Matter

CellPress
REVIEWS

Review

Site-Selective Protein Modification: From Functionalized Proteins to Functional Biomaterials

Jared A. Shadish[1] and Cole A. DeForest[1,2,3,4,*]

Chemists, biologists, and material scientists alike have long sought to control protein presentation, orientation, and activity within biomaterials to dictate reaction catalysis, biological signaling, and cell-fate specification. Such control is most typically achieved through the installation of reactive chemical handles onto proteins that govern static or dynamic biomacromolecule-material associations. Though convenient, stochastic functionalization strategies often yield ill-defined protein samples with severely diminished activity. In contrast, site-selective methodologies permit the controlled installation of material-interacting handles onto fragile proteins while preserving native structure and function. Here, we review methods that afford chemical, regioselective, and site-specific modification of proteins, emphasizing those that have found utility in the creation of functional biomaterials. We discuss cutting-edge strategies involving small-molecule-based labeling, genetic engineering, and chemoenzymatic reactions that provide precise control over the extent and location of protein modification. We assess the progress and look to the future in exploiting these functional proteins to create next-generational biomaterials for tissue engineering, therapeutic, and fundamental biological applications.

INTRODUCTION

Proteins are the main workhorses of biology with nearly infinite possibilities for sequence, structure, chemical reactivity, and function. Given their unique potential to regulate exquisitely virtually all biological functions, there is profound interest in using these powerful biomacromolecules in many biomaterials contexts—for example, as potent therapeutics, for three-dimensional (3D) cell culture, and in disease detection.[1,2] Fully harnessing these complicated biomolecules for materials applications has required constant effort to effectively regulate their orientation and spatiotemporal presentation from and within such biomaterials. This control has been gained, in part, through the installation of reactive functional groups onto proteins that govern their static or dynamic interactions with materials. However, modifying such chemically and structurally rich biopolymers without perturbing their function remains an open challenge.[3] Although many strategies exist to label proteins stochastically, these random modifications can adversely block enzymatic active sites and binding pockets or alter the protein's 3D structure, leading to a decrease in or complete obliteration of activity. Despite these challenges, early pioneering work utilizing stochastic installation of functional groups onto antibodies has revolutionized biology, enabling the foundational advances making up immunohistochemistry and enzyme-linked immunosorbent assays (ELISAs).[4,5] Inspired by these early successes and to expand applications to ever more fragile

Progress and Potential

Proteins are the most abundant biomolecule and play essential roles in all living processes. To better understand their role in regulating biological fates, as well as to promote specific cellular responses for tissue engineering and therapeutic applications, interest in making materials that incorporate functional proteins continues to grow. Although there has been significant progress to date, challenges have arisen in how best to interface proteins with and attach them to materials in a manner that preserves their activity. Spurred by recent innovations in chemical biology that enable precise introduction of reactive handles at perfectly defined protein surface locations, it is now conceivable to control how, when, and where bioactive proteins bind to and are released from materials. Properly developed applications of these emerging technologies will further revolutionize modern medicine, enabling new directions in disease detection, drug delivery, and tissue engineering.

species, techniques to refine modification specificity remain a topic of active development.

Regardless of the functionalization strategy, careful consideration must be paid to the reaction conditions for protein modification. Due to the fragility and massive size of most full-length proteins, the chemistries employed must occur under mild, aqueous conditions and proceed relatively rapidly on substrates at very-low-molar concentrations.[6] Owing to the diversity and arrangement of amino acid residues within each species, the protein targets display a diverse array of chemical functionalities with varying nucleophilicity, pK_a, and redox potential. Targeting a specific functional group from the many that may be available on the protein's surface, let alone a single moiety on a specific residue among many near-identical copies, is a tall order. To increase the specificity of these labeling reactions, three broad strategies have emerged: (1) those that utilize reactive small-molecule chemical reagents to modify endogenous proteins; (2) those that harness or hijack protein translational processes for direct labeling; and (3) those that utilize enzymes for co- or post-translational modification of proteins.

Small-molecule-based labeling strategies exploit differences in nucleophilicity, pK_a, and redox potential of a given side chain to gain chemoselectivity. Lysine and cysteine residues have historically been the most widely targeted amino acids, owing to their relative abundance and highly nucleophilic side chains.[7] Activated esters are commonly employed to label primary amines present on lysine side chains, while maleimides act as electrophiles for conjugation to thiol-bearing cysteine residues (Figure 1). While such labeling methods enable facile creation of functional bioconjugates on highly purified protein samples, difficulties in controlling their extent and location of reactions yield poorly defined samples, often with dramatically reduced activity. Controlling the microenvironment of a particular amino acid through motif insertion can subtly change the reactivity of a particular residue and increase reaction specificity.

Strategies that utilize protein translational processes for direct labeling have proved quite successful in further refining the selectivity of bioconjugation. Here, proteins are able to be modified during ribosomal elongation, allowing for functionalization prior to purification. Genetic code expansion utilizing evolved tRNA/synthetase pairs allows for the incorporation of non-natural amino acids containing functional side chains into a specific location within the primary sequence of a protein.[8] Although improvements continue to be made, genetic code expansion has been traditionally associated with non-trivial protein expression setups and reduced overall yields, limiting its application in a materials context. Residue-specific incorporation, whereby a non-canonical amino acid is substituted into the proteome in the place of a naturally occurring species using native tRNA/synthetase pairs, can boost these yields, but comes at the price of potentially incomplete incorporation and a somewhat limited scope of functional handles that can be installed.[9,10] Chimeric proteins where one of the fusion species preferentially binds to a substrate or can undergo self-labeling are readily obtained in high yield and with perfectly controlled modification, but can exhibit decreased activity or can be too bulky for some applications.

Labeling strategies that utilize enzymes for co- or post-translational modification of proteins have proved exceptionally successful in increasing selectivity and conversion. Most typically, short peptide sequences that can be recognized and acted upon by a variety of enzymes are appended to full-length proteins using conventional protein engineering strategies. Combining the ease of installation of recognition peptide motifs with the exquisite selectivity of enzymes has allowed for

[1]Department of Chemical Engineering, University of Washington, Seattle, WA 98195, USA

[2]Department of Bioengineering, University of Washington, Seattle, WA 98105, USA

[3]Institute for Stem Cell & Regenerative Medicine, University of Washington, Seattle, WA 98109, USA

[4]Molecular Engineering & Sciences Institute, University of Washington, Seattle, WA 98195, USA

*Correspondence: profcole@uw.edu

https://doi.org/10.1016/j.matt.2019.11.011

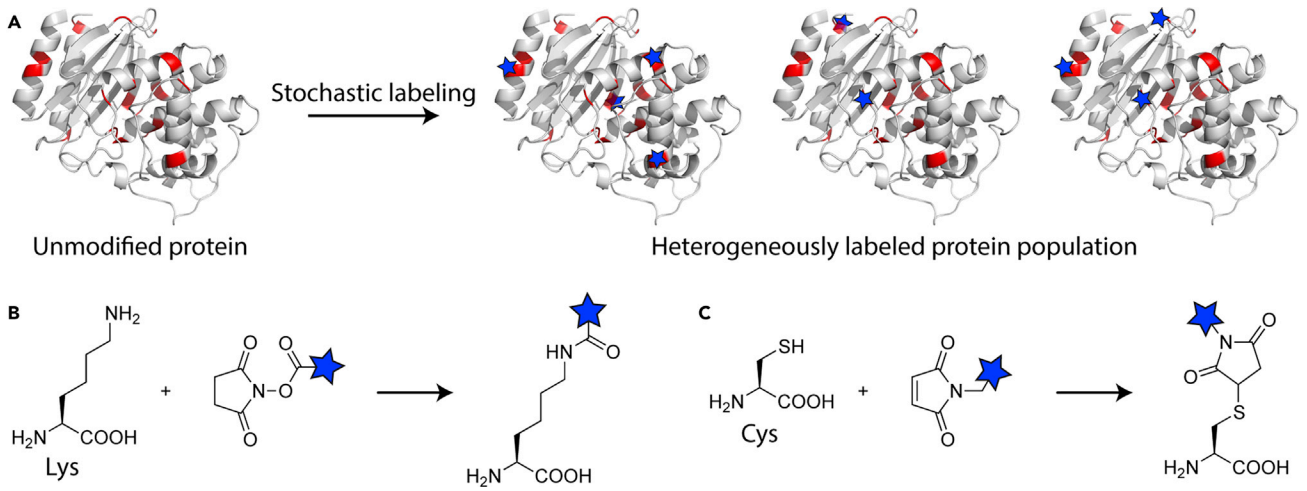

Figure 1. Stochastic Labeling of Endogenous Amino Acids

(A) Taking advantage of reactive functionalities present on endogenous amino acids (red), residue-specific modification (blue stars) can be achieved. Although these reactions are easily implemented and compatible with almost any protein, challenges exist in controlling where and to what extent each species is modified, resulting in a heterogeneously labeled population.

(B) Lysine residues are most commonly modified with activated esters (such as that of *N*-hydroxysuccinimide), whereby lysine's ε-amino group nucleophilically attacks the ester to form a stable amide linkage.

(C) Cystine residues are readily decorated through thiol-maleimide chemistry.

unparalleled control over the site of modification while retaining high conversion and yield. While these strategies are often limited to the termini as sites of potential modification and can be comparatively challenging to implement, progress is being made with newly evolved enzymes and peptide recognition sequences to address both of these concerns.

Ultimately, proteins need to be modified mildly, selectively, and with high conversion to retain their activity for use in biomaterials, especially if the technology is to translate into a clinical setting. Some applications to date have involved controlling the orientation of proteins on surfaces, creating crosslinks within biomaterials using proteins with novel functions, and affording spatiotemporal control over biochemical cue presentation within synthetic tissues. As each application features its own set of constraints, careful consideration for the employed protein labeling strategies must be given. While engineered proteins have gained traction in biocatalysis, metabolic engineering, optogenetics, and as biosensors,[11–14] here we review the many methods that afford site-specific modification of proteins, placing special emphasis on those that have found significant utility and/or hold exceptional promise toward creation of next-generation functional biomaterials.

CHEMOSELECTIVE PROTEIN MODIFICATION

Most historically prevalent examples of protein modification have relied on inherent differences in reactivity among the side chains to dictate conjugation. While these techniques have traditionally had relatively high chemoselectivity, they have typically struggled to maintain regioselectivity with high conversion. Careful engineering and rational design of new protein modification reagents has led to increased specificity, enabling endogenous "off-the-shelf" proteins to be labeled with no need for genetic engineering. Modification of the local microenvironment around desired sites of reaction can enhance specificity but comes at the price of requiring protein engineering. Here, we discuss small-molecule-based labeling strategies that target endogenous amines, thiols, aromatics, and the N and C termini

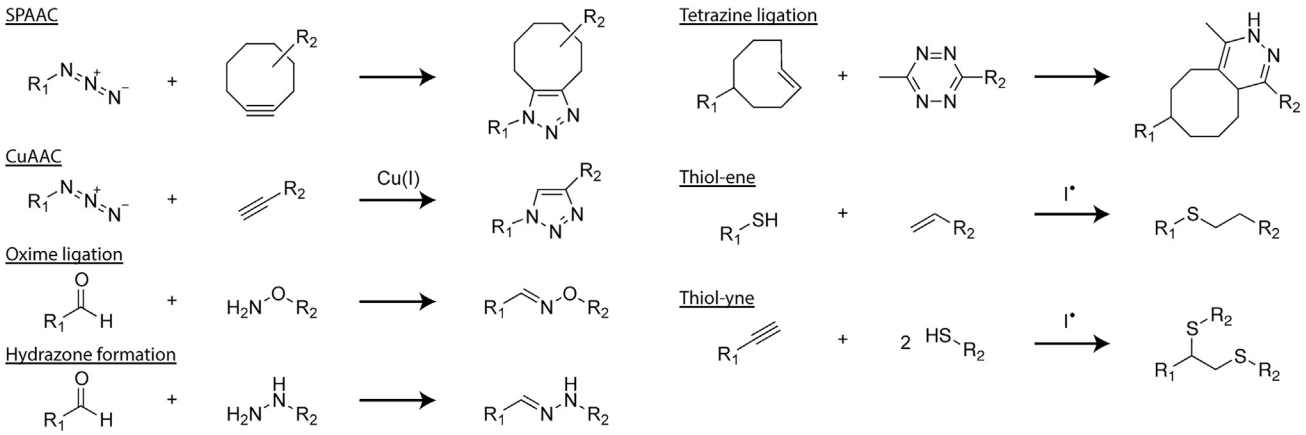

Figure 2. Common Reactions Used to Elaborate Modified Proteins

Reactive handles previously installed through another method (e.g., endogenous labeling, genetic code expansion) can be elaborated through a variety of specific reactions. Here, R_1 and R_2 can each refer to a protein, a biomaterial, or additional functional groups.

on proteins, as well as strategies to enhance labeling specificity using motif insertion. Many of these modifications install reactive handles that can be used for additional elaboration and to directly attach to materials (Figure 2).

Lysine Modification

Lysine has proved to be a popular target for modification due to the nucleophilicity of the ε-amine located on its side chain and its overall ubiquity (~6% of residues in the human proteome). However, this abundance and inherent reactivity also makes site-specific labeling challenging. Building on a long history for amide linkage formation in peptide synthesis,[15] activated esters (including those based on *N*-hydroxysuccinimide [NHS]) are the most widely used electrophiles for conjugation with primary amines on the protein surface. Despite the random and stochastic nature of this labeling strategy, the ease and simplicity of NHS-mediated protein modification has led to wide adoption in the biomaterials community. In one study, the Shoichet lab used NHS chemistry to functionalize epidermal growth factor (EGF) with biotin for immobilization in a streptavidin-patterned hyaluronic acid hydrogel, and found that immobilized EGF could dictate invasion of MDA-MB-231 and MDA-MB-468 breast cancer cells.[16] In another example, DeForest and Tirrell used NHS-based reactions to modify a variety of proteins with aldehydes and photoreactive *ortho*-nitrobenzyl esters used for photopatterned functionalization of poly(ethylene glycol) (PEG)-based hydrogels.[17] By reversibly patterning human mesenchymal stem cell-laden biomaterials with vitronectin, they demonstrated spatiotemporal control over osteogenic differentiation. NHS chemistry offers a quick off-the-shelf method for modifying proteins that has proved successful in many applications.

Despite these successes, stochastic labeling of primary amines often has a deleterious effect on protein function. As a good nucleophile with the ability to act as a base in the proper pH, lysine is often part of the catalytic triad of enzymes and is thus required for protein activity; β-lactamases in particular are known to utilize lysine in their active pocket.[18] Although the enhanced reactivity of lysines in the catalytic domain of enzymes is typically viewed as a drawback, creative strategies have taken advantage of this feature to create site-specifically defined antibody-drug conjugates; the lysine in the catalytic domain of a β-lactamase fused to a monoclonal antibody has been utilized to site-specifically tether a β-lactam-containing chemotherapeutic.[19]

Taking advantage of minute differences in the pK_a between lysine residues, computational approaches have also been exploited to identify electrophiles that would favor site-specific modification. Sulfonyl acrylate reagents were designed to exploit subtle differences in reactivity on native lysine residues to give site-specific incorporation on a variety of proteins, including installation of a fluorophore onto trastuzumab, without destroying its activity.[20] While an impressive demonstration of the power of computational design of small-molecule reagents, there remains significant room for improvement. Conversion occurred at the lysine with the lowest pK_a, which may be a critical or inaccessible residue on some proteins, and high excesses of sulfonyl acrylate reagents (1,000- to 10,000-fold excess) were required for quantitative conversion.

Multi-step reactions have also been used in an attempt to control the extent and location of lysine modification. Using a multi-component reaction using an aldehyde, an acetylene, and a Cu-ligand complex, propargylamine was appended site-specifically onto nine different proteins.[21] While the modified proteins retained their activity, this technique required long reaction times (up to 72 h) with relatively low conversion (generally 40%–60%). Lysine modifications remain one of the most common techniques for protein functionalization, but achieving high conversion while maintaining specificity continues to be a challenge.

Cysteine Modification

Cysteine residues are also a popular target for protein modification due to the high nucleophilicity of its side chain thiol, and the relative lack of free sulfhydryl groups on native proteins (most are tied up in disulfide linkages). Since many proteins do not contain native free thiols that can be targeted, and reduction of disulfide bonds can have deleterious effects on activity by disrupting protein structure, incorporation of a single cysteine residue via genetic engineering is commonly employed to give a functional handle for conjugation at a precise protein location. Traditionally, maleimides have been used for thiol alkylation and protein conjugation. Maleimide chemistry has long been exploited to immobilize antibody antigen-binding fragment (Fab) and horseradish peroxidase onto surfaces for ELISAs.[22] Though a powerful chemistry for *in vitro* applications, it was recently reported that the thioether bond can undergo thiol exchange or ring opening and decompose *in vivo*.[23] Despite this limitation, several maleimide conjugates have been approved by the Food and Drug Administration (e.g., brentuximab vedotin, trastuzumab emtansine, and certolizumab pegol). Michael-type chemistry, often employing acrylates or vinyl sulfones, has also proved useful in labeling free thiols. The Weber group used a thiolated phytochrome-based protein (Cph1) to create protein-polymer hydrogels with cyclic stiffening properties (Figure 3).[24] Cph1 containing a free sulfhydryl was conjugated to an eight-arm PEG through Michael addition to generate a protein-polymer hydrogel. Cph1 would then dimerize under 740-nm light but revert to its monomeric form with 660-nm exposure, leading to reversible material crosslinking.

Thiol-targeting reagents have also been developed that offer increased aqueous and metabolic stability, as well as crosslinkers. Oxetanes have found use for thiol modification, increasing overall protein stability in human plasma and in the presence of thiol-containing nucleophiles such as glutathione and β-mercaptoethanol.[25] Amine-to-thiol crosslinking reagents, such as succinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), have also proved popular. SMCC contains both maleimide and NHS-activated ester functionalities that permit conjugation to lysines on a protein of interest that does not contain free thiols through amine reactivity. The sulfhydryl groups of a second protein are then conjugated using the

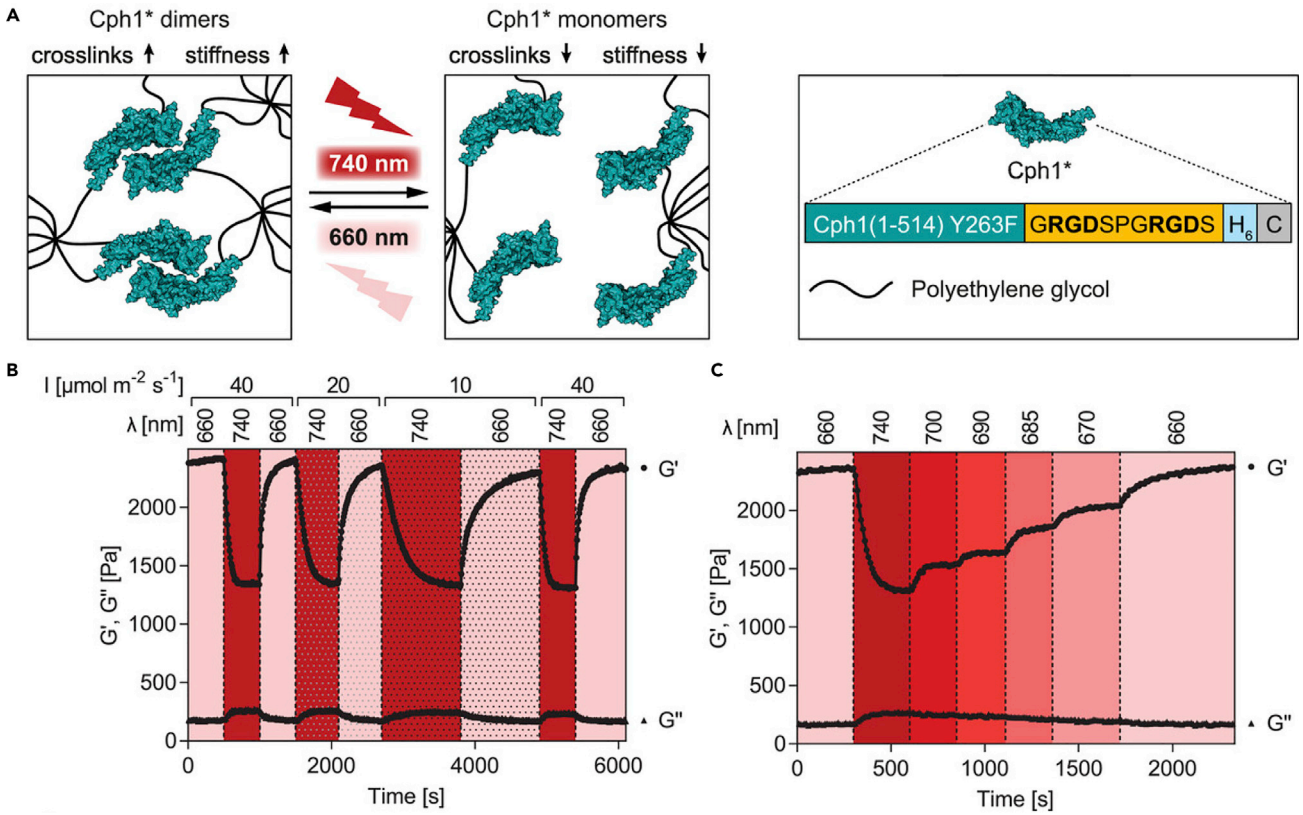

Figure 3. Reversible Gel Stiffness through Protein Photodimerization

(A) Cph1 protein dimerizes with 740-nm light, but reverts to its monomeric form in response to 660-nm light. Multi-arm polymers end-functionalized with Cph1 undergo reversible crosslinking with multi-colored light exposure, yielding biomaterials with tunable stiffness. Cph1 is linked to an eight-arm PEG through Michael addition involving its C-terminal cysteine residue.

(B) Absolute and dynamic changes to moduli can be controlled through cycled exposure to 660- and 740-nm light at different intensities (*I*). Here, *G′* corresponds to the storage moduli of the material, while *G″* is the loss moduli.

(C) Intermediate moduli can be accessed through controlled exposure to different wavelengths between 660 and 740 nm.

Image modified with permission from Hörner et al.[24] Copyright 2019, Wiley-VCH.

maleimide group to create protein-protein heterodimers. Sodium 4-((4-(cyanoethynyl)benzoyl)oxy)-2,3,5,6-tetrafluorobenzenesulfonate is a recently developed amine-to-thiol coupling reagent that was used to produce antibody-drug conjugates with improved plasma stability as compared with maleimide conjugation.[26]

Where free sulfhydryl groups are not present for reaction, disulfide bonds can be targeted using disulfide rebridging; here, bisalkylation is employed to rebridge native disulfide bonds to conserve native structure while including a linker containing exogenous functionality. Following this approach, the Brocchini group demonstrated that interferon-α2a (IFN-α2a) and antibody fragments could be site-specifically PEGylated while retaining the bridged structure required for sustained activity. Interestingly, double alkylation of the protein could also be achieved by increasing the ratio of PEG monosulfone:protein.[27] Further development of alkylation reagents has enabled the creation of compounds capable of reducing and rebridging disulfides with a single reagent,[28] allyl sulfones with improved water solubility for the functionalization of peptides and proteins,[29] and those capable of photomediated disulfide rebridging using UV light.[30] Recently, disulfide

rebridging has been used to conjugate toxic payloads to antibodies for targeted therapeutic delivery.[31] Disulfide rebridging is a powerful technique for modifying proteins with accessible disulfide bonds, especially in cases where maintaining the disulfide linkage is critical to function.

Thiol-ene click chemistry is another class of thiol-based reactions that is amenable to photomediated conjugation.[32] This radical-mediated conjugation reaction leads to anti-Markovnikov addition of a thiol to an alkene with high selectivity and yield, enabling the creation of protein functionalized biomaterials. Photochemical thiol-ene coupling most typically requires the addition of a photoinitiator to generate radicals for radical-mediated addition. A recently developed allyl sulfide chain-transfer agent enables the reversible immobilization of proteins into hydrogels using UV light; thiolated transferrin and ovalbumin were photoimmobilized in the presence of lithium phenyl-2,4,6-trimethylbenzoylphosphinate photoinitiator and released with 3D control using multi-photon lithography.[33] While UV-mediated thiol-ene conjugations proceed rapidly, concerns persist over the cytocompatibility and selectivity of the photoinitiators and radical-mediated reactions if they are to be performed in the presence of living cells.[34]

Aromatic Residue Modification

The relatively low abundance of aromatic residues on solvent-accessible regions of protein surfaces has made surface tyrosine, tryptophan, and phenylalanine residues an attractive target for site-specific modification where targeting nucleophilic side chains is intractable. Despite this interest, maintaining selectivity for one particular aromatic residue over another is made difficult by nearly identical reactivity and chemical environments between the aromatic side chains.

Sodium trifluoromethanesulfinate has been used to selectively install CF_3 groups onto tryptophan residues in full-length proteins for use in ^{19}F nuclear magnetic resonance studies.[35] This technique enables the curious ability to create "Teflon proteins," which are highly hydrophobic (CF_3 is twice as hydrophobic as CH_3), potentially increasing the bioavailability while reducing metabolic detoxification of pharmaceutically active proteins. A photocatalytic method for modification of the β position of tryptophan has also been developed.[36] An iridium photocatalyst, which is active under blue light, was used to incorporate a methyl acrylate onto the tryptophan residue on model peptides with good selectivity over other aromatic amino acids. While relatively unexplored, the unique physiochemical effects of modifying aromatic residues enable novel functionalities that are otherwise inaccessible when modifying nucleophilic residues.

N- and C-Terminal Modification

Much effort has been expended to find reagents suitable for modifying the termini of protein substrates.[37] As every linear polypeptide chain contains a unique N and C terminus, and these ends are often solvent accessible, labeling at these locations affords site-specific modification without the need for extensive protein engineering (Figure 4A). Typically, slight differences in reactivity, pK_a, and redox potential between the N-terminal amine and those on lysine side chains, as well as the C-terminal carboxylic acid with those on aspartic and glutamic acid residues, can be used to bias reaction to the termini. Historically, quantitative derivatization of the termini has been difficult to achieve without leading to off-target modification of the side chains. However, the allure of site-specific modification on native proteins is strong, and many chemistries continue to be developed with great promise in achieving chemo- and regioselectivity.

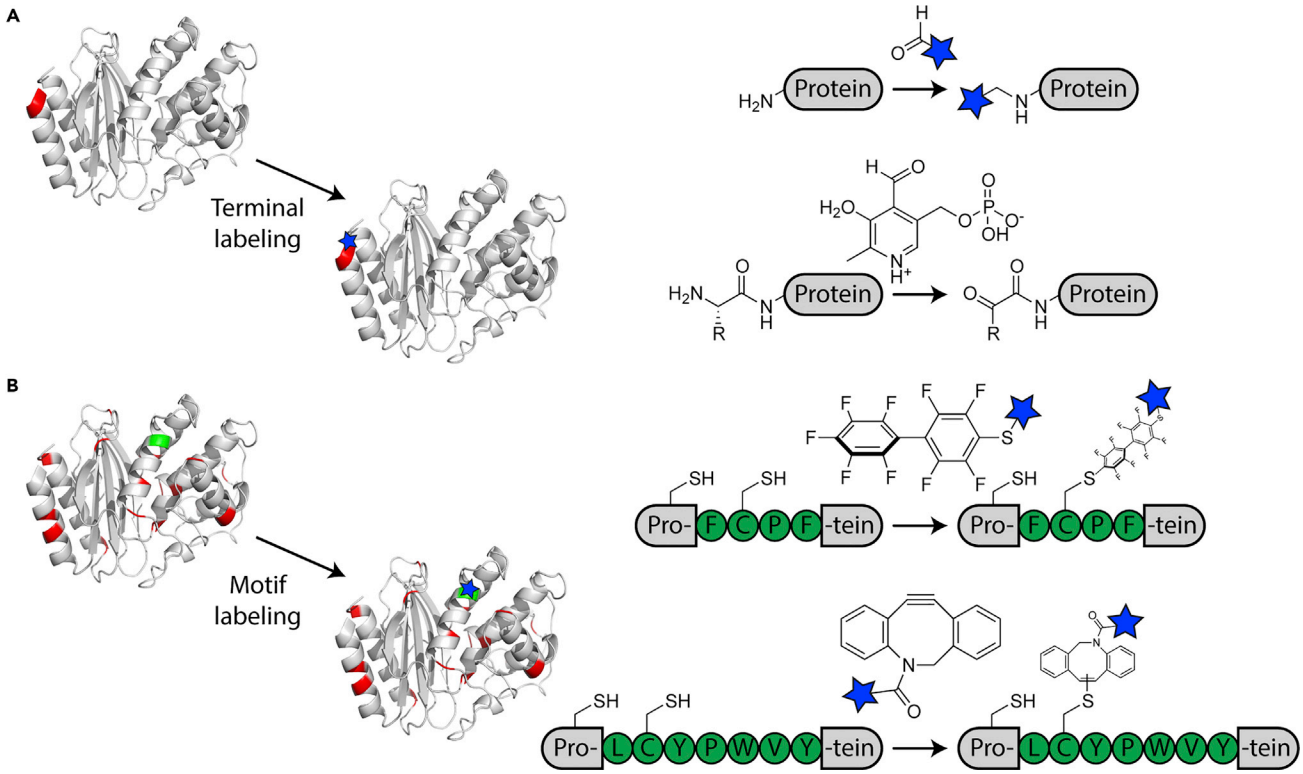

Figure 4. Site-Specific Labeling of Proteins with Reactive Small Molecules

(A) Chemical-based strategies can be used to selectively modify the protein termini. In one example, the primary amine on the N terminus can be derivatized through reductive amination involving an aldehyde-containing reagent (top). Alternatively, a pyridoxal 5′-phosphate (PLP)-mediated transamination reaction oxidizes the N-terminal amine to a ketone or an aldehyde (bottom).

(B) Taking advantage of reactive motifs, small-molecule chemical-based labeling strategies can be confined to specific locations on the protein. A single cysteine moiety can be selectively targeted using a π-clamp-mediated strategy (top) or a DBCO tag (bottom).

Several small-molecule approaches to selectively modify the N terminus have emerged. *O*-Aminophenols react readily with the N terminus of proteins using a potassium ferricyanide catalyst, favoring N-terminal proline residues as a substrate.[38] However, the conversion is relatively low (20%–60%), and free cysteines are also modified by the reaction. Aldehyde-containing reagents such as 2-pyridinecarbaldehyde (2-PCA) have proved popular for selective reductive alkylation of the N terminus.[39] One advantage of this approach is that reductive alkylation preserves the amine on the N terminus, theoretically maintaining native charge distribution on the protein. Antibodies have been site-selectively modified with an azide-functionalized PCA for further click-chemistry conjugation.[40] However, achieving complete functionalization of both chains of the antibody proved difficult, even at a very high excess of the coupling reagent. Applied to biomaterials, polyacrylamide (PAAm) hydrogels were functionalized with 2-PCA and used to immobilize extracellular matrix (ECM) proteins at their N termini.[41] These ECM-functionalized substrates promoted cell adhesion and spreading, and supported the creation of large collagen fibers, a phenotype associated with native collagen that could not be achieved using NHS chemistry for lysine modification.

Another common chemistry is for N-terminal transamination through the use of pyridoxal 5′-phosphate (PLP). Depending on the N-terminal amino acid composition, reaction with PLP oxidizes the N terminus to a ketone or aldehyde that can be

subsequently used for oxime/hydrazone ligation. Originally used to modify peptides, PLP gave 65% conversion of the N terminus into a ketone in 2 h at 37°C.[42] Where modification of native N termini on full-length proteins proves challenging, as has been the case for monoclonal antibodies, site-directed mutagenesis can be used to increase the reactivity of the terminus and achieve high levels of conversion.[43] Streptavidin was immobilized through an N-terminal oxime ligation onto micropatterned aminooxy-containing surfaces.[44] Particularly when working with enzymes, selective control over the location of protein immobilization and their surface orientation is important for many industry applications including biosensing and reactor design.

N-terminal cysteine residues can also provide a unique handle for subsequent modification. Native chemical ligation utilizes the free thiol of the cysteine, which is reacted with a C-terminal thioester. The bond then undergoes an S-N acyl shift to generate an amide linkage. Although a promising way to make protein conjugates containing only native functionality, these approaches face some challenges in implementation: engineering proteins to have an N-terminal cysteine residue is non-trivial as the start-codon-associated methionine residue must be enzymatically excised, and C-terminal thioesters for protein modification can be difficult to synthesize using solid-phase methodologies.[45] Expressed protein ligation, using recombinantly expressed polypeptides that self-cleave to contain the required functionalities, has also seen some development. In one example, expressed protein ligation was used to create site-specifically SUMOylated or methylated histones to investigate the impact of post-translational modifications on histone regulation.[46]

N-terminal cysteines can also react with aldehydes to produce thiazolidines, although developed reactions tend to be inefficient. *ortho*-Boronic acid-substituted benzaldehydes have been used to give site-specific conjugation of N-terminal cysteines with a unique thiazolidino boronate structure.[47] However, this bond was found to degrade under mildly acidic conditions and yielded only modest conversions (~50%).

While not as thoroughly explored as the N-terminal modifications, selective C-terminal modification has also been performed. Mild differences in the oxidation potential of the C terminus versus side chains allowed for decarboxylative alkylation to be performed exclusively at the C terminus.[48] Using a visible-light photoredox catalyst, insulin was modified to add a C-terminal alkyne with 40%–50% conversion. Investigation into post-translational modifications in the proteome reveal proteins with C-terminal methylation and α-amidation, suggesting the possibility of future chemistries to be developed to target this functionality.[49]

Despite these advances, many proteins do not have a solvent-accessible N or C terminus, have termini that are critical to their activity, or have flanking sequences with chemical microenvironments that make modification challenging. In addition, careful control over reaction conditions is necessary for termini labeling, making conjugation in complicated biological media or in the presence of cells a challenge. Therefore, additional strategies for site-selective protein modification remain of great interest.

Motif Insertion

Installation of short peptide domains that impart some specific structural or chemical modality—often referred to as motifs—onto proteins can further enhance reaction selectivity during protein modification (Figure 4B). By carefully controlling the

microenvironment of specific residues, these domains increase the efficiency and specificity of reactions.[50] When used in combination with small-molecule chemistries for protein modification, an impressive degree of control over protein functionalization can be achieved. Although computational modeling of microenvironments and further reagent development are likely to yield continued improvements, these approaches still tend to be fairly limited in scope with non-quantitative conversions.

Perhaps the best-known example of a peptide epitope leading to enhanced material binding is the "His tag." Polyhistidine sequences (usually either 6xHis or 10xHis tags) encoded within proteins allow for non-covalent binding with metal ions (e.g., Ni^{2+}, Co^{2+}, Cu^{2+}), a feature that is regularly exploited for protein purification in immobilized metal-affinity chromatography.[51] It has been found that His-tagged proteins may undergo N-terminal acylation with D-gluconic acid δ-lactone. The GH_6 tag (GHHHHHH) has been developed to enable N-terminal acylation of proteins.[52] Gluconolactone and 4-methoxyphenyl esters were used as acylating agents to PEGylate the N terminus of multiple full-length proteins. This modification was found to be reversible after extended incubation in phosphate-buffered saline.

Phage display typically has been used to determine epitopes that bind specific substrates. Fluorettes form one such short peptide sequence that has been found to bind common fluorophores (e.g., rhodamine red, Oregon green 514, fluorescein). Maleimide-functionalized Texas red dye was found to selectively conjugate to the fluorette domain of an engineered protein, even when other free thiols existed in the sequence.[53]

When combined with chemoselective modification techniques, motif insertion has yielded impressive regioselectivity. One peptide sequence identified through library screenings is the dibenzocyclooctyne (DBCO) tag (LCYPWVY) which, when reacted via thiol-yne conjugation, forms a stable covalent bond between DBCO-containing reagents and the free sulfhydryl group of the encoded cysteine.[54] This engineered microenvironment was found to lead to a 220-fold increase in reaction rate as compared with a cysteine-containing peptide control lacking the tag. Trastuzumab and green fluorescent protein (GFP) with terminal DBCO tags were modified site-selectively with 80%–90% conversion.

Increased reactivity of specific cysteine residues has also been engineered through the addition of a π clamp (FCPF).[55] Reactions with perfluoroaryl compounds and cysteines are typically very slow, except in the presence of a catalytic enzyme. Exploring sequences that provide altered microenvironments around specific cysteine residues that could enhance this reaction, the FCPF π clamp was found to enable site-specific conjugation in the presence of endogenous thiols. The strategy was used to conjugate affinity tags, fluorophores, alkynes, and linear PEG onto protein substrates at either the N or C terminus, or in an internal loop. This technique was found to lead to >95% conversion of the target thiol in 30 min.

Creating sites for efficient disulfide rebridging and subsequent reaction has also been explored through the addition of a Dis tag (CISTCC), which contains three cysteines with engineered spacing to serve as potential sites of modification.[56] The Dis tag allows for stepwise conjugation of maleimide-containing compounds for conjugation of the free thiol followed by the addition of allyl sulfones for disulfide rebridging. Yellow fluorescent protein was engineered to include the Dis tag and subsequently dual functionalized with two different fluorophores. Interleukin-

2 (IL-2), which contains a native disulfide, was also functionalized through rebridging and shown to have no loss in bioactivity; this is particularly noteworthy as reduction of the native disulfide bond of IL-2 leads to a 90% loss in activity.

Histidine bis-alkylation has been used to increase the selectivity of histidine modification. Histidine is a relatively rare residue, and by constraining the conjugation chemistry to require two nearby histidine residues the reaction specificity is increased. The HGH tag was incorporated into IFN-α2a away from the active site and used for site-specific PEGylation.[57] Conversion was found to be modest (~20%–40%), although the activity of modified IFN-α2a (74% compared with wild type) was the highest reported for any PEGylated IFN.

IN-TRANSLATION PROTEIN MODIFICATION

Although small-molecule-based approaches are convenient for modifying endogenous proteins, strategies that exploit protein translational processes for direct labeling offer enhanced labeling specificity during protein synthesis. Here, we highlight methods for residue-specific incorporation of non-canonical amino acids, site-specific labeling with non-natural amino acids through genetic code expansion, and fusion proteins including those that undergo self-labeling.

Metabolic Labeling

While small molecules have proved successful at site-specific protein modification and biomaterial functionalization, they require the exogenous treatment of the protein of interest with small-molecule reagents. There is growing interest in employing existing protein translational tools for incorporation of non-natural functionality during polypeptide elongation to enable co-translational protein modification (Figure 5). An early example of this technique is the metabolic incorporation of azidohomoalanine (Aha), an azide-containing methionine analog that is recognized by the native tRNA synthetase and substituted into the growing polypeptide chain stochastically during protein expression.[58] Similarly, an alkyne-containing methionine analog, homopropargylglycine (Hpg), may also be incorporated in place of methionine.[59] Both of these tags proved useful for subsequent purification and fluorophore conjugation onto newly synthesized proteins using click chemistry, as well as for visualizing newly synthesized proteins intracellularly within cell-laden biomaterials.[60,61] Although non-canonical amino acids containing azide, alkyne, and alkene functionalities can be metabolically incorporated into proteins using the natural translational machinery, directed evolution of endogenous aminoacyl tRNA synthetases has further enabled incorporation of an expanded collection of non-canonical amino acids.[62]

To increase the specificity of metabolic labeling, genetic engineering may be performed to enrich a given domain with the residue to be incorporated. In one example, GFP was engineered to have a C-terminal -HMHMHMHM tag.[63] Protein expression using a methionine-autotrophic bacterial strain in media lacking methionine but supplemented with Aha resulted in proteins whose methionine residues had been replaced with azide-containing Aha. These dually tagged proteins could be purified via immobilized metal-affinity chromatography using the histidine residues and then covalently immobilized onto an alkyne-modified agarose resin through copper-catalyzed azide-alkyne cycloaddition (CuAAC) chemistry.

Genetic Code Expansion

Genetic code expansion may be used to localize the site of modification to a single residue anywhere within the polypeptide backbone (Figure 6). Typically, the amber

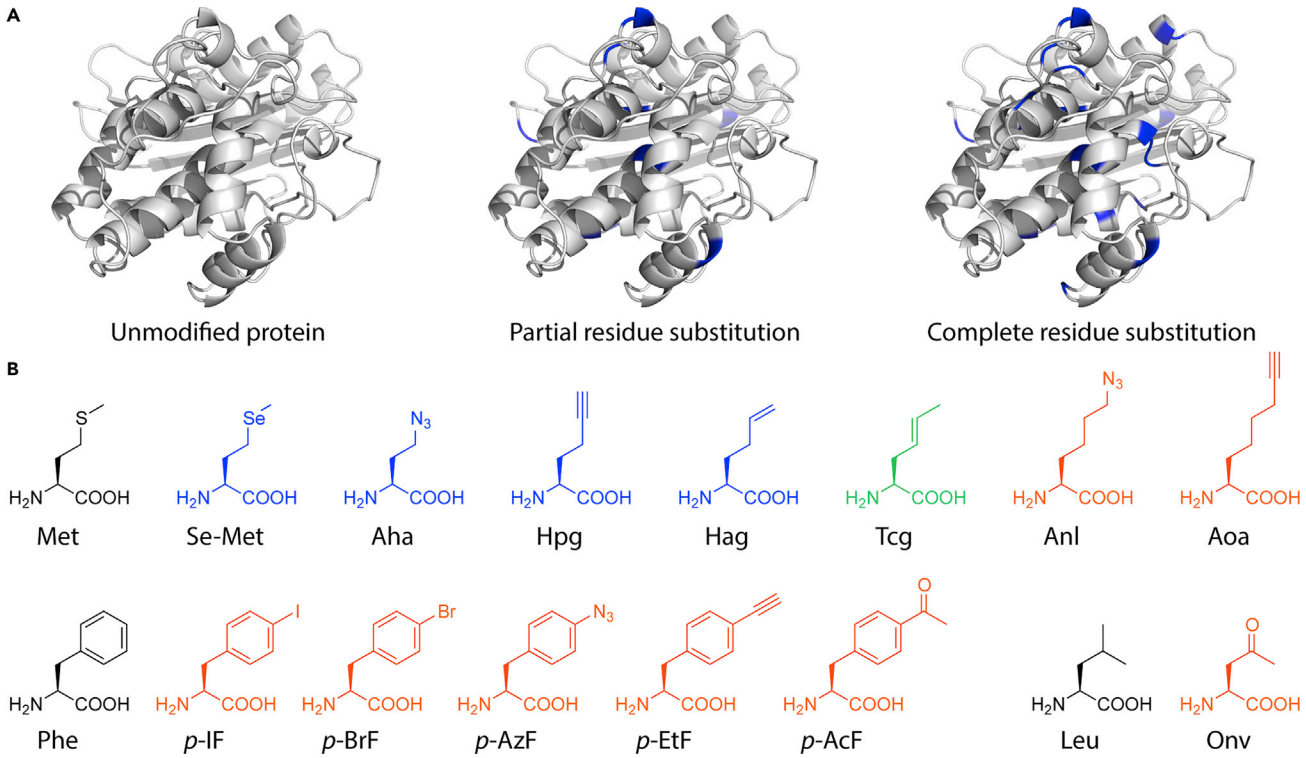

Figure 5. Residue-Specific Modification of Proteins through Metabolic Incorporation

(A) Metabolic incorporation of non-canonical amino acids in the place of their natural analog afford residue-specific substitution. By varying the relative concentrations of the canonical:non-canonical species present during expression, the extent of protein labeling is altered.

(B) Structures of previously reported non-canonical amino acids. Those shown in blue are substrates for the endogenous translational machinery encoding the black canonical residue on its left; the analog shown in green requires overexpression of wild-type tRNA synthetase; those in orange require expression of a mutant tRNA synthetase. Image inspired by Ngo and Tirrell.[62]

stop codon (UAG) is overridden to instead incorporate non-natural functionality using mutated tRNA/tRNA synthetase pairs. The Schultz group pioneered efforts in developing orthogonal tRNA/tRNA synthetase pairs specific to the amber stop codon that lacked cross-recognition by endogenous synthetases, typically utilizing synthetases from archaea and other distantly related organisms.[64] These strategies have resulted in incorporation of a wide variety of functional amino acids, including those that contain native post-translational modifications, clickable handles, and photoreactive groups.

Post-translational modifications that govern protein-protein interactions are highly sensitive to the incorporation site. To investigate this, several genetic code expansion techniques have been targeted toward native amino acids featuring known post-translational modifications. As an example, an ε-*N*-2-hydroxyisobutyryl-lysine (HibK) was site-specifically incorporated into histones to probe the function of 2-hydroxyisobutyryl-modified lysines.[65] In addition, the effects of protein phosphorylation have been investigated through the site-specific incorporation of a genetically encoded phosphotyrosine residue.[66] Protein methylation has been investigated through the incorporation of a dimethylated lysine residue that was site-specifically incorporated into histones and p53 in order to probe the function of epigenetic regulators such as histone demethylase LSD1 and histone acetyltransferase Tip60.[67]

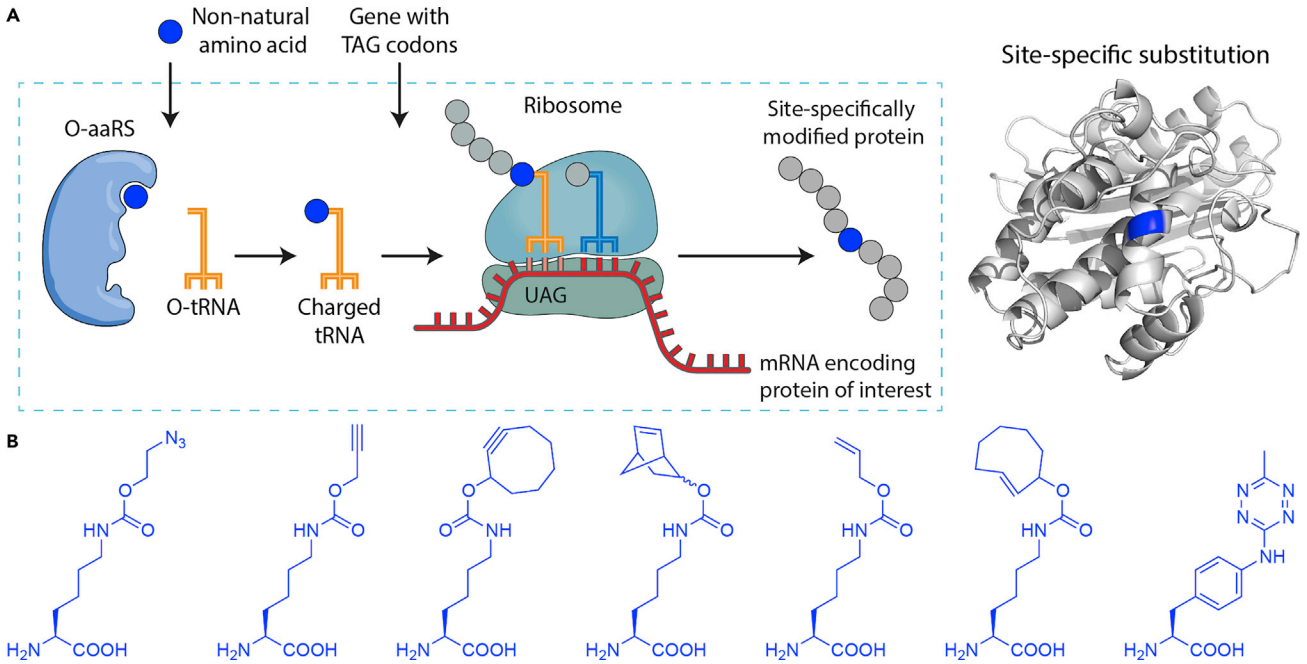

Figure 6. Site-Specific Protein Modification through Genetic Code Expansion

(A) Non-natural amino acids are charged by an orthogonal aminoacyl-tRNA synthetase (O-aaRS) onto an orthogonal tRNA (O-tRNA). The aminoacylated tRNA decodes the UAG codon present within modified protein mRNA, resulting in the site-specific incorporation of the non-natural amino acid.

(B) Structures of several reported non-natural amino acids with reactive side chains that have been incorporated into proteins using genetic code expansion.

Image in (A) inspired by Chin.[8] Copyright (2017), Springer Nature.

Beyond natural post-translational modifications, there is also significant interest in introducing novel functionalities into proteins using genetic engineering. Efforts in genetic code expansion led by the Chin group have been used to introduce amino acids containing azides for CuAAC click chemistry,[68] strained alkynes for SPAAC/SPANC,[69] and norbornenes and *trans*-cyclooctenes for inverse Diels-Alder reactions with tetrazines.[70,71] Recent efforts from the Kloxin lab have immobilized azide-tagged proteins through enzyme-degradable peptide linkages into materials (Figure 7).[72]

Consistent with interest in controlling material properties with light, photosensitive amino acids have proved popular to provide spatiotemporal control over protein function. By protecting a side chain critical for protein activity with a removable photocage, protein activity can be turned on upon exposure to light. In one example, a photocaged lysine was incorporated into live zebrafish embryos to enable spatiotemporal control over the activity of luciferase and various kinases.[73] In another, a photocaged cysteine was used to control tobacco etch virus protease activity in mammalian cells.[74] Moreover, several non-natural amino acids that are able to undergo photomediated ligation have been developed. Genetically encoded photolysines containing a diazirine moiety for UV-mediated crosslinking have been reported.[75] Furthermore, a dual-functional phenylazide/phenylamine side chain (AmAzZLys) was reported with a photoreactive azide for UV-induced crosslinking and a phenylamine for bioconjugation, which was used to site-selectively modify an antibody with a fluorophore and subsequently crosslink it to an antigen.[76] Such techniques are currently underutilized but hold great promise in the biomaterials space.

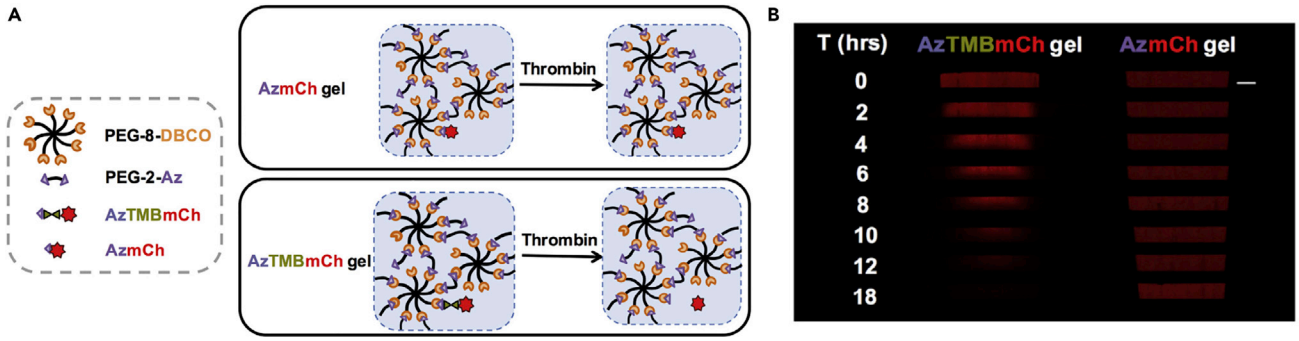

Figure 7. Enzyme-Mediated Release of Site-Specifically Modified Proteins from Gel Biomaterials

(A) Hydrogels formed through SPAAC between multifunctional PEG precursors are uniformly decorated with site-specifically modified full-length proteins. Inclusion of a thrombin-cleavable motif (TMB) between the gel and the protein permits enzyme-triggered release. Here, Az refers to a reactive azide installed onto the N terminus of mCherry (mCh) fluorescent protein using genetic code expansion.

(B) mCh is enzymatically released from gels functionalized with AzTMBmCh, but not from those lacking the thrombin cut-site.

Image modified with permission from Guo et al.[72] Copyright (2017), Elsevier.

While genetic code expansion has enabled many novel functionalities to be incorporated site-specifically into proteins, it remains limited in the substrates that can realistically be incorporated into the tRNA synthetase while maintaining orthogonality, and in maintaining high protein expression yields through multiple non-canonical amino acid insertions. However, the advantage of co-translational protein modification should not be overlooked, enabling protein labeling intracellularly during protein synthesis.

Fusion Proteins and Direct Ligand Immobilization (Self-Labeling)

Proteins have evolved to have high binding affinity selectively for their binding partners in the extremely crowded and complex biological environment within living cells. The ability of proteins to form strong non-covalent interactions or even covalent bonds with their substrates allows for protein modification at the active site of the enzyme. When expressed as a fusion protein with another species, the site of substrate conjugation is localized away from the protein of interest. Protein fusions with antibodies against particular antigens[77] and maltose-binding protein[78] represent full-length protein species that have been exploited by the materials community, primarily in chromatographic protein purification, while short peptide tags (e.g., 6xHis,[79] solid-binding peptides,[80] coiled coils[81]) have also been utilized to control protein/biomaterial formation and interactions. Although such strategies are relatively straightforward to implement, bonding is non-covalent and reversible, which can place some limits on application. As such, fusion partners that form stable covalent adducts represent an area of considerable interest.

Protein fusions have also been created, whereby one species undergoes a self-labeling reaction to install a functional handle for subsequent reaction with materials (Figure 8). Of these direct ligand immobilization tools, the SNAP tag and HaloTag are the most common. The SNAP tag is a modified form of O^6-alkylguanine DNA alkyltransferase that enables conjugation of O^6-benzylguanine analogs to the SNAP domain.[82] This technique has been used to functionalize proteins with ferrocene for immobilization onto cyclodextrin-micropatterned surfaces and vesicles.[83] The HaloTag is an engineered haloalkane dehalogenase that irreversibly binds to chloroalkane substrates. Fusions with the HaloTag have been used for conjugating full-length proteins to small-molecule fluorophores and agarose resins.[84] Cutinase is another self-labeling species, which forms a covalent adduct with phosphonate

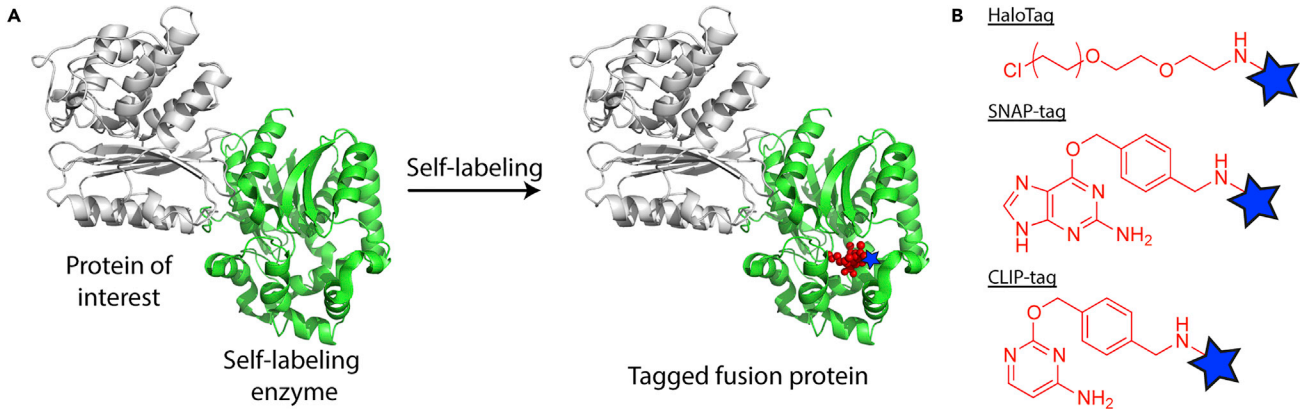

Figure 8. Site-Specific Protein Modification though Enzymatic Self-Labeling

(A) Self-labeling proteins used as fusion partners can afford site-specific protein labeling with functionalized small molecules.

(B) The HaloTag (a modified haloalkane dehalogenase) covalently attaches to haloalkane substrates; the SNAP tag (an engineered *O*-6-methylguanine-DNA methyltransferase) couples benzylguanine derivatives; the CLIP tag (an evolved DNA methyltransferase) ligates benzylcytosine-containing species.

ligands. Antibody-cutinase fusions have been used to create immobilized antibody arrays with controlled orientation.[85]

Owing to its full genetic encodability and rapid reaction, the SpyTag-SpyCatcher system is another self-ligating system that has significant traction in the biomaterials space. Here, the Howarth group has split fibronectin-binding protein FbaB into two fragments capable of forming a covalent isopeptide bond between critical lysine and aspartic acid residues unique to each fragment. Dubbing these fragments as SpyTag and SpyCatcher, isopeptide bond formation was found to form rapidly and near quantitatively in diverse reaction conditions.[86] Protein hydrogels have been formed by creating protein fusions with multiple SpyTag or SpyCatcher domains. In one example, Elastin-like proteins (ELPs) are fused to either three SpyTag domains or two SpyCatcher domains, which form covalent gels within 5 min of mixing.[87] Due to the rapid kinetics, specificity, and mild conditions of the reaction, live cells can be encapsulated within the network with high viability, and proteins such as SpyTag-functionalized mCherry and leukemia inhibitory factor were immobilized into the network during gelation. Similar protein hydrogels based on globular domains (GB1 and FNIII) as fusions with SpyTag and SpyCatcher were also formed and shown to be able to encapsulate viable cells.[88] The Clark group created PEG hydrogels with SpyTag crosslinks to tether SpyCatcher-labeled species, ultimately enabling proteins to be immobilized within gels at very high densities.[89] Although the SpyCatcher/Tag system permits site-specific conjugation, fusions with the bulky enzymatic groups may have deleterious effects on protein function.[90] The FbaB domains have been further split to create a SpyLigase, which acts as an enzyme that covalently conjugates SpyTag (AHIVMVDAYKPTK) with the KTag (ATHIKFSKRD).[91] This technique harnesses the benefits of protein specificity without necessitating the addition of the bulky SpyCatcher group, an advantage that other groups have sought to exploit through engineering other enzymes for protein modification. The SpyCatcher/Tag chemistry has been further modified to permit purification of modified proteins.[92]

CHEMOENZYMATIC PROTEIN MODIFICATION

Enzymes are biocatalysts that are able to perform reactions with unparalleled specificity by utilizing the sterics and unique microenvironment present in large

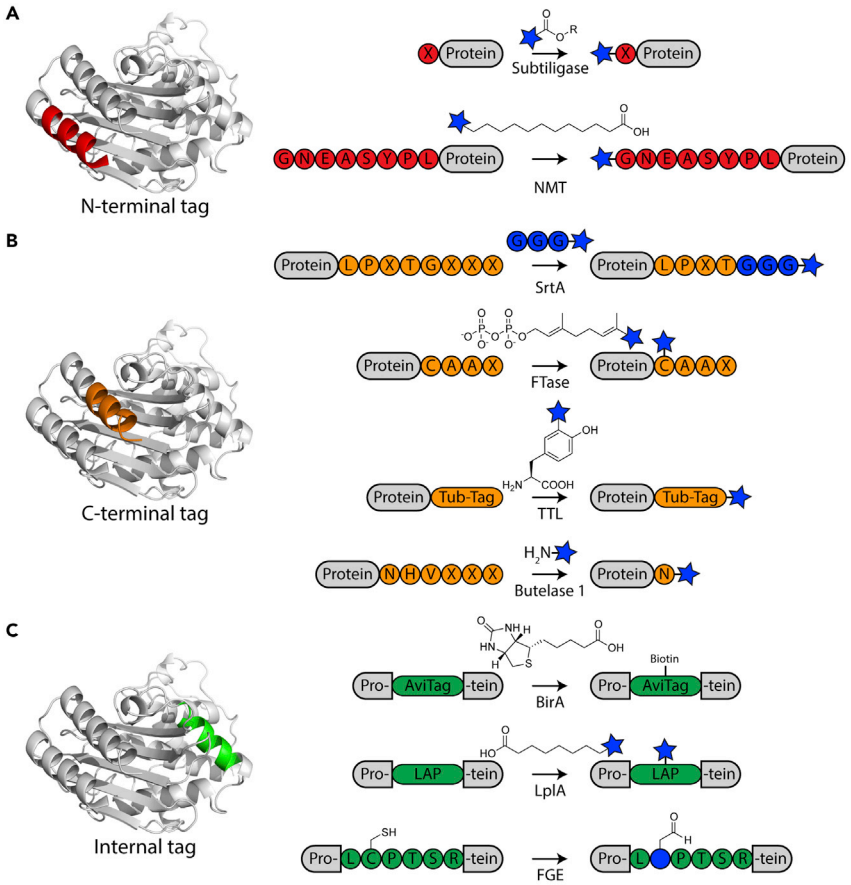

Figure 9. Chemoenzymatic Modification of Reactive Motifs within Proteins Affords Site-Specific Modification

(A) Subtiligase and NMT act upon N-terminal sequences appended onto the protein.

(B) SrtA, FTase, TTL, and butelase 1 are commonly used to derivatize proteins near their C terminus.

(C) BirA, LplA, and FGE can act upon internal motifs within the protein.

biomacromolecules to change the energy landscape of a reaction. Taking advantage of these enzymes for their selectivity, in combination with engineered efforts to expand their functionality, has enabled unprecedented site-selective modification of full-length proteins.[93] Many enzymes have been discovered that uniquely permit creation and modification of biomaterials with bioactive proteins (Figure 9).

Sortase A

Sortase A (SrtA) is a transpeptidase that canonically recognizes the sorting signal (LPXTG) on proteins and conjugates an aminoglycine nucleophile in the presence of calcium. Originally found in *Staphylococcus aureus* as the enzyme responsible for tethering proteins to the cell wall,[94] SrtA's unique functionality has gained significant popularity for the "sortagging" reaction, which is commonly exploited to modify either protein termini.[95] In brief, the cysteine residue in the active pocket of SrtA forms a thioester intermediate by cleaving between the threonine and glycine in the LPXTG motif. The thioester is then displaced by an aminoglycine nucleophile, typically an N-terminal polyglycine sequence on a peptide or a protein, forming a stable amide bond. This reaction reforms the LPXTG sorting signal, making the process reversible. Much work has been performed in evolving the SrtA

enzyme, including variants that have a higher catalytic activity, are calcium independent, and recognize different substrates.[96] A reaction cascade involving sortase modification and split inteins was recently reported to modify proteins without unwanted amino acid "scars."[97]

SrtA has been used not only for PEGylation and in the creation of protein-protein conjugates but also for immobilizing proteins onto and within materials. By tethering a pentaglycine peptide onto surfaces, LPETG-containing proteins could be reversibly conjugated.[98] The Griffith group created synthetic PEG hydrogels that contained an LPRTG-containing sortaggable substrate, which was used to immobilize a polyglycine-EGF species to promote DNA synthesis in primary human hepatocytes and endometrial epithelial cells.[99] In another example, GFP with a C-terminal LPETG tag was immobilized onto polystyrene beads containing a triglycine motif.[100]

Protein engineering to create protein-sortase fusions for intramolecular sortagging has also been performed to create single-step purification/conjugation strategies.[101] Proteins of interest are expressed as an N-terminal fusion to a 6xHis-containing SrtA enzyme with an LPETG spacer. Upon addition of calcium and a GGG-tagged probe, intramolecular transpeptidation by the SrtA enzyme conjugates the GGG-tagged probe to the C terminus of the protein of interest while displacing the 6xHis SrtA domain during immobilized metal affinity chromatography. This technique has been used to conjugate photoactive and clickable moieties onto various proteins for the four-dimensional photoreversible protein patterning of hydrogels (Figure 10).[102] Sortagged and tethered proteins were found to be bioactive and able to spatiotemporally govern cell proliferation, intracellular mitogen-activated protein kinase signaling, and subcellularly resolved receptor endocytosis. In another example, this strategy was used to install stimuli-labile functionalities onto proteins that would dictate how they were released from environmentally responsive biomaterials following Boolean YES/OR/AND logic.[103]

While traditionally SrtA has been used to modify the N and C termini of proteins, recent work has enabled sortagging of internal lysine residues through isopeptide bond formation. Isopeptide bond formation is achieved by encoding the pilin domain (WXXXVXVYPK) to make the ε-amino group of lysine a more reactive nucleophile. This reaction was found to be selective and achieved high conversion (>95%) on peptide substrates and on antibodies.[104] This technique enables the facile incorporation of multiple sortag modifications per protein, and could be used in cases where the protein termini are inaccessible or would have deleterious effects on protein activity.

Subtiligase

Subtiligase is an engineered ligase that conjugates C-terminal activated esters, such as a thioester or a glycolate phenylalanyl amide activated ester, with the N terminus of peptides and proteins.[105] Subtiligase has been used in the synthesis of full-length ribonuclease A from six peptide fragments, as well as in the conjugation of phosphopeptides to phosphatase and tensin homolog PTEN in the investigation of post-translational phosphorylation on protein activity.[106] Conversion is relatively good (~70%) after a 60-min reaction. Subtiligase has found utility in the proteomics community, where labeling free N termini is useful in identifying proteolytic cleavage sites.[107,108] Given its promiscuity in modifying N-terminal substrates, labeling a specific target protein using Subtiligase in a complicated biological mixture or cell-laden environment remains a challenge.

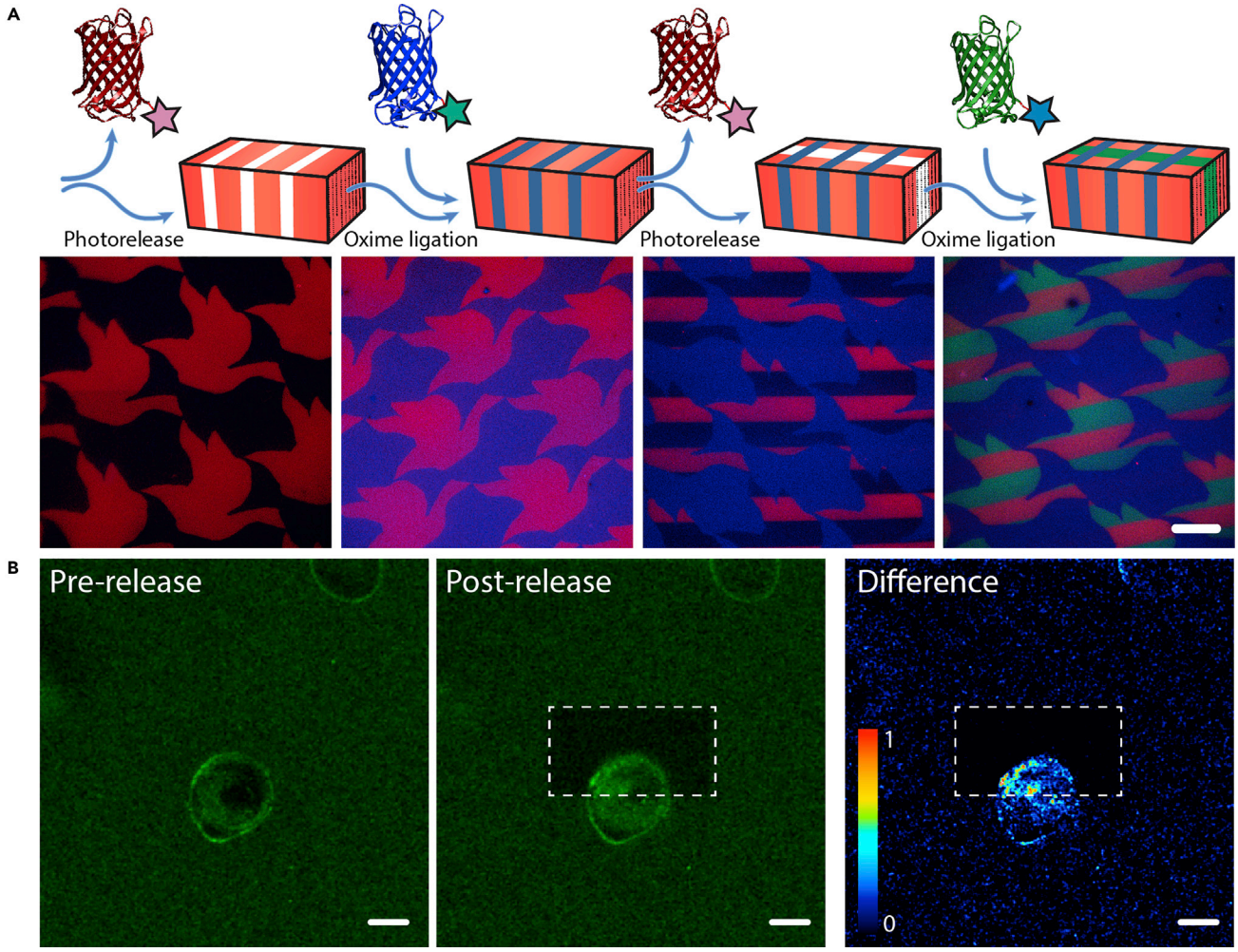

Figure 10. Photoreversible Patterning of Gel Biomaterials with Site-Specifically Modified Proteins

(A) A photomediated oxime ligation/*ortho*-nitrobenzyl cleavage sequence is used to reversibly immobilize site-specifically modified proteins within gels using masked light. Photoreactive functional groups were installed onto proteins C-terminally using SrtA.

(B) Photorelease of a GFP-epidermal growth factor fusion protein (GFP-EGF, green) is confined to gel subvolumes that bisect a single A431 cell in 3D culture. The endosome formation is visible in <5 min and is concentrated in the regions of light exposure. Fluorescent images correspond to time points immediately preceding and 5 min after the protein photorelease within a single gel. Difference calculations between images pre-release and post-release highlight the local GFP-EGF internalization. Scale bars, 10 μm.

Image modified with permission from Shadish et al.[102] Copyright (2019), Springer Nature.

Biotin Ligase

Biotin ligase (BirA) is an *Escherichia coli*-derived enzyme that appends a biotin group to the ε-amine of lysine in an ATP-dependent reaction. Biotin ligase's natural substrate is the biotin carboxyl carrier protein (BCCP), a relatively large domain of 75 amino acids. Although the binding of BirA to BCCP has been used to successfully immobilize immunoglobulin G to gold surfaces, the long length of BCCP may be problematic for some proteins.[109] A more promiscuous biotin ligase has also been developed, BioD2, which when fused to a bait protein allows for biotinylation of proximate proteins for investigation into protein-protein interactions.[110] However, for site-specific labeling, phage display has revealed a short peptide substrate for localized biotinylation.[111] Biotin labeling of full-length proteins has been achieved by adding the AviTag (GLNDIFEAQKIEWHE) to a terminus or internal loop.[112] The

biotin-streptavidin bond is one of the most stable non-covalent bonds, making it a promising strategy for immobilization within streptavidin-containing materials.

In one example exploiting the specificity and orthogonality of this non-covalent interaction, the Shoichet lab labeled ciliary neutrophic factor with the AviTag for biotinylation, and Sonic Hedgehog with Barnstar, which strongly interacts with its binding partner Barnase.[113] Streptavidin and Barnase were subsequently immobilized into agarose gels using two-photon lithography. Upon swelling of biotin- and Barnstar-labeled proteins, exquisite 3D control over protein immobilization was achieved. These immobilized proteins were used to guide migration of adult neural precursor cells within the hydrogels. In another example, the Leipzig lab added an N-terminal AviTag to IFN-γ, platelet-derived growth factor AA, and bone morphogenic protein 2 for site-specific biotinylation of growth factors and subsequent immobilization into streptavidin-functionalized methacrylamide chitosan gels.[114] They found that the site-specifically tethered growth factors led to enhanced lineage-specific differentiation in comparison with adsorbed growth factor treatments.

Microbial Transglutaminase

Microbial transglutaminase facilitates the transamidation of glutamine side chains to lysine residues to create an ε-(γ-glutamyl)lysine isopeptide bond.[115] Although it has found tremendous value for improving texture in foods, the original incarnation lacked specificity. Recent developments in using transglutaminase for protein modification have improved specificity and enabled site-specific protein modification and biomaterial formation.[116] Most typically, increased specificity has been achieved through the engineering of glutamyl- or amine-donor peptides with transglutaminase substrates (e.g., Q-tag, YAHQAHY; K-tag, KKKKKK) for *Bacillus subtilis*.[117] These substrates increase the reactivity of the specific residue, allowing for site-specific transglutaminase conjugation. In one example, biotin ligase with an N-terminal Q-tag was immobilized onto magnetic microspheres and was shown to retain >95% activity.[118] Newly discovered microbial transglutaminases, such as that from *Kutzneria albida*, show high specificity for their substrates (YRYRQ and RYESK) and have been used for the site-specific labeling of antibodies with biotin.[119]

Hydrogel biomaterials have also been formed through transglutaminase-mediated crosslinking. Hyaluronic acid was functionalized with glutamine, and an eight-arm PEG was end functionalized with lysine residues for transglutaminase-assisted crosslinking.[120] Hydrogels formed in this manner showed increased adhesion and spreading of MCF7 and C2C12 cells. To facilitate spatial control over the microbial transglutaminase ligation, the active lysine in the K domain of the substrate of an engineered transglutaminase (FXIIIa) was photocaged.[121,122] This photocaged substrate has been stochastically immobilized into PEG hydrogels and spatiotemporally activated upon exposure to UV light. Q peptides were engineered into proteins (e.g., vascular endothelial growth factor, platelet-derived growth factor), and upon treatment with FXIIIa would be covalently tethered into the photoactivated regions of the hydrogel. Immobilized growth factors were found to drive human mesenchymal stem cell migration. This combination of photocaging a specific residue in a substrate domain with enzymatic labeling is a promising avenue toward spatiotemporal biomaterials generation, one that could potentially be extended to other substrate/enzyme pairs.

Farnesyltransferase

Farnesyltransferase recognizes the CA_1A_2X motif (where A_1 and A_2 are any aliphatic amino acid, and X is any amino acid) on the C terminus of proteins and installs a

farnesyl pyrophosphate group at the cysteine residue to form a thioether bond in a process called prenylation. Farnesyl analogs have been synthesized with a variety of functional handles, including aldehydes, azides, and alkynes.[123,124] Azide- and alkyne-containing analogs showed the best incorporation using farnesyltransferase, with kinetic constants comparable with incorporation of the native substrate.[125] Azide-functionalized GFP was immobilized onto alkyne-modified agarose beads and shown to retain its fluorescence. Further investigation revealed a benzaldehyde-containing substrate that could be efficiently incorporated onto GFP.[126] This benzaldehyde was to create a fluorescence resonance energy transfer pair between GFP/Texas red or GFP/TAMRA, respectively, through hydrazone or oxime ligation. Aldehyde-functionalized proteins could also be immobilized onto hydrazide beads and released through the addition of aminooxy-PEG, leading to a single-step purification/PEGylation strategy.

Other prenyltransferases, such as geranylgeranyltransferase (GGTase), have also been used to enable orthogonal labeling strategies for dual site-specific protein labeling. The Distefano group found that farnesyltransferase best recognized the CVIA peptide sequence, while GGTase recognized CVLL. GGTase was also found to incorporate bulkier and longer isoprenoids, enabling one-pot labeling of GFP-CVLL and RFP-CVIA with an azide-containing farnesyltransferase substrate and an alkyne-containing GGTase substrate.[127] The two functionalized proteins could then be conjugated together through CuAAC click chemistry.

N-Myristoyltransferase

N-Myristoyltransferase (NMT) is a eukaryotic enzyme that installs a fatty acid group to the N terminus of proteins for anchoring to the cell membrane. NMT has been shown to be promiscuous in substrate affinity, and myristic acid analogs containing azides, alkynes, carbonyls, and halogens can be incorporated site-specifically when an N-terminal NMT recognition motif is appended to the protein of interest.[128–130] Expression systems in *E. coli* that co-express the protein of interest, NMT, and a methionine aminopeptidase for excision of the N-terminal methionine permit near-quantitative protein functionalization.[130,131] As *E. coli* does not naturally utilize the NMT enzyme, myristoylation was found to be orthogonal to the *E. coli* proteome. Myristic acid has been installed onto an elastin-like polypeptide to generate temperature-responsive liposomes for doxorubicin and paclitaxel encapsulation,[132] as well as in the synthesis of genetically encoded hybrid self-assembled materials (Figure 11).[133]

NMT has also been used to create modified proteins for direct incorporation into biomaterials. One example used NMT and SrtA to functionalize both the N and C termini of the light-responsive protein LOV2-Jα with reactive azides.[134] The dual-functionalized LOV2-Jα was then used to crosslink a PEG hydrogel, such that the light-driven conformational change of the protein crosslinker would drive changes the bulk gel mechanics. In another example, NMT was used to install a reactive azide onto the N terminus of fusion proteins containing PhoCl and a C-terminal protein of interest.[135] PhoCl is an engineered fluorescent protein from *Clavularia* sp. that undergoes intramolecular bond cleavage upon exposure to violet light.[136] After covalent immobilization in PEG hydrogels through an NMT-installed azide, exposure to violet light causes cleavage in the PhoCl domain and liberation of the protein of interest. This allowed for the facile creation of protein-patterned biomaterials. Among the enzymes that have been utilized for protein modification, NMT is one of few options that can act co-translationally for intracellular protein labeling.

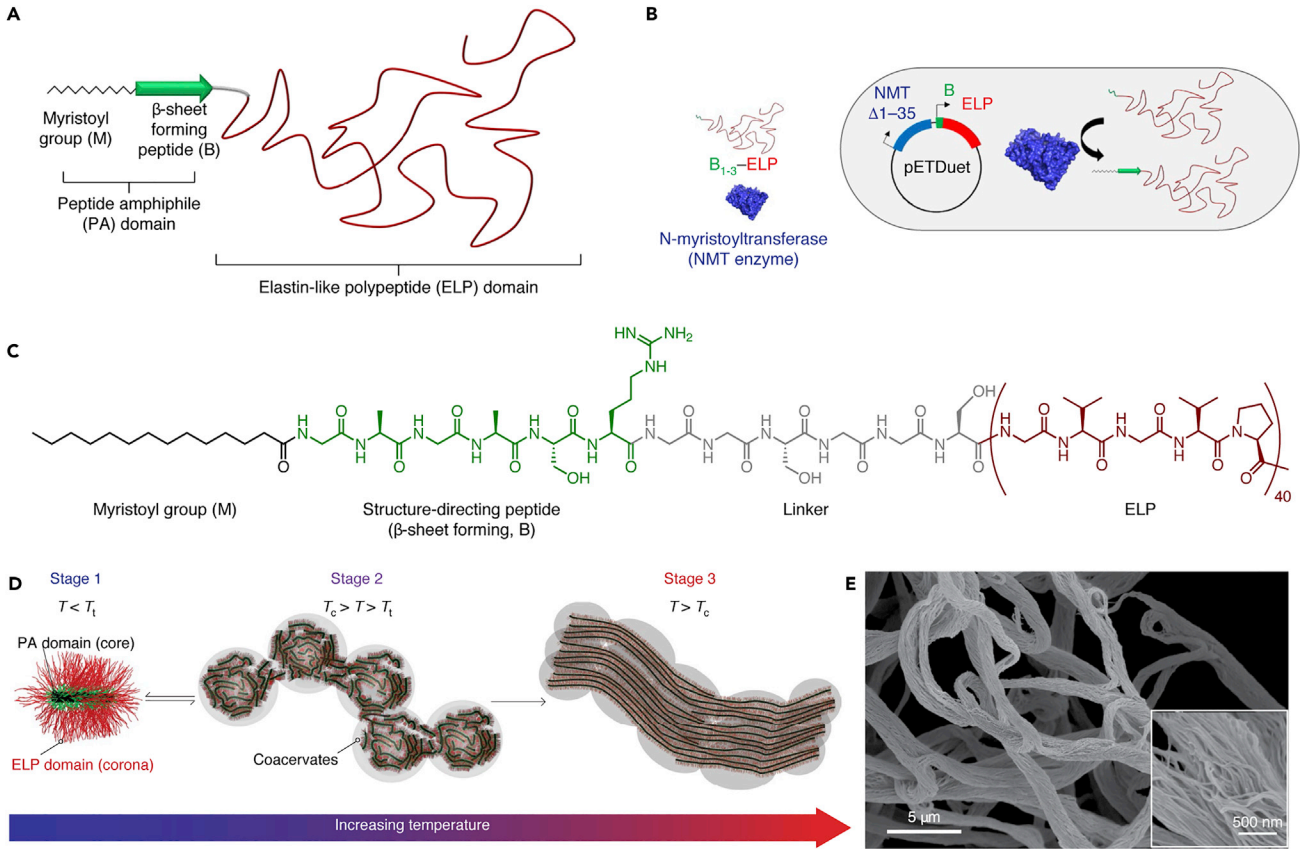

Figure 11. Genetically Encoded Lipid-Polypeptide Temperature-Responsive Biomaterials

(A) Fatty-acid-modified elastin-like polypeptides (FAMEs) consist of a myristoyl group (zigzag chain, M), a β-sheet-forming peptide (green arrow, B), and an ELP domain (red).

(B) A dual expression strategy is employed to create the myristolylated FAME through co-translational modification involving NMT.

(C) Chemical structure of the FAME subunit.

(D) FAMEs undergo a three-staged mechanism of self-assembly. Below the first transition temperature (T_t), FAMEs self-assemble into nanostructures whereby the hydrophobic myristoyl group is buried within the hydrophilic ELP protein corona. Between T_t and a second transition temperature (T_c), the dehydrated ELP domain undergoes a lower critical solution temperature-based phase transition into spherical droplets of a liquid-like coacervate. Above T_c, macroscale self-assembly occurs.

(E) Scanning electron micrograph of macroscopic aggregate formed from FAMEs heated above T_c.

Image modified with permission from Mozhdehi et al.[133] Copyright (2018), Springer Nature.

Phosphopantetheinyl Transferase

Coenzyme A (CoA) analogs can be efficiently labeled onto a serine residue of the DSLEFIASKLA (ybbR) tag via phosphopantetheinyl transferase (Sfp). This tag can be inserted into internal loops of the protein, as well as the N or C terminus, and allows for efficient incorporation of diverse CoA analogs containing biotin, fluorescein, TAMRA, Texas red, and other compounds.[137] Attaching CoA to a PEG-acrylamide resin, luciferase, and glutathione-*S*-transferase have been covalently immobilized directly from cell lysate while retaining high activity.[138] Sfb has also been used to quantitatively biotinylate proteins that could be immobilized on streptavidin surfaces for phage display binding experiments.[139] Despite these successes, Sfb may be less efficient than other strategies in some contexts; in contrast to SrtA, which could functionalize polyglycine-modified surfaces, the Gaub group found that Sfb was not able to immobilize GFP onto a CoA-modified biolayer interferometer surface.[140] These findings further highlight the importance of considering reaction sterics when choosing a chemoenzymatic strategy for protein modification.

Tubulin Tyrosine Ligase

Tubulin tyrosine ligase (TTL) is an enzyme naturally involved in microtubule maintenance.[141,142] It recognizes the "Tub-Tag" (VDSVEGEGEEEGEE) and appends tyrosine analogs including those containing azide- and aldehyde-functional handles to the C terminus.[143] Nanobodies, GFP, and ubiquitin were all modified with high conversion (>99%). Azide-modified proteins were demonstrated susceptible to further derivatization using click chemistry for super-resolution microscopy and immunoprecipitation experiments. Additional investigation of the substrate tolerance of TTL has revealed further promiscuity in substrates, including biotin and fluorophore-containing tyrosine analogs.[144] This allowed for single-step protein functionalization with the fluorescent dye coumarin; TTL coumarin-labeled Annexin V was found to be as active as a commercial Alexa 350-conjugated Annexin V probe.

Lipoic Acid Ligase

Lipoic acid ligase (LplA) recognizes the lipoic acceptor peptide (LAP) sequence (GFEIDKVWYDLDA) and conjugates a lipoic acid group to the side chain of the lysine residue. Pioneered by the Ting group for fluorophore conjugation, LplA has been shown to accept many other substrates.[145] The LAP tag can be added to an internal loop or the N or C terminus of the protein of interest, and treatment with LplA proceeds rapidly and near quantitatively. Further investigation into substrate specificity revealed substrates with clickable handles for subsequent conjugation.

Azide moieties were installed to the N terminus and two internal sites of GFP using a lipoic acid analog (10-azidodecanoic acid).[146] Near-complete conversion of LAP-GFP was achieved after 1 h of incubation with LplA and 60-fold molar excess of substrate. These sites were then PEGylated with an alkyne PEG via CuAAC/SPAAC to give a multiple site-specifically PEGylated protein. Use of a mutant LplA allowed for the incorporation of a stable norbornene derivative for subsequent reaction with tetrazine-containing compounds.[147] The technique has been used for cell-surface labeling of an LAP-containing transmembrane protein.[148]

Formylglycine-Generating Enzyme

Unlike many chemical and enzymatic modifications that append bulky substituents to a protein, formylglycine-generating enzyme (FGE) enables the direct oxidation of a critical cysteine residue to the aldehyde-containing formylglycine.[149] Aldehydes are convenient tags for bioconjugation via oxime ligation or hydrazone formation.[150,151] FGE selectively modifies the cysteine in the LCTPSR motif; as this short peptide tag can be genetically installed anywhere on a protein, FGE can uniquely install an aldehyde on virtually any accessible region of the protein.[152] FGE has been used to modify a haloacid dehalogenase enzyme ST2570 for immobilization into amine-functionalized flow reactors for conversion of L-2-haloalkanoic acid;[153] the immobilized enzyme was found to have a higher catalytic activity and thermal stability than the free enzyme. FGE is a promising, though relatively unexplored, enzyme for materials applications. This may be in part explained by the limited functionality of the enzyme; while other conversion techniques have broad substrate specificity and are able to append many functional groups, FGE is limited to relatively simple aldehyde modification (but powerfully at any location).

Butelase 1 Ligase

Butelase 1 is a recently discovered Asx peptide ligase that is able to append C-terminal NHV or DHV sequences to nucleophilic N-terminal domains.[154] Butelase has very high activity, roughly 20,000 times faster than SrtA, and has been used for peptide ligation and protein cyclization.[155] One potential downside of this technique is

that the enzyme can also act as a peptidase, although careful control of the reaction conditions can limit the hydrolyzed product to <5%. While a rapid and efficient enzyme, its high reactivity means that the N terminus of the NHV- or DHV-containing domain must be blocked or otherwise hindered to prevent cyclization when performing peptide ligation. One possible solution is to engineer an N-terminal proline residue, which is unable to act as a nucleophile for ligation. While butelase 1 has not been used in a materials setting as of authoring this publication, this strategy is likely to find great utility in creating cyclic peptides and proteins, which often have unique therapeutic effects.[156]

CONCLUSION AND FUTURE OUTLOOK

While this review has highlighted many newly developed chemistries that offer unparalleled control over protein functionalization, there is no one-size-fits-all solution for every application. Wild-type proteins may be modified if they are amenable to specific reaction conditions and can be obtained at reasonable purity through the use of small-molecule reagents to label endogenous residues or a terminus, though often with relatively low conversion and selectivity. Otherwise, genetic engineering approaches, either by genetic code expansion or chemoenzymatic strategies, enable exquisite selectivity on otherwise intractable protein domains in complicated biological reaction mixtures such as the inside of a cell. However, use of these techniques typically requires protein engineering and expression expertise. Additional consideration concerning reaction orthogonality must be paid when attempting multiple modifications per protein, whether it be to create protein-based material crosslinks or to add additional functionalities to an enzyme. Ultimately, combinations of several modification strategies may be necessary for engineering advanced materials.

Although peptide screening and phage display can often prove useful in identifying new site-selective modification chemistries, prediction of the effects of the complex tertiary structure of full-length proteins remains challenging. Increasing computational abilities will accelerate the development of new techniques and allow for more accurate predictions of their efficacy in modifying a given protein.[157,158] As even relatively short peptides have an enormous set of potential sequences (e.g., there are more 20-mer peptides based only on canonical amino acids than the number of stars in the observable universe), machine learning methodologies will be critical for focusing synthetic and screening efforts. Artificial intelligence has seen recent adoption by the oncology community, by whom it has been used to screen peptides for treating tumors.[159] Similar techniques applied to a materials context could generate novel peptides and proteins with new and/or enhanced functionality.

Despite the large range of functional proteins and a variety of strategies that exist for their controlled modification, relatively few protein species have been incorporated into the biomaterials space. Opportunities exist to create "smart" biomaterials that respond to their environment by incorporating proteins that undergo physicochemical changes upon exposure to specific stimuli.[160] While there exist thousands of stimuli-responsive proteins, only a few have been used in materials. Calmodulin, a protein that undergoes a conformational change to bind calcium ions, has been used to create hydrogels whose mechanical properties are a function of calcium concentration.[134,161] Moreover, the field of optogenetics has diligently been identifying, engineering, and utilizing a host of light-responsive proteins to spatiotemporally control cell fate;[162–164] these proteins could be co-opted to create phototunable materials.

As highlighted, there is now an expansive toolbox readily available to modify functional proteins for biomaterials applications. There is a trade-off between "off-the-shelf" labeling of endogenous proteins and the specificity of reaction. Although chemoselective methods for amine/thiol modification have been used with great success, they have typically only been applied to robust proteins or in applications where retaining even a fraction of activity is tolerable. Complementary techniques that employ protein engineering and chemoenzymatic chemistries have proved more amenable to modifying fragile proteins, but have required in-house protein expression using techniques that may not be readily available to all biomaterial scientists. Such drawbacks are often overshadowed by the precise control over the site of modification, enabling the creation of protein-based materials with novel functionalities that would be impossible to replicate with small molecules. The protein modification field has seen exponential growth in the number of techniques available, and the time is ripe for applying these methodologies to create the next generation of functional biomaterials.

ACKNOWLEDGMENTS

The authors thank E. Ruskowitz for the helpful feedback during preparation of the manuscript. This work was supported by a University of Washington Faculty startup grant (C.A.D.), a Jaconette L. Tietze Young Scientist Research award (C.A.D.), as well as a standard grant (CBET 1803054, C.A.D.) and a Faculty Early Career Development CAREER award (DMR 1652141, C.A.D.) from the National Science Foundation.

AUTHOR CONTRIBUTIONS

Both authors contributed equally in the preparation of this Review.

REFERENCES

1. Lagassé, H.A.D., Alexaki, A., Simhadri, V.L., Katagiri, N.H., Jankowski, W., Sauna, Z.E., and Kimchi-Sarfaty, C. (2017). Recent advances in (therapeutic protein) drug development. F1000Res. *6*, 113.
2. Leader, B., Baca, Q.J., and Golan, D.E. (2008). Protein therapeutics: a summary and pharmacological classification. Nat. Rev. Drug Discov. *7*, 21–39.
3. Fisher, S.A., Baker, A.E.G., and Shoichet, M.S. (2017). Designing peptide and protein modified hydrogels: selecting the optimal conjugation strategy. J. Am. Chem. Soc. *139*, 7416–7427.
4. Coons, A.H., Creech, J., Norman, R., and Berliner, E. (1942). The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody 1. J. Immunol. *45*, 159–170.
5. Engvall, E., and Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. Immunochemistry *8*, 871–874.
6. Tamura, T., and Hamachi, I. (2019). Chemistry for covalent modification of endogenous/native proteins: from test tubes to complex biological systems. J. Am. Chem. Soc. *141*, 2782–2799.
7. Spicer, C.D., and Davis, B.G. (2014). Selective chemical protein modification. Nat. Commun. *5*, 4740.
8. Chin, J.W. (2017). Expanding and reprogramming the genetic code. Nature *550*, 53–60.
9. Kiick, K.L., Saxon, E., Tirrell, D.A., and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc. Natl. Acad. Sci. U S A *99*, 19–24.
10. Johnson, J.A., Lu, Y.Y., Van Deventer, J.A., and Tirrell, D.A. (2010). Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. Curr. Opin. Chem. Biol. *14*, 774–780.
11. Zeymer, C., and Hilvert, D. (2018). Directed evolution of protein catalysts. Annu. Rev. Biochem. *87*, 131–157.
12. Tobin, P.H., Richards, D.H., Callender, R.A., Wilson, C.J., Haven, N., Haven, N., and Haven, N. (2016). Protein engineering: a new frontier for biological therapeutics. Curr. Drug Metab. *15*, 743–756.
13. Chen, Z., and Zeng, A.P. (2016). Protein engineering approaches to chemical biotechnology. Curr. Opin. Biotechnol. *42*, 198–205.
14. Chernov, K.G., Redchuk, T.A., Omelina, E.S., and Verkhusha, V.V. (2017). Near-infrared fluorescent proteins, biosensors, and optogenetic tools engineered from phytochromes. Chem. Rev. *117*, 6423–6446.
15. Anderson, G.W., Zimmerman, J.E., and Callahan, F.M. (1964). The use of esters of N-hydroxysuccinimide in peptide synthesis. J. Am. Chem. Soc. *86*, 1839–1842.
16. Fisher, S.A., Tam, R.Y., Fokina, A., Mahmoodi, M.M., Distefano, M.D., and Shoichet, M.S. (2018). Photo-immobilized EGF chemical gradients differentially impact breast cancer cell invasion and drug response in defined 3D hydrogels. Biomaterials *178*, 751–766.
17. DeForest, C.A., and Tirrell, D.A. (2015). A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. Nat. Mater. *14*, 523–531.
18. Dodson, G., and Wlodawer, A. (1998). Catalytic triads and their relatives. Trends Biochem. Sci. *23*, 347–352.
19. Nanna, A.R., Li, X., Walseng, E., Pedzisa, L., Goydel, R.S., Hymel, D., Burke, T.R., Roush, W.R., and Rader, C. (2017). Harnessing a catalytic lysine residue for the one-step preparation of homogeneous antibody-drug conjugates. Nat. Commun. *8*, 1112.
20. Matos, M.J., Oliveira, B.L., Martínez-Sáez, N., Guerreiro, A., Cal, P.M.S.D., Bertoldo, J., Maneiro, M., Perkins, E., Howard, J., Deery, M.J., et al. (2018). Chemo- and regioselective lysine modification on native proteins. J. Am. Chem. Soc. *140*, 4004–4017.
21. Chilamari, M., Purushottam, L., and Rai, V. (2017). Site-selective labeling of native

proteins by a multicomponent approach. Chemistry *23*, 3819–3823.

22. Yoshitake, S., Imagawa, M., Ishikawa, E., Niitsu, Y., Urushizaki, I., Nishiura, M., Kanazawa, R., Kurosaki, H., Tachibana, S., Nakazawa, N., and Ogawa, H. (1982). Mild and efficient conjugation of rabbit fab′ and horseradish peroxidase using a maleimide compound and its use for enzyme immunoassay. J. Biochem. *92*, 1413–1424.

23. Fontaine, S.D., Reid, R., Robinson, L., Ashley, G.W., and Santi, D.V. (2015). Long-term stabilization of maleimide-thiol conjugates. Bioconjug. Chem. *26*, 145–152.

24. Hörner, M., Raute, K., Hummel, B., Madl, J., Creusen, G., Thomas, O.S., Christen, E.H., Hotz, N., Gübeli, R.J., Engesser, R., et al. (2019). Phytochrome-based extracellular matrix with reversibly tunable mechanical properties. Adv. Mater. *31*, 1806727.

25. Boutureira, O., Martínez-Sáez, N., Brindle, K.M., Neves, A.A., Corzana, F., and Bernardes, G.J.L. (2017). Site-selective modification of proteins with oxetanes. Chemistry *23*, 6483–6489.

26. Kolodych, S., Koniev, O., Baatarkhuu, Z., Bonnefoy, J.-Y., Debaene, F., Cianférani, S., Van Dorsselaer, A., and Wagner, A. (2015). CBTF: new amine-to-thiol coupling reagent for preparation of antibody conjugates with increased plasma stability. Bioconjug. Chem. *26*, 197–200.

27. Shaunak, S., Godwin, A., Choi, J.-W., Balan, S., Pedone, E., Vijayarangam, D., Heidelberger, S., Teo, I., Zloh, M., and Brocchini, S. (2006). Site-specific PEGylation of native disulfide bonds in therapeutic proteins. Nat. Chem. Biol. *2*, 312–313.

28. Lee, M.T.W., Maruani, A., Baker, J.R., Caddick, S., and Chudasama, V. (2016). Next-generation disulfide stapling: reduction and functional re-bridging all in one. Chem. Sci. *7*, 799–802.

29. Wang, T., Riegger, A., Lamla, M., Wiese, S., Oeckl, P., Otto, M., Wu, Y., Fischer, S., Barth, H., Kuan, S.L., and Weil, T. (2016). Water-soluble allyl sulfones for dual site-specific labelling of proteins and cyclic peptides. Chem. Sci. *7*, 3234.

30. Griebenow, N., Dilmaç, A.M., Greven, S., and Bräse, S. (2016). Site-specific conjugation of peptides and proteins via rebridging of disulfide bonds using the Thiol-Yne coupling reaction. Bioconjug. Chem. *27*, 911–917.

31. Ruddle, B., Fleming, R., Wu, H., Gao, C., and Dimasi, N. (2019). Characterization of disulfide bond rebridged fab-drug conjugates prepared using a dual maleimide pyrrolobenzodiazepine cytotoxic payload. ChemMedChem *14*, 1185–1195.

32. Hoyle, C.E., and Bowman, C.N. (2010). Thiol-Ene click chemistry. Angew. Chem. Int. Ed. *49*, 1540–1573.

33. Grim, J.C., Brown, T.E., Aguado, B.A., Chapnick, D.A., Viert, A.L., Liu, X., and Anseth, K.S. (2018). A reversible and repeatable Thiol-Ene bioconjugationfor dynamic patterning of signaling proteins in hydrogels. ACS Cent. Sci. *4*, 909.

34. Williams, C.G., Malik, A.N., Kim, T.K., Manson, P.N., and Elisseeff, J.H. (2005). Variable cytocompatibility of six cell lines with photoinitiators used for polymerizing hydrogels and cell encapsulation. Biomaterials *26*, 1211–1218.

35. Imiołek, M., Karunanithy, G., Ng, W.-L., Baldwin, A.J., Gouverneur, V., and Davis, B.G. (2018). Selective radical trifluoromethylation of native residues in proteins. J. Am. Chem. Soc. *140*, 1568–1571.

36. Yu, Y., Zhang, L.-K., Buevich, A.V., Li, G., Tang, H., Vachal, P., Colletti, S.L., and Shi, Z.-C. (2018). Chemoselective peptide modification via photocatalytic tryptophan β-position conjugation. J. Am. Chem. Soc. *140*, 6797–6800.

37. Rosen, C.B., and Francis, M.B. (2017). Targeting the N terminus for site-selective protein modification. Nat. Chem. Biol. *13*, 697–705.

38. Obermeyer, A.C., Jarman, J.B., and Francis, M.B. (2014). N-terminal modification of proteins with *o*-aminophenols. J. Am. Chem. Soc. *136*, 9572–9579.

39. Chen, D., Disotuar, M.M., Xiong, X., Wang, Y., and Chou, D.H.-C. (2017). Selective N-terminal functionalization of native peptides and proteins. Chem. Sci. *8*, 2717–2722.

40. Li, D., Han, B., Wei, R., Yao, G., Chen, Z., Liu, J., Poon, T.C.W., Su, W., Zhu, Z., Dimitrov, D.S., and Zhao, Q. (2018). N-terminal α-amino group modification of antibodies using a site-selective click chemistry method. MAbs *10*, 712–719.

41. Lee, J.P., Kassianidou, E., MacDonald, J.I., Francis, M.B., and Kumar, S. (2016). N-terminal specific conjugation of extracellular matrix proteins to 2-pyridinecarboxaldehyde functionalized polyacrylamide hydrogels. Biomaterials *102*, 268–276.

42. Gilmore, J.M., Scheck, R.A., Esser-Kahn, A.P., Joshi, N.S., and Francis, M.B. (2006). N-terminal protein modification through a biomimetic transamination reaction. Angew. Chem. Int. Ed. *45*, 5307–5311.

43. Witus, L.S., and Francis, M. (2010). Site-specific protein bioconjugation via a pyridoxal 5′-phosphate-mediated N-terminal transamination reaction. Curr. Protoc. Chem. Biol. *2*, 125–134.

44. Christman, K.L., Broyer, R.M., Tolstyka, Z.P., and Maynard, H.D. (2007). Site-specific protein immobilization through N-terminal oxime linkages. J. Mater. Chem. *17*, 2021.

45. Conibear, A.C., Watson, E.E., Payne, R.J., and Becker, C.F.W. (2018). Native chemical ligation in protein synthesis and semi-synthesis. Chem. Soc. Rev. *47*, 9046–9068.

46. Dhall, A., Weller, C.E., Chu, A., Shelton, P.M.M., and Chatterjee, C. (2017). Chemically sumoylated histone H4 stimulates intranucleosomal demethylation by the LSD1-CoREST complex. ACS Chem. Biol. *12*, 2275–2280.

47. Bandyopadhyay, A., Cambray, S., and Gao, J. (2016). Fast and selective labeling of N-terminal cysteines at neutral pH via thiazolidino boronate formation. Chem. Sci. *7*, 4589–4593.

48. Bloom, S., Liu, C., Kölmel, D.K., Qiao, J.X., Zhang, Y., Poss, M.A., Ewing, W.R., and MacMillan, D.W.C. (2018). Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. Nat. Chem. *10*, 205–211.

49. Marino, G., Eckhard, U., and Overall, C.M. (2015). Protein termini and their modifications revealed by positional proteomics. ACS Chem. Biol. *10*, 1754–1764.

50. Lotze, J., Reinhardt, U., Seitz, O., and Beck-Sickinger, A.G. (2016). Peptide-tags for site-specific protein labelling in vitro and in vivo. Mol. Biosyst. *12*, 1731–1745.

51. Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., and Schäfer, F. (2009). Immobilized-metal affinity chromatography (IMAC): a review. Methods Enzymol. *463*, 439–473.

52. Martos-Maldonado, M.C., Hjuler, C.T., Sørensen, K.K., Thygesen, M.B., Rasmussen, J.E., Villadsen, K., Midtgaard, S.R., Kol, S., Schoffelen, S., and Jensen, K.J. (2018). Selective N-terminal acylation of peptides and proteins with a Gly-His tag sequence. Nat. Commun. *9*, 3307.

53. Sunbul, M., Nacheva, L., and Jäschke, A. (2015). Proximity-induced covalent labeling of proteins with a reactive fluorophore-binding peptide tag. Bioconjug. Chem. *26*, 1466–1469.

54. Zhang, C., Dai, P., Vinogradov, A.A., Gates, Z.P., and Pentelute, B.L. (2018). Site-selective cysteine-cyclooctyne conjugation. Angew. Chem. Int. Ed. *57*, 6459–6463.

55. Zhang, C., Welborn, M., Zhu, T., Yang, N.J., Santos, M.S., Van Voorhis, T., and Pentelute, B.L. (2016). π-Clamp-mediated cysteine conjugation. Nat. Chem. *8*, 120–128.

56. Agrawalla, B.K., Wang, T., Riegger, A., Domogalla, M.P., Steinbrink, K., Dörfler, T., Chen, X., Boldt, F., Lamla, M., Michaelis, J., et al. (2018). Chemoselective dual labeling of native and recombinant proteins. Bioconjug. Chem. *29*, 29–34.

57. Peciak, K., Laurine, E., Tommasi, R., Choi, J., and Brocchini, S. (2019). Site-selective protein conjugation at histidine. Chem. Sci. *10*, 427–439.

58. Dieterich, D.C., Lee, J.J., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2007). Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. Nat. Protoc. *2*, 532–540.

59. tom Dieck, S., Müller, A., Nehring, A., Hinz, F.I., Bartnik, I., Schuman, E.M., and Dieterich, D.C. (2012). Current Protocols in Cell Biology (John Wiley & Sons, Inc.), Chapter 7, Unit 7.11.

60. Ferreira, S.A., Motwani, M.S., Faull, P.A., Seymour, A.J., Yu, T.T.L., Enayati, M., Taheem, D.K., Salzlechner, C., Haghighi, T., Kania, E.M., et al. (2018). Bi-directional cell-pericellular matrix interactions direct stem cell fate. Nat. Commun. *9*, 4049.

61. Loebel, C., Mauck, R.L., and Burdick, J.A. (2019). Local nascent protein deposition and remodelling guide mesenchymal stromal cell mechanosensing and fate in three-dimensional hydrogels. Nat. Mater. *18*, 883–891.

62. Ngo, J.T., and Tirrell, D.A. (2011). Noncanonical amino acids in the interrogation of cellular protein synthesis. Acc. Chem. Res. *44*, 677–685.

63. Zhang, X., Lu, W., Kwan, K., Bhattacharyya, D., and Wei, Y. (2017). Dual-functional-tag-facilitated protein labeling and immobilization. ACS Omega *2*, 522–528.

64. Wang, L., Brock, A., Herberich, B., and Schultz, P.G. (2001). Expanding the genetic code of *Escherichia coli*. Science *292*, 498–500.

65. Xiao, H., Xuan, W., Shao, S., Liu, T., and Schultz, P.G. (2015). Genetic incorporation of ε-N-2-hydroxyisobutyryl-lysine into recombinant histones. ACS Chem. Biol. *10*, 1599–1603.

66. Luo, X., Fu, G., Wang, R.E., Zhu, X., Zambaldo, C., Liu, R., Liu, T., Lyu, X., Du, J., Xuan, W., et al. (2017). Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. Nat. Chem. Biol. *13*, 845–849.

67. Wang, Z.A., Zeng, Y., Kurra, Y., Wang, X., Tharp, J.M., Vatansever, E.C., Hsu, W.W., Dai, S., Fang, X., and Liu, W.R. (2017). A genetically encoded allysine for the synthesis of proteins with site-specific lysine dimethylation. Angew. Chem. Int. Ed. *56*, 212–216.

68. Chin, J.W., Santoro, S.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. (2002). Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli*. J. Am. Chem. Soc. *124*, 9026–9027.

69. Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D.J., Blackman, M.L., Fox, J.M., and Chin, J.W. (2012). Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. J. Am. Chem. Soc. *134*, 10317–10320.

70. Chin, J.W., Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J.W. (2012). Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. Nat. Chem. *4*, 298–304.

71. Darko, A., Wallace, S., Dmitrenko, O., Machovina, M.M., Mehl, R.A., Chin, J.W., and Fox, J.M. (2014). Conformationally strained trans-cyclooctene with improved stability and excellent reactivity in tetrazine ligation. Chem. Sci. https://doi.org/10.1039/c4sc01348d.

72. Guo, C., Kim, H., Ovadia, E.M., Mourafetis, C.M., Yang, M., Chen, W., and Kloxin, A.M. (2017). Bio-orthogonal conjugation and enzymatically triggered release of proteins within multi-layered hydrogels. Acta Biomater. *56*, 80–90.

73. Liu, J., Hemphill, J., Samanta, S., Tsang, M., and Deiters, A. (2017). Genetic code expansion in zebrafish embryos and its application to optical control of cell signaling. J. Am. Chem. Soc. *139*, 9100–9103.

74. Nguyen, D.P., Mahesh, M., Elsässer, S.J., Hancock, S.M., Uttamapinant, C., and Chin, J.W. (2014). Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells. J. Am. Chem. Soc. *136*, 2240–2243.

75. Yang, T., Li, X.-M., Bao, X., Fung, Y.M.E., and Li, X.D. (2016). Photo-lysine captures proteins that bind lysine post-translational modifications. Nat. Chem. Biol. *12*, 70–72.

76. Yamaguchi, A., Matsuda, T., Ohtake, K., Yanagisawa, T., Yokoyama, S., Fujiwara, Y., Watanabe, T., Hohsaka, T., and Sakamoto, K. (2016). Incorporation of a doubly functionalized synthetic amino acid into proteins for creating chemical and light-induced conjugates. Bioconjug. Chem. *27*, 198–206.

77. Weller, M.G. (2000). Immunochromatographic techniques—a critical review. Fresenius J. Anal. Chem. *366*, 635–645.

78. Riggs, P. (2000). Expression and purification of recombinant proteins by fusion to maltose-binding protein. Mol. Biotechnol. *15*, 51–63.

79. Lata, S., and Piehler, J. (2005). Stable and functional immobilization of histidine-tagged proteins via multivalent chelator headgroups on a molecular poly(ethylene glycol) brush. Anal. Chem. *77*, 1096–1105.

80. Baneyx, F., and Schwartz, D.T. (2007). Selection and analysis of solid-binding peptides. Curr. Opin. Biotechnol. *18*, 312–317.

81. Altunbas, A., and Pochan, D.J. (2012). Peptide-based and polypeptide-based hydrogels for drug delivery and tissue engineering. Top Curr. Chem. *310*, 135–167.

82. Cole, N.B. (2013). Site-specific protein labeling with SNAP-tags. Curr. Protoc. Protein Sci. *73*. Unit 30.1.

83. Wasserberg, D., Uhlenheuer, D.A., Neirynck, P., Cabanas-Danés, J., Schenkel, J.H., Ravoo, B.J., An, Q., Huskens, J., Milroy, L.-G., Brunsveld, L., and Jonkheijm, P. (2013). Immobilization of ferrocene-modified SNAP-fusion proteins. Int. J. Mol. Sci. *14*, 4066–4080.

84. Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem. Biol. *3*, 373–382.

85. Kwon, Y., Han, Z., Karatan, E., Mrksich, M., and Kay, B.K. (2004). Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. Anal. Chem. *76*, 5713–5720.

86. Zakeri, B., Fierer, J.O., Celik, E., Chittock, E.C., Schwarz-Linek, U., Moy, V.T., and Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc. Natl. Acad. Sci. U S A *109*, E690–E697.

87. Sun, F., Zhang, W.B., Mahdavi, A., Arnold, F.H., and Tirrell, D.A. (2014). Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. Proc. Natl. Acad. Sci. U S A *111*, 11269–11274.

88. Gao, X., Fang, J., Xue, B., Fu, L., and Li, H. (2016). Engineering protein hydrogels using SpyCatcher-SpyTag chemistry. Biomacromolecules *17*, 39.

89. Lim, S., Jung, G.A., Muckom, R.J., Glover, D.J., and Clark, D.S. (2019). Engineering bioorthogonal protein-polymer hybrid hydrogel as a functional protein immobilization platform. Chem. Commun. (Camb.) *55*, 806–809.

90. Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. *60*, 523–533.

91. Fierer, J.O., Veggiani, G., and Howarth, M. (2014). SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. Proc. Natl. Acad. Sci. U S A *111*, E1176–E1181.

92. Khairil Anuar, I.N.A., Banerjee, A., Keeble, A.H., Carella, A., Nikov, G.I., and Howarth, M. (2019). Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. Nat. Commun. *10*, 1734.

93. Zhang, Y., Park, K.-Y., Suazo, K.F., and Distefano, M.D. (2018). Recent progress in enzymatic protein labelling techniques and their applications. Chem. Soc. Rev. *47*, 9106–9136.

94. Mazmanian, S.K., Liu, G., Hung, T.T., and Schneewind, O. (1999). *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. Science *285*, 760–763.

95. Popp, M.W., Antos, J.M., and Ploegh, H.L. (2009). Site-specific protein labeling via sortase-mediated transpeptidation. Curr. Protoc. Protein Sci. *Chapter 15*. Unit 15.3.

96. Antos, J.M., Truttmann, M.C., and Ploegh, H.L. (2016). Recent advances in sortase-catalyzed ligation methodology. Curr. Opin. Struct. Biol. *38*, 111–118.

97. Thompson, R.E., Stevens, A.J., and Muir, T.W. (2019). Protein engineering through tandem transamidation. Nat. Chem. *11*, 737–743.

98. Ham, H.O., Qu, Z., Haller, C.A., Dorr, B.M., Dai, E., Kim, W., Liu, D.R., and Chaikof, E.L. (2016). In situ regeneration of bioactive coatings enabled by an evolved Staphylococcus aureus sortase A. Nat. Commun. *7*, 11140.

99. Cambria, E., Renggli, K., Ahrens, C.C., Cook, C.D., Kroll, C., Krueger, A.T., Imperiali, B., and Griffith, L.G. (2015). Covalent modification of synthetic hydrogels with bioactive proteins via sortase-mediated ligation. Biomacromolecules *16*, 2316–2326.

100. Parthasarathy, R., Subramanian, S., and Boder, E.T. (2007). Sortase A as a novel molecular "stapler" for sequence-specific protein conjugation. Bioconjug. Chem. *18*, 469–476.

101. Warden-Rothman, R., Caturegli, I., Popik, V., and Tsourkas, A. (2013). Sortase-tag

expressed protein ligation: combining protein purification and site-specific bioconjugation into a single step. Anal. Chem. *85*, 11090–11097.

102. Shadish, J.A., Benuska, G.M., and DeForest, C.A. (2019). Bioactive site-specifically modified proteins for 4D patterning of gel biomaterials. Nat. Mater. *18*, 1005–1014.

103. Gawade, P.M., Shadish, J.A., Badeau, B.A., and DeForest, C.A. (2019). Logic-based delivery of site-specifically modified proteins from environmentally responsive hydrogel biomaterials. Adv. Mater. *31*, 1902492.

104. McConnell, S.A., Amer, B.R., Muroski, J., Fu, J., Chang, C., Ogorzalek Loo, R.R., Loo, J.A., Osipiuk, J., Ton-That, H., and Clubb, R.T. (2018). Protein labeling via a specific lysine-isopeptide bond using the pilin polymerizing sortase from *Corynebacterium diphtheriae*. J. Am. Chem. Soc. *140*, 8420–8423.

105. Chang, T.K., Jackson, D.Y., Burnier, J.P., and Wells, J.A. (1994). Subtiligase: a tool for semisynthesis of proteins. Proc. Natl. Acad. Sci. U S A *91*, 12544–12548.

106. Henager, S.H., Chu, N., Chen, Z., Bolduc, D., Dempsey, D.R., Hwang, Y., Wells, J., and Cole, P.A. (2016). Enzyme-catalyzed expressed protein ligation. Nat. Methods *13*, 925–927.

107. Yoshihara, H.A.I., Mahrus, S., and Wells, J.A. (2008). Tags for labeling protein N-termini with subtiligase for proteomics. Bioorg. Med. Chem. Lett. *18*, 6000–6003.

108. Wiita, A.P., Seaman, J.E., and Wells, J.A. (2014). Global analysis of cellular proteolysis by selective enzymatic labeling of protein N-termini. Methods Enzymol. *544*, 327–358.

109. Miyao, H., Ikeda, Y., Shiraishi, A., Kawakami, Y., and Sueda, S. (2015). Immobilization of immunoglobulin-G-binding domain of Protein A on a gold surface modified with biotin ligase. Anal. Biochem. *484*, 113–121.

110. Kim, D.I., Jensen, S.C., Noble, K.A., Kc, B., Roux, K.H., Motamedchaboki, K., and Roux, K.J. (2016). An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell *27*, 1188–1196.

111. Beckett, D., Kovaleva, E., and Schatz, P.J. (2008). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Sci. *8*, 921–929.

112. Fairhead, M., and Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. Methods Mol. Biol. *1266*, 171–184.

113. Wylie, R.G., Ahsan, S., Aizawa, Y., Maxwell, K.L., Morshead, C.M., and Shoichet, M.S. (2011). Spatially controlled simultaneous patterning of multiple growth factors in three-dimensional hydrogels. Nat. Mater. *10*, 799–806.

114. Li, H., Koenig, A.M., Sloan, P., and Leipzig, N.D. (2014). In vivo assessment of guided neural stem cell differentiation in growth factor immobilized chitosan-based hydrogel scaffolds. Biomaterials *35*, 9049–9057.

115. Ando, H., Adachi, M., Umeda, K., Matsuura, A., Nonaka, M., Uchio, R., Tanaka, H., and Motoki, M. (1989). Purification and characteristics of a novel transglutaminase derived from microorganisms. Agric. Biol. Chem. *53*, 2613–2617.

116. Strop, P. (2014). Versatility of microbial transglutaminase. Bioconjug. Chem. *25*, 855–862.

117. Oteng-Pabi, S.K., Clouthier, C.M., and Keillor, J.W. (2018). Design of a glutamine substrate tag enabling protein labelling mediated by *Bacillus subtilis* transglutaminase. PLoS One *13*, e0197956.

118. Yu, C.-M., Zhou, H., Zhang, W.-F., Yang, H.-M., and Tang, J.-B. (2016). Site-specific, covalent immobilization of BirA by microbial transglutaminase: a reusable biocatalyst for in vitro biotinylation. Anal. Biochem. *511*, 10–12.

119. Steffen, W., Ko, F.C., Patel, J., Lyamichev, V., Albert, T.J., Benz, J., Rudolph, M.G., Bergmann, F., Streidl, T., Kratzsch, P., et al. (2017). Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins. J. Biol. Chem. *292*, 15622–15635.

120. Ranga, A., Lutolf, M.P., Hilborn, J., and Ossipov, D.A. (2016). Hyaluronic acid hydrogels formed in situ by transglutaminase-catalyzed reaction. Biomacromolecules *17*, 1553–1560.

121. Mosiewicz, K.A., Kolb, L., van der Vlies, A.J., Martino, M.M., Lienemann, P.S., Hubbell, J.A., Ehrbar, M., and Lutolf, M.P. (2013). In situ cell manipulation through enzymatic hydrogel photopatterning. Nat. Mater. *12*, 1071–1077.

122. Griffin, D.R., Borrajo, J., Soon, A., Acosta-Velez, G.F., Oshita, V., Darling, N., Mack, J., Barker, T., Iruela-Arispe, M.L., and Segura, T. (2014). Hybrid photopatterned enzymatic reaction (HyPER) for in situ cell manipulation. Chembiochem *15*, 233–242.

123. Labadie, G.R., Viswanathan, R., and Poulter, C.D. (2007). Farnesyl diphosphate analogues with ω-bioorthogonal azide and alkyne functional groups for protein farnesyl transferase-catalyzed ligation reactions. J. Org. Chem. *72*, 9291–9297.

124. Jennings, B.C., Danowitz, A.M., Wang, Y.-C., Gibbs, R.A., Distefano, M.D., and Fierke, C.A. (2016). Analogs of farnesyl diphosphate alter CaaX substrate specificity and reactions rates of protein farnesyltransferase. Bioorg. Med. Chem. Lett. *26*, 1333–1336.

125. Duckworth, B.P., Xu, J., Taton, T.A., Guo, A., and Distefano, M.D. (2006). Site-specific, covalent attachment of proteins to a solid surface. Bioconjug. Chem. *17*, 967–974.

126. Rashidian, M., Song, J.M., Pricer, R.E., and Distefano, M.D. (2012). Chemoenzymatic reversible immobilization and labeling of proteins without prior purification. J. Am. Chem. Soc. *134*, 8455–8467.

127. Zhang, Y., Blanden, M.J., Sudheer, C., Gangopadhyay, S.A., Rashidian, M., Hougland, J.L., and Distefano, M.D. (2015). Simultaneous site-specific dual protein labeling using protein prenyltransferases. Bioconjug. Chem. *26*, 2542–2553.

128. Devadas, B., Lu, T., Katoh, A., Kishore, N.S., Wade, A.C., Mehta, P.P., Rudnick, D.A., Bryant, M.L., Adams, S.P., and Li, Q. (1992). Substrate specificity of *Saccharomyces cerevisiae* myristoyl-CoA: protein N-myristoyltransferase. Analysis of fatty acid analogs containing carbonyl groups, nitrogen heteroatoms, and nitrogen heterocycles in an in vitro enzyme assay and subsequent identifi. J. Biol. Chem. *267*, 7224–7239.

129. Heal, W.P., Wright, M.H., Thinon, E., and Tate, E.W. (2012). Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. Nat. Protoc. *7*, 105–117.

130. Kulkarni, C., Kinzer-Ursem, T.L., and Tirrell, D.A. (2013). Selective functionalization of the protein N terminus with N-myristoyl transferase for bioconjugation in cell lysate. ChemBioChem *14*, 1958–1962.

131. Van Valkenburgh, H.A., and Kahn, R.A. (2002). Coexpression of proteins with methionine aminopeptidase and/or N-myristoyltransferase in *Escherichia coli* to increase acylation and homogeneity of protein preparations. Methods Enzymol. *344*, 186–193.

132. Luginbuhl, K.M., Mozhdehi, D., Dzuricky, M., Yousefpour, P., Huang, F.C., Mayne, N.R., Buehne, K.L., and Chilkoti, A. (2017). Recombinant synthesis of hybrid lipid-peptide polymer fusions that self-assemble and encapsulate hydrophobic drugs. Angew. Chem. Int. Ed. *56*, 13979–13984.

133. Mozhdehi, D., Luginbuhl, K.M., Simon, J.R., Dzuricky, M., Berger, R., Varol, H.S., Huang, F.C., Buehne, K.L., Mayne, N.R., Weitzhandler, I., et al. (2018). Genetically encoded lipid-polypeptide hybrid biomaterials that exhibit temperature-triggered hierarchical self-assembly. Nat. Chem. *10*, 496–505.

134. Liu, L., Shadish, J.A., Arakawa, C.K., Shi, K., Davis, J., and DeForest, C.A. (2018). Cyclic stiffness modulation of cell-laden protein-polymer hydrogels in response to user-specified stimuli including light. Adv. Biosyst. *2*, 1800240.

135. Shadish, J.A., Strange, A.C., and DeForest, C.A. (2019). Genetically encoded photocleavable linkers for patterned protein release from biomaterials. J. Am. Chem. Soc. *141*, 15619–15625.

136. Zhang, W., Lohman, A.W., Zhuravlova, Y., Lu, X., Wiens, M.D., Hoi, H., Yaganoglu, S., Mohr, M.A., Kitova, E.N., Klassen, J.S., et al. (2017). Optogenetic control with a photocleavable protein, PhoCl. Nat. Methods *14*, 391–394.

137. Yin, J., Straight, P.D., McLoughlin, S.M., Zhou, Z., Lin, A.J., Golan, D.E., Kelleher, N.L., Kolter, R., and Walsh, C.T. (2005). Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. Proc. Natl. Acad. Sci. U S A *102*, 15815–15820.

138. Wong, L.S., Thirlway, J., and Micklefield, J. (2008). Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase. J. Am. Chem. Soc. *130*, 12456–12464.

139. Zhao, B., Zhang, K., Bhuripanyo, K., Wang, Y., Zhou, H., Zhang, M., and Yin, J. (2015). Phage selection assisted by Sfp

phosphopantetheinyl transferase-catalyzed site-specific protein labeling. Methods Mol. Biol. *1266*, 161–170.

140. Ott, W., Durner, E., and Gaub, H.E. (2018). Enzyme-mediated, site-specific protein coupling strategies for surface-based binding assays. Angew. Chem. Int. Ed. *57*, 12666–12669.

141. Ersfeld, K., Wehland, J., Plessmann, U., Dodemont, H., Gerke, V., and Weber, K. (1993). Characterization of the tubulin-tyrosine ligase. J. Cell Biol. *120*, 725–732.

142. Janke, C. (2014). The tubulin code: molecular components, readout mechanisms, and functions. J. Cell Biol. *206*, 461–472.

143. Schumacher, D., Helma, J., Mann, F.A., Pichler, G., Natale, F., Krause, E., Cardoso, M.C., Hackenberger, C.P.R., and Leonhardt, H. (2015). Versatile and efficient site-specific protein functionalization by tubulin tyrosine ligase. Angew. Chem. Int. Ed. *54*, 13787–13791.

144. Schumacher, D., Lemke, O., Helma, J., Gerszonowicz, L., Waller, V., Stoschek, T., Durkin, P.M., Budisa, N., Leonhardt, H., Keller, B.G., and Hackenberger, C.P.R. (2017). Broad substrate tolerance of tubulin tyrosine ligase enables one-step site-specific enzymatic protein labeling. Chem. Sci. *8*, 3471–3478.

145. Uttamapinant, C., White, K.A., Baruah, H., Thompson, S., Fernández-Suárez, M., Puthenveetil, S., and Ting, A.Y. (2010). A fluorophore ligase for site-specific protein labeling inside living cells. Proc. Natl. Acad. Sci. U S A *107*, 10914–10919.

146. Plaks, J.G., Falatach, R., Kastantin, M., Berberich, J.A., and Kaar, J.L. (2015). Multisite clickable modification of proteins using lipoic acid ligase. Bioconjug. Chem. *26*, 1104–1112.

147. Best, M., Degen, A., Baalmann, M., Schmidt, T.T., and Wombacher, R. (2015). Two-step protein labeling by using lipoic acid ligase with norbornene substrates and subsequent inverse-electron demand diels-alder reaction. ChemBioChem *16*, 1158–1162.

148. Baalmann, M., Best, M., and Wombacher, R. (2018). Site-specific protein labeling utilizing lipoic acid ligase (LplA) and bioorthogonal inverse electron demand diels-alder reaction. Methods Mol. Biol. *1728*, 365–387.

149. Dierks, T., Schmidt, B., Borissenko, L.V., Peng, J., Preusser, A., Mariappan, M., and von Figura, K. (2003). Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. Cell *113*, 435–444.

150. Kölmel, D.K., and Kool, E.T. (2017). Oximes and hydrazones in bioconjugation: mechanism and catalysis. Chem. Rev. *117*, 10358–10376.

151. Farahani, P.E., Adelmund, S.M., Shadish, J.A., and DeForest, C.A. (2017). Photomediated oxime ligation as a bioorthogonal tool for spatiotemporally-controlled hydrogel formation and modification. J. Mater. Chem. B *5*, 4435–4442.

152. Carlson, B.L., Ballister, E.R., Skordalakes, E., King, D.S., Breidenbach, M.A., Gilmore, S.A., Berger, J.M., and Bertozzi, C.R. (2008). Function and structure of a prokaryotic formylglycine-generating enzyme. J. Biol. Chem. *283*, 20117–20125.

153. Jian, H., Wang, Y., Bai, Y., Li, R., and Gao, R. (2016). Site-specific, covalent immobilization of dehalogenase ST2570 catalyzed by formylglycine-generating enzymes and its application in batch and semi-continuous flow reactors. Molecules *21*, https://doi.org/10.3390/molecules21070895.

154. Nguyen, G.K.T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J.P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat. Chem. Biol. *10*, 732–738.

155. Nguyen, G.K.T., Kam, A., Loo, S., Jansson, A.E., Pan, L.X., and Tam, J.P. (2015). Butelase 1: a versatile ligase for peptide and protein macrocyclization. J. Am. Chem. Soc. *137*, 15398–15401.

156. Joo, S.H. (2012). Cyclic peptides as therapeutic agents and biochemical tools. Biomol. Ther. (Seoul) *20*, 19–26.

157. Strumillo, M., and Beltrao, P. (2015). Towards the computational design of protein post-translational regulation. Bioorg. Med. Chem. *23*, 2877–2882.

158. Narayanam, M.K., Liang, Y., Houk, K.N., and Murphy, J.M. (2016). Discovery of new mutually orthogonal bioorthogonal cycloaddition pairs through computational screening. Chem. Sci. *7*, 1257–1261.

159. Chen, J.-Q., Chen, H.-Y., Dai, W., Lv, Q.-J., and Chen, C.Y.-C. (2019). Artificial intelligence approach to find lead compounds for treating tumors. J. Phys. Chem. Lett. *10*, 4382–4400.

160. Kopecek, J. (2007). Hydrogel biomaterials: a smart future? Biomaterials *28*, 5185–5192.

161. Ehrick, J.D., Deo, S.K., Browning, T.W., Bachas, L.G., Madou, M.J., and Daunert, S. (2005). Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. Nat. Mater. *4*, 298–302.

162. Zhou, X.X., Chung, H.K., Lam, A.J., and Lin, M.Z. (2012). Optical control of protein activity by fluorescent protein domains. Science *338*, 810–814.

163. Levskaya, A., Weiner, O.D., Lim, W.A., and Voigt, C.A. (2009). Spatiotemporal control of cell signalling using A light-switchable protein interaction. Nature *461*, 997–1001.

164. Wang, H., Vilela, M., Winkler, A., Tarnawski, M., Schlichting, I., Yumerefendi, H., Kuhlman, B., Liu, R., Danuser, G., and Hahn, K.M. (2016). LOVTRAP: an optogenetic system for photoinduced protein dissociation. Nat. Methods *13*, 755–758.