy reagents to form oximes.*

catalysts, although the requisite high concentrations of these reagents (up to 0.1 M) might preclude their use in living systems, wherein toxicity becomes a concern.[10] Further, although ketones and aldehydes are absent from the cell surface and from macromolecules within the cell, these functional groups are present within many intracellular metabolites (e.g. glucose, pyruvate) and can thus be thought of as semi-bioorthogonal. For these reasons, the use of ketone/aldehyde ligations for labeling of biomolecules within living systems has been somewhat limited.

3.2.2 Staudinger Ligation of Phosphines and Azides

A major breakthrough occurred in 2000, with the introduction of the azide as a functional handle for bio-orthogonal chemical reactions. Unlike the ketone, the azide is truly bio-orthogonal – it is essentially unreactive with any biological functional groups under physiological conditions. The slow reaction of alkyl azides with thiols at physiological pH is usually insignificant on the timescale of biological experiments.[11] The azide is also absent from virtually all biological systems. The lone exception reported to date is an azide-containing natural product isolated from the dinoflagellate *Gymnodinium breve*, a species of red algae that is responsible for the production of toxic red tides along the Gulf Coast of Florida.[12,13]

Despite its kinetic stability, the azide is thermodynamically a high-energy species prone to specific reactivity both as a soft electrophile and as a 1,3-dipole.[14] The Staudinger ligation of phosphines and azides takes advantage of the former type of reactivity. This reaction is a modification of the classical Staudinger reduction of azides with phosphines, in which an aza-ylide intermediate collapses in water to amine and phosphine oxide products. However, construction of a triarylphosphine reagent with an ester positioned *ortho* to the phosphorus atom enabled trapping of the aza-ylide intermediate, in aqueous conditions, as an amide, thereby covalently 'ligating' the two molecules (Figure 3.4).[15] The Staudinger ligation has many positive attributes: like azides, phosphines are absent from biological systems, and due to the relative dearth of soft electrophiles in biology, triarylphosphines have no crossreactivity with endogenous biomolecules. Furthermore, the phosphine reagents – derivatized as conjugates of many different epitopes such as biotin, fluorophores, and peptides – demonstrate no toxicity.

The mechanism of the Staudinger ligation allows for a wide variety of modifications and clever applications. For example, inversion of the orientation of the ester has enabled the development of a so-called 'traceless' Staudinger ligation in which an amide is formed and the phosphine oxide byproduct is expelled (Figure 3.5).[16–18] As this variant forms native amides, it is a powerful tool for peptide ligation; however, lower yields and slower kinetics have hindered its use as a bioorthogonal reaction for sensitive detection of azides in living systems.

Further, the Staudinger ligation can be used to create 'smart' phosphine reagents that become fluorescent upon reaction with azides. Such fluorogenic reagents can be tremendously advantageous for biological imaging experiments because they permit dynamic monitoring of labeling reactions and eliminate problems associated with washing away unreacted phosphine probe. Two different approaches have been taken to tackle this challenge, each exploiting a different aspect of the reaction mechanism (Figure 3.6). First, a fluorogenic coumarin–phosphine reagent was synthesized that is nonfluorescent due to quenching by

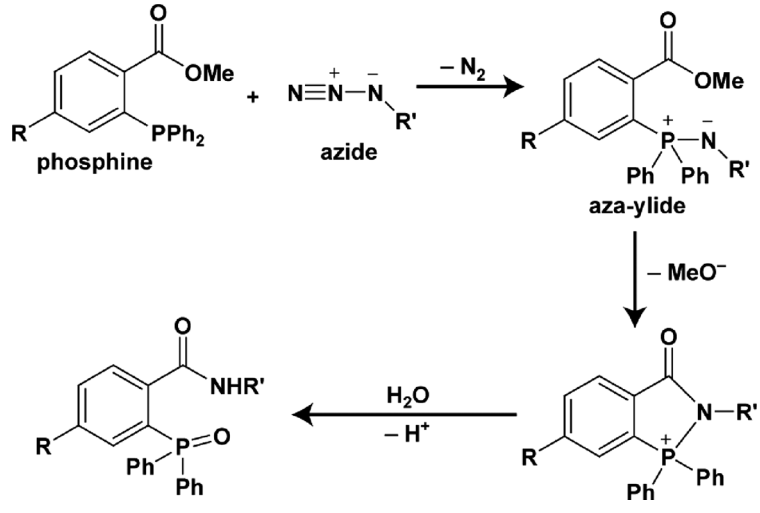


Figure 3.4 *The Staudinger ligation of azides and triarylphosphines. The mechanism begins with nucleophilic attack of the phosphine on the terminal nitrogen atom of the azide, liberating N_2. The resultant aza-ylide intermediate is trapped intramolecularly by the methyl ester, releasing methanol and forming an amide and the phosphine oxide.*

the phosphine lone pair of electrons. Upon Staudinger ligation with azides, the conversion to a phosphine oxide eliminates this quenching and the fluorophore 'turns on'.[19] Unfortunately, this reagent has not seen widespread use because nonspecific air oxidation of the phosphine also results in fluorophore turn-on, and oxygen cannot be avoided in most biological systems. A second approach utilizes intramolecular fluorescence resonance energy transfer (FRET) quenching. Here, a phosphine–fluorophore conjugate contains a FRET

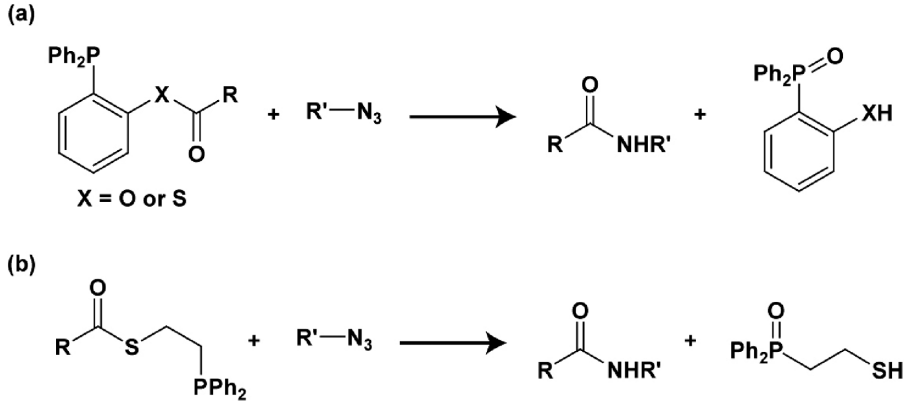


Figure 3.5 *'Traceless' Staudinger ligation strategies that enable amide formation concomitant with release of the phosphine oxide by-product.*

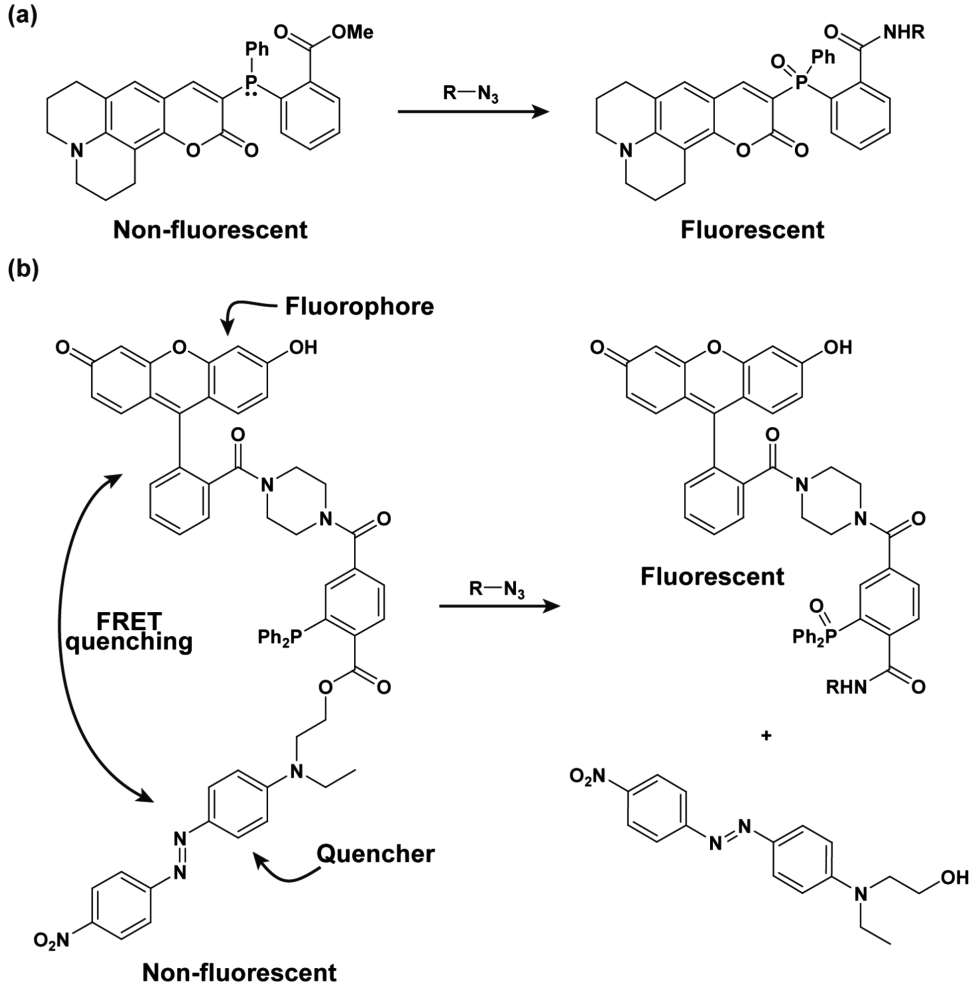


Figure 3.6 *Two approaches to the design of fluorogenic phosphine probes. (a) A coumarin phosphine conjugate is nonfluorescent due to lone-pair quenching of the fluorophore; upon Staudinger ligation with azides, the product is fluorescent because the phosphine is oxidized. (b) A FRET-based fluorescein phosphine conjugate is nonfluorescent due to the action of an intramolecular FRET quencher (Disperse Red 1), attached via an ester. Upon Staudinger ligation, the quencher is liberated and the product is fluorescent.*

quencher attached via the ester and is hence nonfluorescent. Upon reaction with azides, the ester is cleaved, liberating the quencher and turning on fluorescence. As nonspecific phosphine oxidation generates a silent (though nonfunctional) reagent, this probe has been successfully employed for imaging azide-functionalized glycans in living cells.[20]

Despite the myriad applications that the Staudinger ligation enables by virtue of its exquisite selectivity, it suffers from relatively slow reaction kinetics.[21] Accordingly, cell labeling reactions can often require an hour or longer. Thus, for studying dynamic

biological processes that occur on faster timescales, improvements to intrinsic reaction kinetics have been sought. A mechanistic study revealed that, for triarylphosphine and alkyl azide reactants, the rate-determining step is the initial nucleophilic attack of the phosphine on the terminal nitrogen atom of the azide.[21] Electron-rich phosphine reagents that could accelerate this step have been designed. Although these reagents performed faster in the Staudinger ligation, they were also highly susceptible to nonspecific air oxidation, which both lowered reaction yields and made the reagents difficult to handle and store.[21] Fundamentally, it is difficult, if not impossible, to decouple improved kinetics in reactions with azides from increased propensity toward nonspecific air oxidation, as both properties ultimately arise from the nucleophilic character of the phosphine. Therefore, for the study for dynamic biological processes, it became necessary to investigate the alternative mode of reactivity of the azide as a 1,3-dipole.

3.2.3 Copper-free Azide–Alkyne Cycloadditions

In addition to being a soft electrophile, the azide is also a 1,3-dipole, and as such, it can undergo [3 + 2] cycloadditions with alkynes as first reported by Michael in 1893[22] and studied extensively by Huisgen.[23] The thermal cycloaddition of azides and unactivated alkynes to form 1,2,3-triazoles is a highly exergonic process ($\Delta G \approx -61$ kcal/mol), although it is kinetically hindered by an activation barrier of 26 kcal/mol.[24] Thus, elevated temperatures or pressures are necessary to accelerate the reaction of azides and simple alkynes. As mentioned earlier, the use of copper catalysis is a highly effective method to lower the activation energy and effect room-temperature triazole formation between azides and terminal alkynes; however, for biological labeling applications, CuAAC is not an ideal bio-orthogonal ligation due to the cytotoxicity of copper.

In an effort to activate the alkyne component for [3 + 2] cycloaddition with azides in a biocompatible manner, we explored the use of ring strain. In 1961, Wittig and Krebs demonstrated that cyclooctyne, the smallest stable cycloalkyne, reacts with azides to form the corresponding 1,2,3-triazole (Figure 3.7).[25] This reaction's fast kinetics – the authors wrote that it 'proceeded like an explosion' – are due to roughly 18 kcal/mol of ring strain

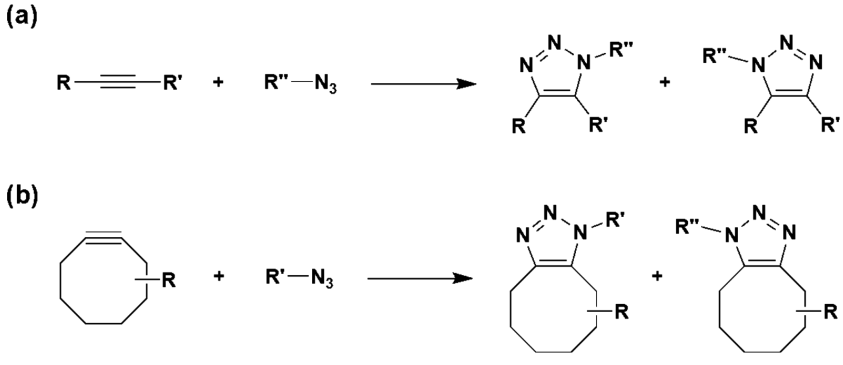


Figure 3.7 *Huisgen 1,3-dipolar cycloadditions of azides and alkynes to form regioisomeric mixtures of 1,2,3-triazoles. (a) Cycloaddition involving azides and linear alkynes. (b) Copper-free, strain-promoted cycloaddition between azides and cyclooctynes.*

in the cyclooctyne starting material, much of which is released in the transition state of the reaction.[26] Building upon this precedent, we synthesized a biotinylated cyclooctyne conjugate and demonstrated that it could selectively label azide-modified proteins *in vitro* and on live cell surfaces.[27] Unlike the case of CuAAC, which exclusively forms the 1,4-disubstituted triazoles, this 'strain-promoted cycloaddition' forms a roughly 1:1 mixture of regioisomeric 1,2,3-triazoles. Importantly, this reagent required no copper catalyst and displayed no toxicity; however, its kinetics were still considerably slower than those of CuAAC and comparable to those of the Staudinger ligation.[27]

We thus set out to improve the kinetics of this strain-promoted cycloaddition by installing LUMO-lowering, electron-withdrawing fluorine atoms adjacent to the cyclooctyne, producing monofluorinated and difluorinated reagents that were roughly 2-fold and 40-fold faster than the original reagent, respectively (Figure 3.8).[28,29] In particular, this last reagent, termed DIFO (for difluorinated cyclooctyne), labels azide-bearing proteins with similar kinetics to CuAAC and has been particularly useful for imaging cultured cells and live animals such as developing zebrafish embryos (see Section 3.3). Other recent additions to the cyclooctyne toolkit include second-generation DIFO reagents with more facile synthetic routes[30] and a highly water-soluble azacyclooctyne reagent designed to eliminate problems associated with nonspecific binding of the more hydrophobic cyclooctyne reagents to

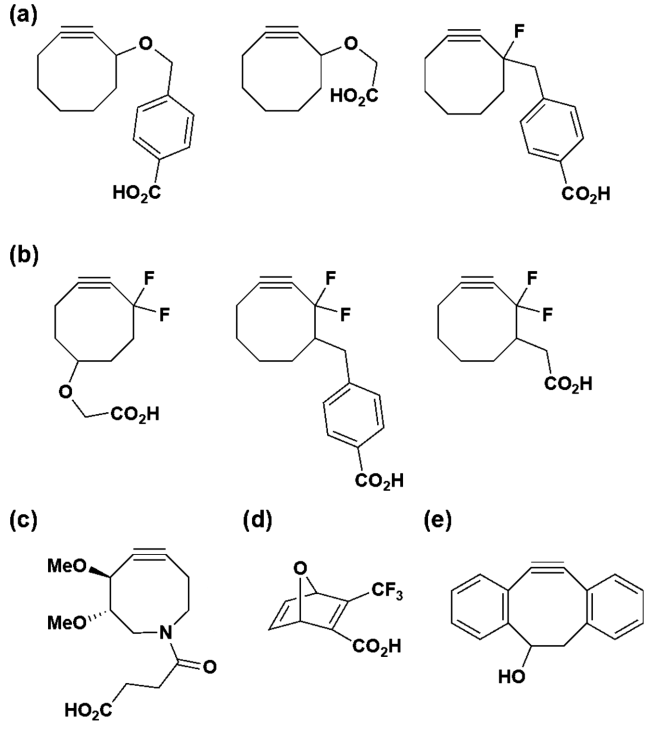


Figure 3.8 *Structures of strained alkynes or alkenes for Cu-free [3 + 2] cycloadditions with azides in biological systems. These include (a) simple cyclooctyne, (b) difluorinated cyclooctyne (DIFO), (c) azacyclooctyne, (d) oxanorbornadiene, and (e) dibenzocyclooctyne probes.*

biological membranes and serum proteins *in vivo* (Figure 3.8).[31] In addition, strained oxanorbornadiene and dibenzocyclooctyne reagents have emerged for selectively labeling azido biomolecules (Figure 3.8).[32,33] Their kinetic rate constants range from comparable to the original cyclooctyne for the former and comparable to DIFO for the latter. Collectively, these studies suggest a rich future for the development of Cu-free reagents for sensitive detection of azides in biological systems.

3.2.4 Bioorthogonal Ligations of Alkenes

Ideally, a slew of bio-orthogonal ligations would exist that are each orthogonal to one another. This scenario would permit the simultaneous use of multiple reactions to label different components in the same system. Thus, to complement the plethora of azide-specific bio-orthogonal ligations, researchers have been recently looking to alternatives. One emerging example is so-called 'photoclick' chemistry. This reaction involves the light-induced decomposition of 2,5-diaryltetrazoles to form dipolar nitrile imine intermediates, which undergo subsequent [3 + 2] cycloaddition with alkenes to yield stable pyrazoline adducts (Figure 3.9). Like the [3 + 2] azide–alkyne cycloaddition, this reaction was first explored by Huisgen many decades ago and recently revisited in the context of bio-orthogonal labeling. Lin and coworkers have demonstrated that various diaryltetrazole probes can, upon irradiation with 302 nm light, selectively label alkene-containing proteins *in vitro* and in living bacterial cells with rapid kinetics.[34,35] Although the irradiation time was limited to a few minutes, this relatively high-energy ultraviolet light can be damaging to living cells. Thus, Lin and coworkers developed second-generation diaryltetrazole reagents that are activated at a more biocompatible 365 nm;[36] in the future, this approach may be extended to biomolecule labeling in mammalian cells.

Davis and coworkers demonstrated that olefin metathesis can be used to selectively target allyl sulfides on isolated proteins.[37] To accomplish this task, cysteine residues were first converted to the corresponding allyl sulfides, and alkene-functionalized probes (e.g. sugars, polyethyleneglycol) were appended via cross-metathesis using the Hoveyda–Grubbs second-generation catalyst. The efficacy of this reaction in living systems has not yet been established, however, and toxicity is always a concern when dealing with organometallic catalysts.

Other bio-orthogonal ligations involving alkenes include the ligation of strained alkenes such as *E*-cyclooctene or norbornene with tetrazine reagents.[38,39] These inverse electron-demand [4 + 2] Diels–Alder reactions proceed with rapid kinetics and in one case displays

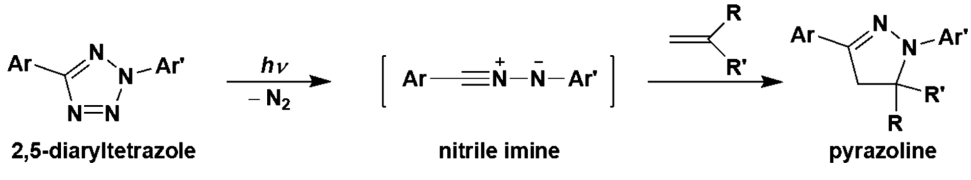


Figure 3.9 *'Photoclick' chemistry. Light-induced decomposition of 2,5-diaryltetrazoles creates dipolar nitrile imine intermediates, which react* in situ *with alkenes to form pyrazoline products.*

bioorthogonality in the context of cell-surface labeling.[39] However, both components in this reaction (i.e. the strained alkene and the tetrazine reagent) are quite large compared with an azide or a ketone. This characteristic presents a potential limitation for metabolic labeling experiments, where biosynthetic enzymes often tolerate only minor structural changes to their natural substrates (see Sections 3.3.2–3.3.4).

3.3 Applications of Copper-free Click Chemistries

In order to take advantage of the click reactions for biomolecule labeling, one of the two bioorthogonal functional groups that participates in the click reaction (e.g. ketone, azide, etc.), often termed a 'chemical reporter', must first be introduced into biomolecules. Various methods have been developed for this purpose.[1] This section will discuss the use of activity-based inhibitors for labeling classes of enzymes (Section 3.1), genetically encoded peptide tags and unnatural amino acids for labeling proteins (Section 3.2), unnatural monosaccharides for metabolic labeling of glycans (Section 3.3) and unnatural lipids and nucleotide precursors for labeling lipids and nucleic acids (Section 3.4).

3.3.1 Activity-based Profiling of Enzymes

Covalent inhibitors of enzyme activity have seen widespread use as both pharmaceutical agents and tools in basic research. In many cases, inhibitors are not selective for a specific enzyme but instead can label entire classes of enzymes (e.g. serine proteases). Cravatt has pioneered the use of functionalized enzyme inhibitors to probe enzymes based on their activity. This technique, termed activity-based protein profiling (ABPP), involves the covalent labeling of active enzymes within living cells or lysates, and even within live animals, with a mechanism-based inhibitor derivatized either directly with an affinity tag or fluorophore or with a bio-orthogonal chemical reporter (Figure 3.10).[40,41] In the latter case, a second step, the bio-orthogonal click reaction, can be employed to append the desired tag or probe.

This strategy has been highly successful for labeling classes of enzymes with well-characterized covalent, active site-directed inhibitors (e.g. serine hydrolases[42,43] and glycosidases[44]). In early studies using ABPP, active enzymes were targeted from cell or tissue lysates and characterized by mass spectrometry following affinity capture. However,

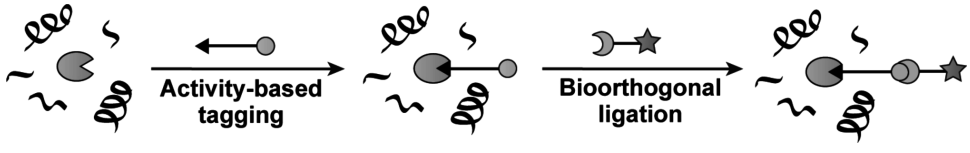


Figure 3.10 *Activity-based protein profiling. An enzyme is selectively labeled within a live cell or cell lysate with an activity-based probe that contains a mechanism-based 'warhead' for covalent labeling of the enzyme (triangle) and a bio-orthogonal chemical reporter (circle). The labeled enzymes can be detected in a subsequent bio-orthogonal ligation with a conjugate of a suitable reaction partner (crescent) and a biophysical probe (star).*

labeling active enzymes with imaging probes within intact cells or whole organisms could provide critical spatial information that is lost in cell and tissue lysates. Ploegh and coworkers demonstrated the use of an azido probe and the Staudinger ligation to visualize active cathepsins, an important class of cysteine proteases, in cells.[45] In principle, any enzyme class can be studied using ABPP provided that it can be targeted with a selective inhibitor.

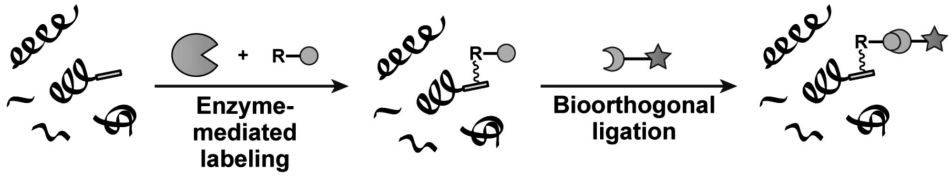
3.3.2 Site-specific Labeling of Proteins

For dynamic imaging of enzymes, activity-based approaches – which by design involve the abolition of catalytic activity – may not be desirable, as they irreversibly perturb the physiological system under study. As well, many proteins involved in important biological processes are not enzymes, including many cell-surface receptors and transporters, transcription factors, and structural proteins. Thus, more general and less invasive approaches to visualizing proteins are necessary. Classically, proteins can be labeled at the genetic level by fusion to fluorescent proteins.[46] Although these probes have been widely used for biological imaging studies, their large size can interfere with the functions of many proteins.

As an alternative to fluorescent proteins, many small peptide-based methods have been developed for site-specific labeling of proteins.[47,48] These methods typically exploit short peptide sequences that can direct the chemical or enzymatic attachment of probes. In addition to the small size of the peptide and small molecule probe, other advantages of these approaches include temporally controlled labeling, multicolor labeling and the use of imaging modalities other than fluorescence. Here, we will discuss peptide tags that employ bioorthogonal chemistry as a central component of the labeling strategy.

In pioneering work, Tsien and coworkers demonstrated that bioorthogonality could be achieved by arranging natural amino acids to form a unique chemical environment. They showed in 1998 that proteins containing an engineered tetracysteine motif (CCXXCC) could be selectively tagged in live cells with biarsenical derivatives of the organic fluorophores.[49] The two most common probes used in this technology, known as FlAsH (fluorescein arsenical hairpin binder) and ReAsH (resorufin arsenical hairpin binder), are initially nonfluorescent but become fluorescent upon chelation to tetracysteine-tagged proteins in live cells.[50] Treatment of cells expressing a tetracysteine-containing protein first with FlAsH and then with ReAsH in a pulse-chase manner enables identification of ‘old’ and ‘new’ populations of the same protein.[51] Inspired by this work, Schepartz and coworkers recently developed an analogous technology for selectively labeling tetraserine motifs on engineered proteins within living cells using a fluorogenic, bis-boronic acid derivative of rhodamine.[52]

The FlAsH/ReAsH tetracysteine method has been widely used in biological studies, including the imaging of mRNA translation,[53] G-protein-coupled receptor activation,[54] amyloid formation,[55] viral trafficking patterns,[56] bacterial secretion systems[57] and membrane protein conformational changes.[58] In addition to its fluorescent properties, ReAsH can initiate the photoconversion of diaminobenzidine, which ultimately provides contrast for electron microscopy, permitting a high-resolution image of the labeled proteins in fixed cells.[50] In some instances, however, application of the biarsenical dyes can lead to background staining due to a moderate affinity for endogenous monothiol and dithiol motifs found within the cell.[59] Thus, alternative techniques have emerged to address this important



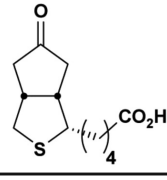
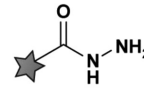
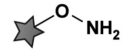
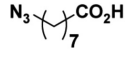
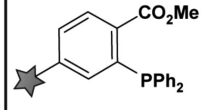
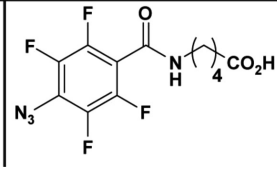
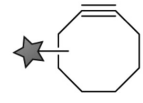
Enzyme	Probe	Peptide tag	Bioorthogonal reagents
Biotin ligase		Acceptor peptide (15 aa)	 Hydrazides
Formylglycine-generating enzyme	N/A	Aldehyde tag (6 aa)	 Aminooxy reagents
Lipoic acid ligase		Lipoic acceptor peptide (17 aa)	 Phosphines
Lipoic acid ligase (W37V)			 Cyclooctynes

Figure 3.11 *Site-specific labeling of proteins via recombinantly fused peptide tags. Recombinant proteins bearing a genetically fused peptide tag are selectively endowed with bioorthogonal functionality within a live cell or cell lysate by the action of a specific ligase or enzyme. Shown are enzyme-based methods for site-specifically labeling proteins with ketones, aldehydes, and azides.*

issue of specificity, and critical to the success of these techniques is the use of bioorthogonal ligations discussed in Section 2.

Ting and coworkers have developed two systems that capitalize on the ability of a bacterial cofactor ligase to covalently label a specific lysine residue within in a short peptide acceptor sequence with unnatural versions of the cofactor (Figure 3.11). The first approach utilizes biotin ligase, an enzyme that can attach biotin via its carboxylic acid to lysine side chains within a 15-residue consensus sequence. The target protein of interest is genetically modified to contain this 15-amino acid 'acceptor peptide', to which biotin ligase catalyzes the attachment of a synthetic ketone-containing biotin isostere.[60] The ketone biotin ligase technology has been used to image receptor dynamics in live cells.[60] A more recent addition to the peptide-labeling toolkit is the use of lipoic acid ligase, which accepts a variety of azide and alkyne-containing analogs of lipoic acid, a naturally occurring cofactor.[61] This

approach is an improvement over the biotin ligase technology as it allows the use of the superior bioorthogonal ligations of azides; further, rational design of lipoic acid ligase has enabled use of larger reporters such as aryl azides, which can be used as photoaffinity labels for dissecting protein–protein interactions.[62]

Both of these enzyme-mediated methods require the chemical synthesis of an unnatural cofactor, although the short ω-azido fatty acids recognized by lipoic acid ligase are simple to prepare. Another approach involves direct conversion of a specific amino acid side chain within a consensus peptide to a bioorthogonal functionality; this concept is embodied in the 'aldehyde tag' technology. This method takes advantage of the formylglycine-generating enzyme (FGE), an enzyme whose natural function is to co-translationally convert cysteine to an aldehyde-containing formylglycine residue in the active site of sulfatases.[63] FGE recognizes a six-residue consensus sequence (LCTPSR), which we discovered can be modified in the context of heterologous proteins. Thus, expression of recombinant proteins bearing this six-residue 'aldehyde tag' sequence leads to the production of proteins bearing bio-orthogonal aldehyde residues in both bacterial and mammalian systems.[64,65]

An alternative approach to site-specific protein labeling is to re-engineer the protein synthesis machinery to incorporate an unnatural amino acid at precisely one position in the protein (Figure 3.12). To accomplish this, Schultz and others have made use of the 'amber' stop codon and its corresponding tRNA, termed the amber suppressor. Mutants of amino acid tRNA-synthetase (aaRS) enzymes were evolved that charge the amber suppressor tRNA with various unnatural amino acids instead of the cognate amino acid.[66] This approach has been highly successful for the site-specific incorporation of unnatural amino acids that possess fluorophores[67–69] and photochemical crosslinkers,[70,71] as well as bio-orthogonal chemical reporters such as ketones, azides, and terminal alkynes.[70,72–74] This method for site-specific protein labeling, widely applied in bacteria and yeast, has recently been expanded for use in mammalian systems.[72,75]

To achieve site-specific protein labeling, the use of the amber suppressor technology requires considerable genetic manipulation. By contrast, a 'residue-specific' method for globally replacing one amino acid with an unnatural surrogate is much simpler to employ (Figure 3.12)[76]. In this approach, unnatural analogs of methionine[77,78] or phenylalanine[79,80] bearing azides and alkynes (as well as numerous other unnatural functional groups) are recognized by the cell's endogenous translational machinery and utilized for protein synthesis when their natural counterparts are in short supply. Pioneered by Tirrell and coworkers,[78,81] residue-specific protein labeling has been used to create novel protein-based biomaterials.[76] Additionally, it can be used to monitor global *de novo* protein synthesis. Termed bio-orthogonal noncanonical amino acid tagging (BONCAT), this technique can be used to visualize[82] and identify[81] newly synthesized proteins in mammalian cells and has been applied to generate static snapshots in fixed cells using CuAAC. By employing the nontoxic copper-free [3 + 2] cycloaddition with cyclooctyne probes, Tirrell and coworkers used the technique in living bacterial cells to discover novel catalytic activities through a series of *in vivo* evolution experiments.[83]

3.3.3 Metabolic Labeling of Glycans

Residue-specific protein labeling essentially constitutes the metabolic labeling of proteins with unnatural amino acids, and analogous technologies exist for probing other classes of

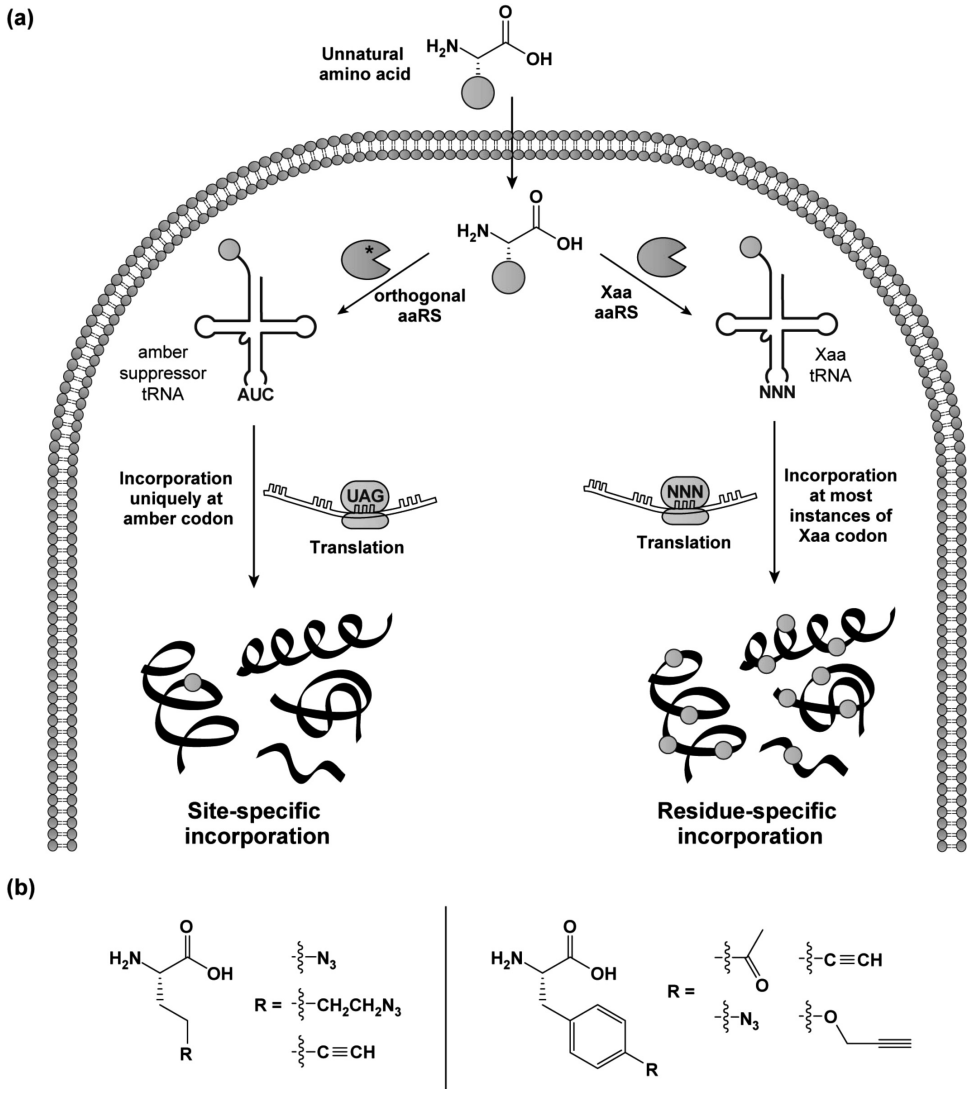


Figure 3.12 *Incorporation of unnatural amino acids into newly synthesized proteins. (a) General schematic for site-specific incorporation of unnatural amino acids using amber suppressor technology (left side) or residue-specific incorporation of unnatural amino acids (right side). (b) Structures of unnatural amino acids containing ketones, azides, and alkynes for incorporation into proteins using these methods.*

biomolecules as well. Glycans, which are linear or branched chains of sugars, constitute a diverse class of biomolecules that are often found as posttranslational modifications of proteins or covalently bonded to lipids. These biopolymers, found both within the cell and at the cell surface, participate in many physiological processes, including organ development,[84] cancer[85] and host–pathogen interactions.[86]

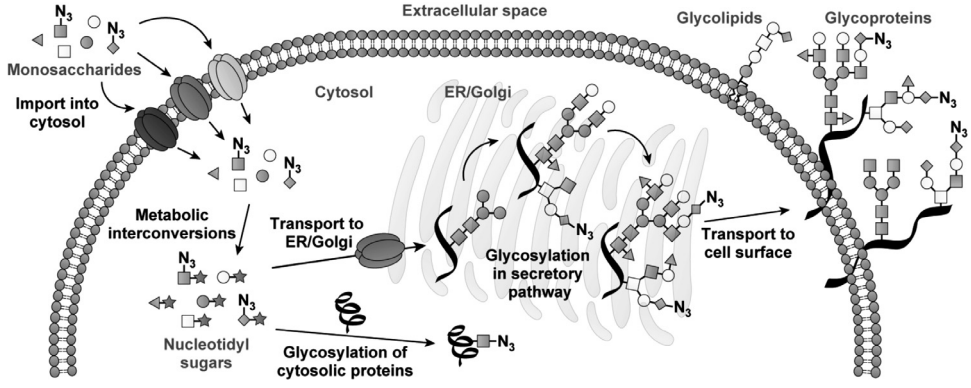


Figure 3.13 *Metabolic oligosaccharide engineering for labeling glycans with unnatural sugars. Monosaccharides are imported into the cytosol (or biosynthesized* de novo*) and metabolically converted to activated nucleotidyl sugars. Unnatural monosaccharides are represented here for clarity as containing the azide group (N_3), but many other unnatural monosaccharides have been employed for metabolic labeling of glycans (see Figure 3.14). These activated sugars are either directly attached to cytosolic proteins or transported into the ER and/or Golgi apparatus and then appended onto protein and lipid scaffolds by the action of glycosyltransferase enzymes. The majority of the resultant glycoproteins and glycolipids are then transported to the cell surface, although some are retained in the ER and Golgi apparatus or trafficked to organelles such as the lysosome, and others still are secreted into the extracellular space (not shown).*

Glycans can be labeled with unnatural monosaccharides bearing chemical reporters in a process termed 'metabolic oligosaccharide engineering'.[87] This approach enables both the visualization of glycans on cells and tissues as well as the profiling of glycosylation at the proteomic level. Analogous to residue-specific protein labeling, synthetic keto, azido or alkynyl sugars can hijack the glycan biosynthetic machinery and label glycans both on cell surfaces and intracellularly (Figure 3.13). Much work in this area has been devoted to the study of cell-surface sialic acids, which can be modified by metabolism of keto,[88] azido[15] and alkynyl[89] analogs of two natural precursors, *N*-acetylmannosamine (ManNAc) and *N*-acetylneuraminic acid (Figure 3.14).[90,91] Other classes of glycoconjugates that have been labeled with unnatural sugars bearing chemical reporters include fucosylated glycans,[89,92,93] mucin-type *O*-glycans[94,95] and cytosolic/nuclear *O*-GlcNAcylated proteins[96] (Figure 3.14).

Imaging of glycans has been performed using a variety of fluorophore conjugates of phosphines[20,97] and cyclooctynes.[29] In particular, the rapid kinetics of the Cu-free [3 + 2] cycloaddition using DIFO enabled us to measure, for the first time in living cells, the dynamics of glycan trafficking (Figure 3.15).[29] We have also extended metabolic oligosaccharide engineering to living animals. Mice and zebrafish treated with appropriate azidosugars can be metabolically labeled at their sialic acid residues and mucin-type *O*-glycans.[98–100] In zebrafish, changes in glycosylation during embryonic development were imaged using DIFO-fluorophore reagents.[100] Further, the Staudinger ligation and strain-promoted [3 + 2] azide–alkyne cycloaddition proceed readily within living mice, suggesting future *in vivo* imaging efforts using this powerful model organism for human disease.[29,98]

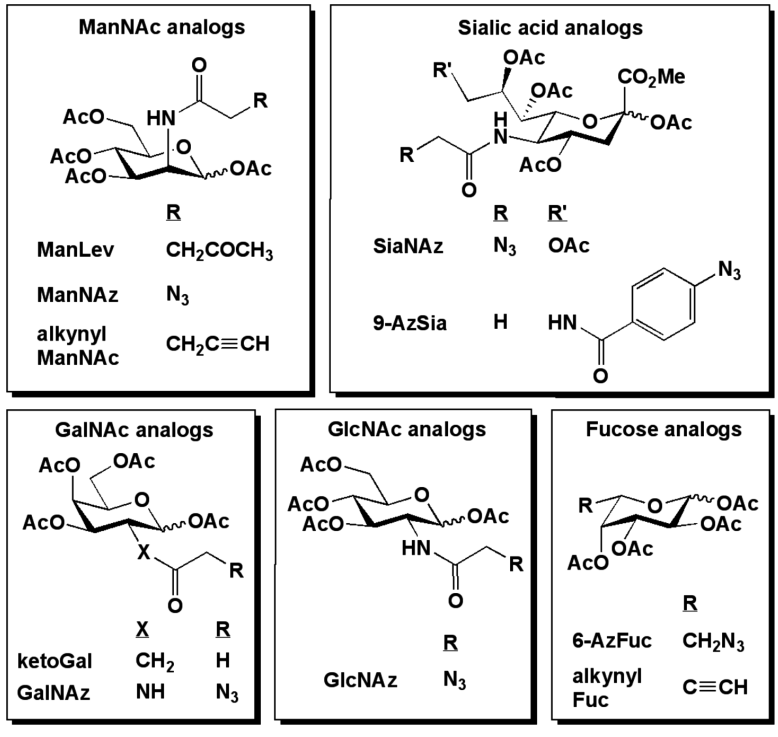


Figure 3.14 *Structures of unnatural monosaccharides containing ketone, azide or alkyne functionality used as substrates for metabolic labeling of glycans.*

3.3.4 Metabolic Targeting of Other Biomolecules with Chemical Reporters

In addition to proteins and glycans, other classes of biomolecules are being targeted using the bio-orthogonal chemical reporter strategy. Most notable among these are lipid-modified proteins, which possess *N*-myristoylation of *N*-terminal glycine residues, or *S*-palmitoylation or farnesylation of cysteine residues. Various groups have demonstrated metabolic labeling of lipidated proteins with azidolipid precursors, which allows for probing the modification in living systems.[101–104] Another protein posttranslational modification that has been probed using chemical reporters is the pantetheinylation. Burkart and coworkers have employed azido and alkynyl pantetheine analogs for metabolic labeling of acyl carrier proteins in living cells.[105] Rajski and coworkers have used azide-derivatized, aziridine-containing *S*-adenosylmethionine analogs for probing substrates of DNA methyltransferases *in vitro*.[106] Finally, Salic and Mitchison reported unnatural nucleotides, ethynyl and azido deoxyuridine, that are incorporated into replicating DNA within living cells and tissues.[107] These probes, after subsequent derivatization with CuAAC reagents, permit the mild detection of dividing cells within fixed cells and tissues and serve as a modern replacement for the bromodeoxyuridine (BrdU) assay that is routinely used to detect dividing cells. Using a similar approach, Jao and Salic were able to probe RNA synthesis in cells and animal tissues using ethynyl uridine. This unnatural nucleotide was incorporated into

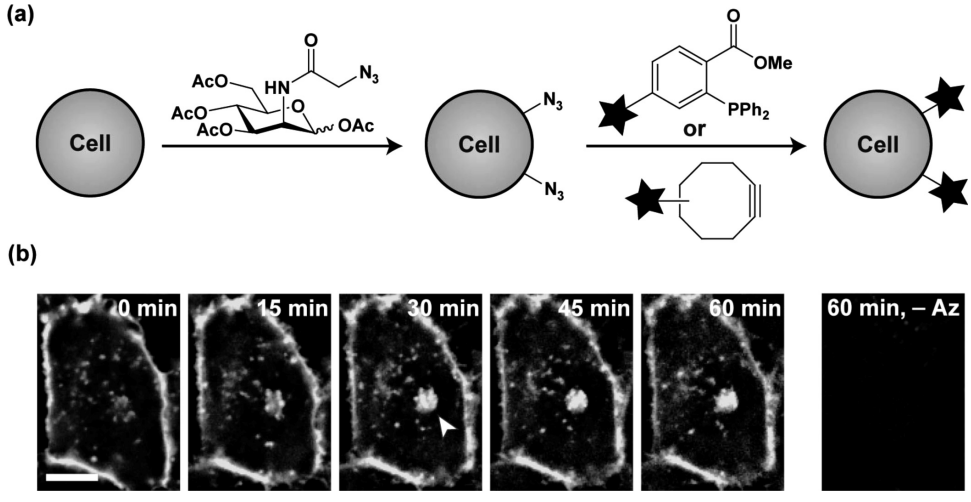


Figure 3.15 *Dynamic imaging of glycans in living cells using metabolic labeling and bioorthogonal chemistries. (a) Experimental setup: cells are metabolically labeled with an azidosugar (ManNAz is shown here as an example), rinsed and then reacted with a fluorescent phosphine or cyclooctyne conjugate. (b) Dynamic imaging of sialic acids using ManNAz and DIFO-Alexa Fluor 488. Chinese hamster ovary cells were metabolically labeled with ManNAz for 3 days, rinsed, reacted with DIFO-Alexa Fluor 488 for 1 min, and imaged by epifluorescence microscopy. Images were acquired every 15 min for 1 h. The image on the right (60 min, –Az) indicates a separate sample metabolically labeled with the control sugar ManNAc. The arrowhead indicates a population of glycans that have undergone endocytosis after the Cu-free click labeling reaction. Scale bar: 5 μm.*

newly synthesized RNA during transcription, and its presence was detected by CuAAC using azido fluorophores.[108]

3.4 Summary and Outlook

In this chapter, we have discussed the development and implementation of bio-orthogonal ligations for labeling biomolecules in living systems. The strategy relies both on the incorporation of a chemical reporter into biomolecules and the subsequent detection of the reporter using bio-orthogonal chemistries. Initial efforts were centered on condensation reactions of ketones and aldehydes, and more recent work has focused on reactions of azides with triarylphosphines and activated alkynes. In particular, copper-free [3 + 2] cycloaddition with strained alkynes presents a promising area for immediate application and future reagent development using the principles of physical organic chemistry. Reagents with faster kinetics could enable more sensitive detection of azides *in vivo*, and fluorogenic alkynes for copper-free [3 + 2] cycloaddition would permit real-time visualization of biological events.

Despite the numerous advances propelled by these azide-specific chemistries, there is a need for alternatives to azides. Currently, methods to image multiple biomolecules

simultaneously are limited by the number of ligation chemistries that are orthogonal to one another. Fertile ground for the search for new biocompatible, copper-free click chemistries includes other pericyclic reactions and photochemistry. The initial reports on 'photoclick' chemistry, a cross-section of these two areas, represent the tip of the iceberg, as a vast literature of classic organic chemistry can be revisited in the modern context of bio-orthogonal ligations. This expanded toolkit of copper-free click chemistries should see wide application in chemical biology in applications ranging from biomaterial design to *in vivo* imaging.

Acknowledgments

C.R.B. thanks the National Institutes of Health (GM058867) for financial support of the work from the authors' laboratory, and J.M.B. was supported by National Science Foundation and National Defense Science and Engineering predoctoral fellowships. We thank Pamela Chang and Matthew Hangauer for critical reading of the chapter.

References

(1) J.A. Prescher and C.R. Bertozzi, (2005), Chemistry in living systems, *Nat. Chem. Biol.*, **1**, 13–21.
(2) H.C. Kolb, M.G. Finn and K.B. Sharpless, (2001), Click chemistry: diverse chemical function from a few good reactions, *Angew. Chem. Int. Edn*, **40**, 2004–2021.
(3) V.V. Rostovtsev, L.G. Green, V.V. Fokin and K.B. Sharpless, (2002), A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes, *Angew. Chem. Int. Edn*, **41**, 2596–2599.
(4) C.W. Tornoe, C. Christensen and M. Meldal, (2002), Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides, *J. Org. Chem.*, **67**, 3057–3064.
(5) H.C. Kolb and K.B. Sharpless, (2003), The growing impact of click chemistry on drug discovery, *Drug Discov. Today*, **8**, 1128–1137.
(6) P. Wu and V.V. Fokin, (2007), Catalytic azide–alkyne cycloaddition: reactivity and applications, *Aldrichimica Acta*, **40**, 7–17.
(7) J.F. Lutz, (2007), 1,3-Dipolar cycloadditions of azides and alkynes: a universal ligation tool in polymer and materials science, *Angew. Chem. Int. Edn*, **46**, 1018–1025.
(8) J.M. Baskin and C.R. Bertozzi, (2007), Bioorthogonal click chemistry: covalent labeling in living systems, *QSAR Comb. Sci.*, **26**, 1211–1219.
(9) W.P. Jencks, (1959), Studies on the mechanism of oxime and semicarbazone formation, *J. Am. Chem. Soc.*, **81**, 475–481.
(10) A. Dirksen, T.M. Hackeng and P.E. Dawson, (2006), Nucleophilic catalysis of oxime ligation, *Angew. Chem. Int. Edn*, **45**, 7581–7584.
(11) I.L. Cartwright, D.W. Hutchinson and V.W. Armstrong, (1976), The reaction between thiols and 8-azidoadenosine derivatives, *Nucl. Acids Res.*, **3**, 2331–2339.
(12) M.B. Hossain, D. Vanderhelm, R. Sanduja and M. Alam, (1985), Structure of 6-azidotetrazolo[5,1-a]phthalazine, $C_8H_4N_8$, isolated from the toxic dinoflaggelate Gymnodinium breve, *Acta Cryst.*, **C41**, 1199–1202.
(13) R.J. Griffin, (1994), The medicinal chemistry of the azido group, *Prog. Med. Chem.*, **31**, 121–232.
(14) S. Brase, C. Gil, K. Knepper and V. Zimmermann, (2005), Organic azides: an exploding diversity of a unique class of compounds, *Angew. Chem. Int. Edn*, **44**, 5188–5240.

(15) E. Saxon and C.R. Bertozzi, (2000), Cell surface engineering by a modified Staudinger reaction, *Science*, **287**, 2007–2010.
(16) E. Saxon, J.I. Armstrong and C.R. Bertozzi, (2000), A 'traceless' Staudinger ligation for the chemoselective synthesis of amide bonds, *Org. Lett.*, **2**, 2141–2143.
(17) B.L. Nilsson, L.L. Kiessling and R.T. Raines, (2000), Staudinger ligation: a peptide from a thioester and azide, *Org. Lett.*, **2**, 1939–1941.
(18) B.L. Nilsson, L.L. Kiessling and R.T. Raines, (2001), High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide, *Org. Lett.*, **3**, 9–12.
(19) G.A. Lemieux, C.L. De Graffenried and C.R. Bertozzi, (2003), A fluorogenic dye activated by the Staudinger ligation, *J. Am. Chem. Soc.*, **125**, 4708–4709.
(20) M.J. Hangauer and C.R. Bertozzi, (2008), A FRET-based fluorogenic phosphine for live-cell imaging with the Staudinger ligation, *Angew. Chem. Int. Edn*, **47**, 2394–2397.
(21) F.L. Lin, H.M. Hoyt, H. Van Halbeek, R.G. Bergman and C.R. Bertozzi, (2005), Mechanistic investigation of the Staudinger ligation, *J. Am. Chem. Soc.*, **127**, 2686–2695.
(22) A. Michael, (1893), Ueber die einwirkung von diazodenzolimid auf acetylendicarbonsauremethylester, *J. Prakt. Chem.*, **48**, 94–95.
(23) R. Huisgen, (1963), 1.3-Dipolare Cycloadditionen: Ruckschau Und Ausblick, *Angew. Chem. Int. Edn*, **75**, 604–637.
(24) F. Himo, T. Lovell, R. Hilgraf, V.V. Rostovtsev, L. Noodleman, K.B. Sharpless and V.V. Fokin, (2005), Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates, *J. Am. Chem. Soc.*, **127**, 210–216.
(25) G. Wittig and A. Krebs, (1961), Zur Existenz Niedergliedriger Cycloalkine 1, *Chem. Ber.*, **94**, 3260–3275.
(26) R.B. Turner, P. Goebel, B.J. Mallon and A.D. Jarrett, (1973), Heats of hydrogenation 9. Cyclic acetylenes and some miscellaneous olefins, *J. Am. Chem. Soc.*, **95**, 790–792.
(27) N.J. Agard, J.A. Prescher and C.R. Bertozzi, (2004), A strain-promoted [3 + 2] azide–alkyne cycloaddition for covalent modification of blomolecules in living systems, *J. Am. Chem. Soc.*, **126**, 15046–15047.
(28) N.J. Agard, J.M. Baskin, J.A. Prescher, A. Lo and C.R. Bertozzi, (2006), A comparative study of bioorthogonal reactions with azides, *ACS Chem. Biol.*, **1**, 644–648.
(29) J.M. Baskin, J.A. Prescher, S.T. Laughlin, N.J. Agard, P.V. Chang, I.A. Miller, A. Lo, J.A. Codelli and C.R. Bertozzi, (2007), Copper-free click chemistry for dynamic *in vivo* imaging, *Proc. Natl Acad. Sci. USA*, **104**, 16793–16797.
(30) J.A. Codelli, J.M. Baskin, N.J. Agard and C.R. Bertozzi, (2008), Second-generation difluorinated cyclooctynes for copper-free click chemistry, *J. Am. Chem. Soc.*, ASAP Article.
(31) E.M. Sletten and C.R. Bertozzi, (2008), A hydrophilic azacyclooctyne for Cu-free click chemistry, *Org. Lett.*, **10**, 3097–3099.
(32) S.S. van Berkel, A.T. Dirks, M.F. Debets, F.L. van Delft, J.J. Cornelissen, R.J. Nolte and F.P. Rutjes, (2007), Metal-free triazole formation as a tool for bioconjugation, *Chembiochem*, **8**, 1504–1508.
(33) X. Ning, J. Guo, M.A. Wolfert and G.J. Boons, (2008), Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast Huisgen cycloadditions, *Angew. Chem. Int. Edn*, **47**, 2253–2255.
(34) W. Song, Y. Wang, J. Qu, M.M. Madden and Q. Lin, (2008), A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins, *Angew. Chem. Int. Edn*, **47**, 2832–2835.
(35) W. Song, Y. Wang, J. Qu and Q. Lin, (2008), Selective functionalization of a genetically encoded alkene-containing protein via 'photoclick chemistry' in bacterial cells, *J. Am. Chem. Soc.*, **130**, 9654–9655.
(36) Y. Wang, W.J. Hu, W. Song, R.K. Lim and Q. Lin, (2008), Discovery of long-wavelength photoactivatable diaryltetrazoles for bioorthogonal 1,3-dipolar cycloaddition reactions, *Org. Lett.*
(37) Y.A. Lin, J.M. Chalker, N. Floyd, G.J. Bernardes and B.G. Davis, (2008), Allyl sulfides are privileged substrates in aqueous cross-metathesis: Application to site-selective protein modification, *J. Am. Chem. Soc.*, **130**, 9642–9643.

(38) M.L. Blackman, M. Royzen and J.M. Fox, (2008), Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels–Alder reactivity, *J. Am. Chem. Soc.*, **130**, 13518–13519.
(39) N.K. Devaraj, R. Weissleder and S.A. Hilderbrand, (2008), Tetrazine-based cycloadditions: application to pretargeted live cell imaging, *Bioconjugate Chem.*, **19**, 2297–2299.
(40) A.E. Speers and B.F. Cravatt, (2004), Chemical strategies for activity-based proteomics, *Chembiochem*, **5**, 41–47.
(41) M.J. Evans and B.F. Cravatt, (2006), Mechanism-based profiling of enzyme families, *Chem. Rev.*, **106**, 3279–3301.
(42) N. Jessani and B.F. Cravatt, (2004), The development and application of methods for activity-based protein profiling, *Curr. Opin. Chem. Biol.*, **8**, 54–59.
(43) A.M. Sadaghiani, S.H. Verhelst and M. Bogyo, (2007), Tagging and detection strategies for activity-based proteomics, *Curr. Opin. Chem. Biol.*, **11**, 20–28.
(44) D.J. Vocadlo and C.R. Bertozzi, (2004), A strategy for functional proteomic analysis of glycosidase activity from cell lysates, *Angew. Chem. Int. Edn*, **43**, 5338–5342.
(45) H.C. Hang, J. Loureiro, E. Spooner, A.W. van der Velden, Y.M. Kim, A.M. Pollington, R. Maehr, M.N. Starnbach and H.L. Ploegh, (2006), Mechanism-based probe for the analysis of cathepsin cysteine proteases in living cells, *ACS Chem. Biol.*, **1**, 713–723.
(46) B.N.G. Giepmans, S.R. Adams, M.H. Ellisman and R.Y. Tsien, (2006), The fluorescent toolbox for assessing protein location and function, *Science*, **312**, 217–224.
(47) I. Chen and A.Y. Ting, (2005), Site-specific labeling of proteins with small molecules in live cells, *Curr. Opin. Biotechnol.*, **16**, 35–40.
(48) K.M. Marks and G.P. Nolan, (2006), Chemical labeling strategies for cell biology, *Nat. Meth.*, **3**, 591–596.
(49) B.A. Griffin, S.R. Adams and R.Y. Tsien, (1998), Specific covalent labeling of recombinant protein molecules inside live cells, *Science*, **281**, 269–272.
(50) S.R. Adams, R.E. Campbell, L.A. Gross, B.R. Martin, G.K. Walkup, Y. Yao, J. Llopis and R.Y. Tsien, (2002), New biarsenical ligands and tetracysteine motifs for protein labeling *in vitro* and *in vivo*: synthesis and biological applications, *J. Am. Chem. Soc.*, **124**, 6063–6076.
(51) G. Gaietta, T.J. Deerinck, S.R. Adams, J. Bouwer, O. Tour, D.W. Laird, G.E. Sosinsky, R.Y. Tsien and M.H. Ellisman, (2002), Multicolor and electron microscopic imaging of connexin trafficking, *Science*, **296**, 503–507.
(52) T.L. Halo, J. Appelbaum, E.M. Hobert, D.M. Balkin and A. Schepartz, (2009), Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid, *J. Am. Chem. Soc.*, **131**, 438–439.
(53) A.J. Rodriguez, S.M. Shenoy, R.H. Singer and J. Condeelis, (2006), Visualization of mRNA translation in living cells, *J. Cell. Biol.*, **175**, 67–76.
(54) C. Hoffmann, G. Gaietta, M. Bunemann, S.R. Adams, S. Oberdorff-Maass, B. Behr, J.P. Vilardaga, R.Y. Tsien, M.H. Ellisman and M.J. Lohse, (2005), A FlAsH-based FRET approach to determine G protein-coupled receptor activation in living cells, *Nat. Meth.*, **2**, 171–176.
(55) M.J. Roberti, C.W. Bertoncini, R. Klement, E.A. Jares-Erijman and T.M. Jovin, (2007), Fluorescence imaging of amyloid formation in living cells by a functional, tetracysteine-tagged alpha-synuclein, *Nat. Meth.*, **4**, 345–351.
(56) N. Arhel, A. Genovesio, K.A. Kim, S. Miko, E. Perret, J.C. Olivo-Marin, S. Shorte and P. Charneau, (2006), Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes, *Nat. Meth.*, **3**, 817–824.
(57) F. Senf, J. Tommassen and M. Koster, (2008), Polar secretion of proteins via the Xcp type II secretion system in *Pseudomonas aeruginosa*, *Microbiology*, **154**, 3025–3032.
(58) S. Chaumont and B.S. Khakh, (2008), Patch-clamp coordinated spectroscopy shows P2X2 receptor permeability dynamics require cytosolic domain rearrangements but not Panx-1 channels, *Proc. Natl Acad. Sci. USA*, **105**, 12063–12068.
(59) K. Stroffekova, C. Proenza and K.G. Beam, (2001), The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins, *Pflugers Arch.*, **442**, 859–866.
(60) I. Chen, M. Howarth, W. Lin and A.Y. Ting, (2005), Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase, *Nat. Meth.*, **2**, 99–104.

(61) M. Fernandez-Suarez, H. Baruah, L. Martinez-Hernandez, K.T. Xie, J.M. Baskin, C.R. Bertozzi and A.Y. Ting, (2007), Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes, *Nat. Biotechnol.*, **25**, 1483–1487.
(62) H. Baruah, S. Puthenveetil, Y.A. Choi, S. Shah and A.Y. Ting, (2008), An engineered aryl azide ligase for site-specific mapping of protein–protein interactions through photo-cross-linking, *Angew. Chem. Int. Edn*.
(63) T. Dierks, B. Schmidt, L.V. Borissenko, J. Peng, A. Preusser, M. Mariappan and K. von Figura, (2003), Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme, *Cell*, **113**, 435–444.
(64) I.S. Carrico, B.L. Carlson and C.R. Bertozzi, (2007), Introducing genetically encoded aldehydes into proteins, *Nat. Chem. Biol.*, **3**, 321–322.
(65) P. Wu, W. Shui, B.L. Carlson, N. Hu, D. Rabuka, J. Lee and C.R. Bertozzi, (2009), Site-specific chemical modification of recombinant proteins produced in mammalian cells using the genetically encoded aldehyde tag, *Proc. Natl Acad. Sci. USA*, **106**, 3000–3005.
(66) J.W. Chin, T.A. Cropp, J.C. Anderson, M. Mukherji, Z.W. Zhang and P.G. Schultz, (2003), An expanded eukaryotic genetic code, *Science*, **301**, 964–967.
(67) J. Wang, J. Xie and P.G. Schultz, (2006), A genetically encoded fluorescent amino acid, *J. Am. Chem. Soc.*, **128**, 8738–8739.
(68) D. Summerer, S. Chen, N. Wu, A. Deiters, J.W. Chin and P.G. Schultz, (2006), A genetically encoded fluorescent amino acid, *Proc. Natl Acad. Sci. USA*, **103**, 9785–9789.
(69) Z.W. Zhang, L. Alfonta, F. Tian, B. Bursulaya, S. Uryu, D.S. King and P.G. Schultz, (2004), Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells, *Proc. Natl Acad. Sci. USA*, **101**, 8882–8887.
(70) J.W. Chin, S.W. Santoro, A.B. Martin, D.S. King, L. Wang and P.G. Schultz, (2002), Addition of *p*-azido-L-phenylalanine to the genetic code of Escherichia coli, *J. Am. Chem. Soc.*, **124**, 9026–9027.
(71) J.W. Chin, A.B. Martin, D.S. King, L. Wang and P.G. Schultz, (2002), Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*, *Proc. Natl Acad. Sci. USA*, **99**, 11020–11024.
(72) W. Liu, A. Brock, S. Chen, S. Chen and P.G. Schultz, (2007), Genetic incorporation of unnatural amino acids into proteins in mammalian cells, *Nat. Meth.*, **4**, 239–244.
(73) A. Deiters, T.A. Cropp, M. Mukherji, J.W. Chin, J.C. Anderson and P.G. Schultz, (2003), Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*, *J. Am. Chem. Soc.*, **125**, 11782–11783.
(74) A. Deiters and P.G. Schultz, (2005), *In vivo* incorporation of an alkyne into proteins in *Escherichia coli*, *Bioorg. Med. Chem. Lett.*, **15**, 1521–1524.
(75) K. Sakamoto, A. Hayashi, A. Sakamoto, D. Kiga, H. Nakayama, A. Soma, T. Kobayashi, M. Kitabatake, K. Takio, K. Saito, M. Shirouzu, I. Hirao and S. Yokoyama, (2002), Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells, *Nucl. Acids Res.*, **30**, 4692–4699.
(76) A.J. Link, M.L. Mock and D.A. Tirrell, (2003), Non-canonical amino acids in protein engineering, *Curr. Opin. Biotechnol.*, **14**, 603–609.
(77) K.L. Kiick, J.C. van Hest and D.A. Tirrell, (2000), Expanding the scope of protein biosynthesis by altering the methionyl-tRNA synthetase activity of a bacterial expression host, *Angew. Chem. Int. Edn*, **39**, 2148–2152.
(78) K.L. Kiick, E. Saxon, D.A. Tirrell and C.R. Bertozzi, (2002), Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proc. Natl Acad. Sci. USA*, **99**, 19–24.
(79) D. Datta, P. Wang, I.S. Carrico, S.L. Mayo and D.A. Tirrell, (2002), A designed phenylalanyl-tRNA synthetase variant allows efficient *in vivo* incorporation of aryl ketone functionality into proteins, *J. Am. Chem. Soc.*, **124**, 5652–5653.
(80) K. Kirshenbaum, I.S. Carrico and D.A. Tirrell, (2002), Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues, *Chembiochem*, **3**, 235–237.
(81) D.C. Dieterich, A.J. Link, J. Graumann, D.A. Tirrell and E.M. Schuman, (2006), Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal

noncanonical amino acid tagging (BONCAT), *Proc. Natl Acad. Sci. USA*, **103**, 9482–9487.
(82) K.E. Beatty, J.C. Liu, F. Xie, D.C. Dieterich, E.M. Schuman, Q. Wang and D.A. Tirrell, (2006), Fluorescence visualization of newly synthesized proteins in mammalian cells, *Angew. Chem. Int. Edn*, **45**, 7364–7367.
(83) A.J. Link, M.K.S. Vink, N.J. Agard, J.A. Prescher, C.R. Bertozzi and D.A. Tirrell, (2006), Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids, *Proc. Natl Acad. Sci. USA*, **103**, 10180–10185.
(84) R.S. Haltiwanger and J.B. Lowe, (2004), Role of glycosylation in development, *Annu. Rev. Biochem.*, **73**, 491–537.
(85) D.H. Dube and C.R. Bertozzi, (2005), Glycans in cancer and inflammation – potential for therapeutics and diagnostics, *Nat. Rev. Drug. Discov.*, **4**, 477–488.
(86) P.M. Rudd, T. Elliott, P. Cresswell, I.A. Wilson and R.A. Dwek, (2001), Glycosylation and the immune system, *Science*, **291**, 2370–2376.
(87) D.H. Dube and C.R. Bertozzi, (2003), Metabolic oligosaccharide engineering as a tool for glycobiology, *Curr. Opin. Chem. Biol.*, **7**, 616–625.
(88) L.K. Mahal, K.J. Yarema and C.R. Bertozzi, (1997), Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis, *Science*, **276**, 1125–1128.
(89) T.L. Hsu, S.R. Hanson, K. Kishikawa, S.K. Wang, M. Sawa and C.H. Wong, (2007), Alkynyl sugar analogs for the labeling and visualization of glycoconjugates in cells, *Proc. Natl Acad. Sci. USA*, **104**, 2614–2619.
(90) S.J. Luchansky, S. Goon and C.R. Bertozzi, (2004), Expanding the diversity of unnatural cell-surface sialic acids, *Chembiochem*, **5**, 371–374.
(91) S. Han, B.E. Collins, P. Bengtson and J.C. Paulson, (2005), Homomultimeric complexes of CD22 in B cells revealed by protein–glycan cross-linking, *Nat. Chem. Biol.*, **1**, 93–97.
(92) D. Rabuka, S.C. Hubbard, S.T. Laughlin, S.P. Argade and C.R. Bertozzi, (2006), A chemical reporter strategy to probe glycoprotein fucosylation, *J. Am. Chem. Soc.*, **128**, 12078–12079.
(93) M. Sawa, T.L. Hsu, T. Itoh, M. Sugiyama, S.R. Hanson, P.K. Vogt and C.H. Wong, (2006), Glycoproteomic probes for fluorescent imaging of fucosylated glycans *in vivo*, *Proc. Natl Acad. Sci. USA*, **103**, 12371–12376.
(94) H.C. Hang and C.R. Bertozzi, (2001), Ketone isosteres of 2-*N*-acetamidosugars as substrates for metabolic cell surface engineering, *J. Am. Chem. Soc.*, **123**, 1242–1243.
(95) H.C. Hang, C. Yu, D.L. Kato and C.R. Bertozzi, (2003), A metabolic labeling approach toward proteomic analysis of mucin-type *O*-linked glycosylation, *Proc. Natl Acad. Sci. USA*, **100**, 14846–14851.
(96) D.J. Vocadlo, H.C. Hang, E.J. Kim, J.A. Hanover and C.R. Bertozzi, (2003), A chemical approach for identifying *O*-GlcNAc-modified proteins in cells, *Proc. Natl Acad. Sci. USA*, **100**, 9116–9121.
(97) P.V. Chang, J.A. Prescher, M.J. Hangauer and C.R. Bertozzi, (2007), Imaging cell surface glycans with bioorthogonal chemical reporters, *J. Am. Chem. Soc.*, **129**, 8400–8401.
(98) J.A. Prescher, D.H. Dube and C.R. Bertozzi, (2004), Chemical remodelling of cell surfaces in living animals, *Nature*, **430**, 873–877.
(99) D.H. Dube, J.A. Prescher, C.N. Quang and C.R. Bertozzi, (2006), Probing mucin-type *O*-linked glycosylation in living animals, *Proc. Natl Acad. Sci. USA*, **103**, 4819–4824.
(100) S.T. Laughlin, J.M. Baskin, S.L. Amacher and C.R. Bertozzi, (2008), *In vivo* imaging of membrane-associated glycans in developing zebrafish, *Science*, **320**, 664–667.
(101) Y. Kho, S.C. Kim, C. Jiang, D. Barma, S.W. Kwon, J. Cheng, J. Jaunbergs, C. Weinbaum, F. Tamanoi, J. Falck and Y. Zhao, (2004), A tagging-via-substrate technology for detection and proteomics of farnesylated proteins, *Proc. Natl Acad. Sci. USA*, **101**, 12479–12484.
(102) H.C. Hang, E.J. Geutjes, G. Grotenbreg, A.M. Pollington, M.J. Bijlmakers and H.L. Ploegh, (2007), Chemical probes for the rapid detection of fatty-acylated proteins in mammalian cells, *J. Am. Chem. Soc.*, **129**, 2744–2745.
(103) D.D. Martin, G.L. Vilas, J.A. Prescher, G. Rajaiah, J.R. Falck, C.R. Bertozzi and L.G. Berthiaume, (2008), Rapid detection, discovery, and identification of post-translationally

myristoylated proteins during apoptosis using a bio-orthogonal azidomyristate analog, *FASEB J.*, **22**, 797–806.
(104) M.A. Kostiuk, M.M. Corvi, B.O. Keller, G. Plummer, J.A. Prescher, M.J. Hangauer, C.R. Bertozzi, G. Rajaiah, J.R. Falck and L.G. Berthiaume, (2008), Identification of palmitoylated mitochondrial proteins using a bio-orthogonal azido-palmitate analogue, *FASEB J.*, **22**, 721–732.
(105) J.L. Meier, A.C. Mercer, H. Rivera, Jr and M.D. Burkart, (2006), Synthesis and evaluation of bioorthogonal pantetheine analogues for *in vivo* protein modification, *J. Am. Chem. Soc.*, **128**, 12174–12184.
(106) L.R. Comstock and S.R. Rajski, (2005), Conversion of DNA methyltransferases into azidonucleosidyl transferases via synthetic cofactors, *Nucl. Acids Res.*, **33**, 1644–1652.
(107) A. Salic and T.J. Mitchison, (2008), A chemical method for fast and sensitive detection of DNA synthesis *in vivo*, *Proc. Natl Acad. Sci. USA*, **105**, 2415–2420.
(108) C.Y. Jao and A. Salic, (2008), Exploring RNA transcription and turnover *in vivo* by using click chemistry, *Proc. Natl Acad. Sci. USA*, **105**, 15779–15784.