Chemical Modifications of Proteins and their Applications in Metalloenzyme Studies

Nathchar Naowarojna, † Ronghai Cheng, † Juan Lopez, Christina Wong, Lu Qiao, and Pinghua Liu*

Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States;
*Correspondence: pinghua@bu.edu
†Contributed equally to the work.

Abstract: Proteins chemical modifications are important tools for elucidating the chemical and biological functions of proteins. Several strategies have been developed to implement these modifications, including enzymatic tailoring reactions, unnatural amino acid incorporation using the expanded genetic code, and recognition-driven transformations. These technologies have been applied in metalloenzyme studies, specifically in dissecting their mechanisms, improving the activities, and creating artificial enzymes with non-natural activities. Herein, we summarize some of the recent efforts in these areas, along with a few metalloenzyme case studies.

Keywords: metalloenzymes, biorthogonal reactions, protein engineering, non-canonical amino acids, biocatalysis.

Proteins perform a remarkable range of catalytic, structural, and regulatory functions using the 20 canonical amino acids (cAAs) as their building blocks. The functions of proteins are further diversified to a level well beyond those possessed by the 20 cAAs through enzymatic-driven post-translational modifications (PTMs), such as methylations, sulfations, phosphorylations, and glycosylations. Moreover, chemical modifications of a protein of interest (POI) with special probes enable functional characterizations of the POI, including dynamics, localization, and protein-protein interactions [1, 2]. In pharmaceutical settings, the modification of protein therapeutics with polyethylene glycol (PEG) enhances their stability and circulation half-time [3]. Due to these reasons, a wide variety of protein modification methodologies have been explored to acquire exquisite control of these macromolecules. In this review, some of the recent strategies for protein modifications are presented and several metalloenzymes will be discussed as case studies in the areas of mechanistic studies and enzyme engineering.

1. Chemical modifications of proteins

Protein chemical modification approaches can be roughly classified into three categories: 1) modifications via the reactivities of cAAs; 2) ribosomal mediated incorporation of noncanonical amino acids (ncAAs); 3) modifications via affinity-driven ligand-directed

reactions. This section presents the recent progress in each of these approaches. The reactions discussed in this section are summarized in **Error! Reference source not found.**.

1.1 Chemical modifications of canonical amino acids (cAAs)

Amino acid chemical modifications offer straightforward approaches to modify proteins. However, the available functional groups in cAAs are limited. Both the reactivity and relative abundance of these residues play crucial roles in determining the selectivity of the chemical modifications. Reactions targeting highly abundant residues such as Lys often result in a mixture of modified proteins; nevertheless, significant progress has been made to obtain site-selective Lys modifications [1, 2]. On the other hand, low abundance residues such as Cys and Tyr might offer an opportunity for site-selective protein modifications. Many of these reactions can be performed within living cells as part of routine experiments and methods applicable for studies in multicellular organisms are currently of great interest in this field.

1.1.1. Lysine modifications

Among the 20 cAAs, Lys is an attractive choice for modifications due to both its high natural abundance and the number of biocompatible chemical reactions with its nucleophilic primary amine. Lys can be modified using electrophiles, such as squaric acids, isothiocyanates, sulfonyl chlorides, activated esters (including *N*-hydroxysuccinimide-based [NHS]), and 2-amino-2-methoxyethyl (**Error! Reference source not found.**) [1, 4, 5]. Due to its high natural abundance, Lys modifications are often useful in cases where site-selectivity is deemed not to be crucial or where multiple modification sites are desired. For instance, NHS-based *ortho*-nitrobenzyl esters and aldehydes are used to modify vitronectin, an extracellular matrix glycoprotein, in photopatternable hydrogels. This system provides simplified environments for studying cellular responses to physiological cues [6]. By reversibly patterning the presentation of vitronectin on the hydrogel, spatial and temporal control over osteogenic stem cell differentiation are achieved.

Due to its high natural abundance, site-specific modifications of Lys residues are challenging. In recent years, methods for achieving site-selectivity have emerged. In one case, by taking advantage of the differential reactivities of Lys residues in different local microenvironments, Matos *et al.* computationally designed sulfonyl acrylates which afforded chemo- and regioselective modifications [7]. Using the designed sulfonyl acrylates, a single Lys residue in several proteins, including the therapeutic antibody Trastuzumab, was modified without disturbing the structure and function of the POIs. With this method, an improved site-selectivity is achieved because the reaction favors Lys residues with the lowest pK_a in the POIs (Lys33 of lysozyme with a pK_a of 9.5, Lys300 of annexin V with a pK_a of 10.3, and Lys100 of C2Am with a pK_a of 10.1). However, this strategy might not be applicable in some POIs where the Lys with the lowest pK_a is inaccessible. In another case, in order to prepare an antibody-drug

conjugate (ADC), Rader *et al.* site-specifically conjugated β-lactam derivatives to a Lys residue (pK_a ~6) in the h38C2 antibody via aldol and retro-aldol reactions [8].

1.1.2. Protein's *N*-terminal modifications

In addition to Lys residues, the *N*-terminal primary amine of a protein also displays unique reactivity. This site-selective modification has been applied in native chemical ligation (NCL) reactions, which are commonly used in peptide synthesis (**Error! Reference source not found.**a) [9, 10]. In NCL reactions, the thiolate of the *N*-terminal Cys residue of one peptide reacts with another *C*-terminal thioester containing peptide. The resulting thioester intermediate undergoes an acyl shift to form a native peptide bond (**Error! Reference source not found.**a). In a second method, a recombinant protein can be conjugated to a synthetic peptide containing the desired modification using the expressed protein ligation (EPL) method (**Error! Reference source not found.**b) [11-13]. This approach enables site-selective modification when the desired site is close to the *N*- or *C*-terminus of a POI. For example, the *N*-terminal ultrafast split intein was genetically fused to the *C*-terminal of histone H2B, which then reacted with a synthetic *C* intein fragment bearing a pre-ubiquitinated Lys120, resulting in a ubiquitinated H2B (H2B-K120Ub) [14]. Through this approach, the semisynthetic H2B-K120Ub is operative and can induce chromatin signaling in isolated nuclei. Additionally, the *N*-terminus has been used to generate reactive ketones though a pyridoxal-5'-phosphate (PLP) catalyzed transamination reaction; however, this reaction is not compatible with some *N*-terminal residues, such as Lys and Gln due to side reactions [15-17]. The resulting ketone or aldehyde can be used for the oxime/hydrazine ligation. For instance, the *N*-terminus of streptavidin was modified to an α-ketoamide allowing its immobilization onto micropatterned surfaces through oxime ligation [18]. This approach immobilizes the proteins in a specific orientation, which is critical for maintaining their bioactivity. Moreover, strategies have been developed to modify the *N*-termini with small molecules, such as a reductive alkylation using an aldehyde-containing 2-pyridinecarbaldehyde (2-PCA) [19]. This *N*-alkylation reaction has been applied to install a secondary amine on human insulin. The resulting secondary amine preserves the positive charge on the *N*-terminus of human insulin, and the bioactivity of the modified insulin is comparable to that of unmodified insulin. Despite these advances, there remain many challenges. For example, the *N*-termini of many proteins may not be accessible for modification or might be essential for their biological/catalytic activities. In addition, *N*-terminal modifications also require careful control over the reaction conditions.

1.1.3. Cysteine modifications

Cys is another residue of choice for chemical modifications due to the high nucleophilicity of its thiol and its low natural abundance in proteins, all of which favor single site modification on the POIs [1, 5, 20]. Under well-controlled pH conditions, selective modification

of a Cys residue over other nucleophilic residues such as Lys and His can be obtained [21]. The thiol of Cys can be modified using electrophilic α-halocarbonyls (iodoacetyl, bromoacetyl, or chloroacetyl) [22, 23] and Michael acceptors (maleimides[24, 25] or vinyl sulfones, **Error! Reference source not found.**a) [26]. Maleimides has been widely used to synthesize ADCs, including the FDA-approved ADCs—brentuximab vedotin, trastuzumab emtansine, and certolizumab pegol [2]. However, the thioether of maleimide conjugates can undergo cleavage via the thiol exchange or hydrolysis-driven reactions. The resulting products from the thiol exchange lead to premature loss of the drug's efficacy and an increased toxicity *in vivo* [27]. These issues can be addressed by minimizing the thiol exchange reaction or by converting the maleimide thio-adducts to stable ring-open conjugates. The latter has been accomplished through the development of new self-hydrolyzing maleimides, which exhibit superior pharmacokinetic properties [28]. In recent efforts, new generations of thiol-targeting modification reagents have been developed, including electron-deficient alkynes [29], 3-arylpropiolonitrile [30, 31], allenamides [32], and the thiol-yne reactions [33]. For example, Wagner *et al.* reported an amine-to-thiol coupling reagent, sodium 4-((4-(cyanoethynyl)benzoyl)oxy)-2,3,5,6-tetrafluorobenzenesulfonate (CBTF), which contains an arylpropionitrile functional group instead of a maleimide (**Error! Reference source not found.**b). The resulting conjugates exhibit superior plasma stability compared to the maleimide conjugates [31].

In cases where the POIs lack a thiol functional group, disulfide-targeting reagents can be used instead. Multiple reagents including bissulfones, allyl sulfones, alkynes, and 3,4-disubstituted maleimides have been developed to site-selectively modify the disulfides of proteins as summarized in recent reviews [5, 34]. For instance, using 2,3-dibromo maleimide with a C-2 (glycine derived) linker, Doxorubicin (DOX)-antibody conjugates were produced through a bis-alkylation reaction (**Error! Reference source not found.**c). This approach results in homogenous ADCs with enhanced pharmacological properties [35]. Oxetane, a four-membered ring with an oxygen atom, has also be used to modify protein's disulfides via a site-selective bis-alkylation reaction [36, 37]. In one of the reports, oxetane was installed onto a genetically detoxified diphtheria toxin (CRM_{197} protein) and the resulting modified protein has increase immunogenicity *in vivo* [36].

1.1.4. Aromatic residues modifications

In addition to Cys, the relatively low natural abundance of aromatic residues, including His, Tyr, Trp, and Phe, offer alternative targets for site-specific modifications. However, obtaining a site-specific modification for one aromatic residue over another is still challenging. The reactivity of the ionizable side chain of Tyr is dependent on its protonation state, which allows the reactivity of Tyr to be modulated by controlling the pH of the reactions. Under acidic conditions, the aromatic ε-carbons adjacent to the hydroxyl group may undergo diazonium couplings (**Error! Reference source not found.**a) [38]. In a recent report, the salmon's

calcitonin was conjugated to linear monomethoxy PEG using this approach. The resulting conjugates maintain the ability to reduce the concentration of calcium ions in the plasma. In conditions where the pH approaches the pK_a of the phenol side chain of the Tyr residue (pK_a of ~10), alkylation or acylation reactions of the oxygen can occur [39-41]. The selective Tyr alkylation is accomplished through the use of π-allylpalladium complexes (**Error! Reference source not found.**b) [39]. Using this method, a novel polarity-sensitive fluorescent probe was installed onto Tyr108 of a bovine Cu/Zn superoxide dismutase (SOD) which allows the conformational changes of the Tyr-containing domain to be monitored. This example might serve as a general approach for studying protein conformational changes if the domain of interest has a Tyr residue.

In the case of Trp, modifications of its side chain via metal-catalyzed reactions have been developed (**Error! Reference source not found.**), including alkynylations [42], C-H arylations [43], and photoinduced cycloadditions of tetrazoles [44]. However, these approaches often involve non-biocompatible reaction conditions such as the use of organic solvents and high temperatures. An alternative strategy for selective modification of Trp consists of an organic radical 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxy (*keto*-ABNO) coupled with $NaNO_2$ under ambient conditions (**Error! Reference source not found.**) [45]. Kanai *et al.* accomplished Trp-selective modification using this approach on myoglobin, lysozyme, BSA, and the β_2-microglobulin antibody. This study opens new avenues for Trp modifications without requiring toxic metals under biocompatible conditions, paving the way for functional modulation of therapeutic targets such as antibodies. In addition to the aforementioned cases, efforts for site-specific modifications of other proteinogenic residues are well summarized in a recent review [46].

1.2. Protein modifications using ncAAs

Protein modifications using chemical reactions are powerful approaches to selectively install the desired chemical handles on POIs. Manipulation of the protein translation machinery for direct incorporation of ncAAs is an alternative approach, which enables the introduction of novel functionalities and reactivities beyond those offered by the 20 cAAs. For example, the incorporation of ncAAs in proteins has been widely applied in biocatalysis to enhance the activity and selectivity of enzymes, and even to obtain novel catalytic reactions that are naturally unavailable [47-49]. This section presents some of the recent advances in the use of the genetic code expansion for the incorporation of ncAAs into POIs and the further chemical modifications of these ncAAs.

1.2.1. Genetic code expansion

The natural genetic code is comprised of 64 codons encoded by four nucleobases. Among them, three serve as stop codons, while the remaining 61 codons are recognized by transfer

RNAs (tRNAs), which are charged by cognate tRNA synthetases (aaRSs) using one of the cAAs. All organisms encode the same 20 amino acids (AAs) with the exceptions of pyrrolysine [50] and selenocysteine [51]. The ability to employ and modify the existing protein translation machinery to incorporate ncAAs into proteins has opened new avenues for protein modifications. One of the strategies relies on the promiscuity of some tRNAs that can be charged with ncAAs that are structurally analogous to their cognate amino acids (**Error! Reference source not found.**a) [52]. In this method, an auxotrophic bacterial host cannot produce the cAA and relies on the exogenously supplied ncAA, resulting in a global replacement in the proteome. The global replacement has proved to be useful in substituting Met with Se-Met to incorporate heavy atoms into POIs for crystallographic phasing [53]. This strategy is straightforward and does not require genetic manipulation of the protein translation components. It relies on the polyspecificity of the aaRS-tRNA pair, meaning that only ncAAs with structural similarities to the cognate amino acids can be incorporated. Nevertheless, the global incorporation of ncAAs might alter protein folding, which could result in a perturbation of the activity and/or stability of the POIs.

A powerful alternative approach is the genetic code expansion to site-specifically incorporate ncAAs into POIs (**Error! Reference source not found.**b). In this method, the amber stop codon (UAG) is used to encode the ncAA of interest because of its low usage in *Escherichia coli* (~7-8%). The amber codon is recognized by an engineered aaRS-tRNA pair for the ncAA of interest. The aaRS-tRNA pair must also be orthogonal, i.e., not interfering with the endogenous translation system (**Error! Reference source not found.**b). For example, the tyrosyl-tRNA synthetase TyrRS-$tRNA_{CUA}$ pair from *Methanocaldococcus jannaschii* is orthogonal in *E. coli* and other bacteria; the TyrRS-$tRNA_{CUA}$ and LeuRS-$tRNA_{CUA}$ pairs from *E. coli* are orthogonal in eukaryotic cells; the pyrrolysyl-tRNA synthetase PylRS-$tRNA_{CUA}$ pairs from *Methanosarcina barkeri* and *Methanosarcina mazei* are orthogonal in both bacteria and eukaryotic cells [48, 54, 55]. Site-specifically modified POIs can be obtained, but the production yield is normally limited by the expression level of the exogenous aaRS-tRNA pairs or the presence of release factor 1 (RF-1), which recognizes the UAG triplet and terminates translation. Recently, an *E. coli* host has been created by removing RF-1 from the *E. coli* genome. Additionally, 95 out of the 273 amber stop codons were replaced with other more frequently used stop codons. After this engineering, the growth defects of the *E. coli* host are minimized when it is used to produce ncAA containing proteins [56, 57]. Most importantly, the ncAA incorporation efficiency is >98% in this engineered host strain, allowing a scalable production of the target ncAA-containing protein.

1.2.2 Next-generation genetic code expansion

To date, more than 200 ncAAs have been incorporated into POIs using the amber suppression method, thereby expanding the chemical functionalities and reactivities of proteins [48, 58]. Thus far, the vast majority of studies employing this technology is restricted to the

incorporation of a single ncAA into POIs. The ability to incorporate multiple ncAAs into a protein might offer new opportunities for advanced biophysical studies and the synthesis of enhanced protein-based therapeutics. To achieve such goals, the enhanced specificity and orthogonality of aaRS-tRNA pairs are essential. Orthogonal aaRS-tRNA pairs can charge multiple ncAAs during the translation process. Thus, to site-selectively incorporate multiple ncAAs, the orthogonal aaRS-tRNA pairs must be highly selective. Several engineering approaches have been developed to address these challenges. Liu *et al.* employed the phage-assisted continuous evolution (PACE) technology to evolve a PylRS variant with 45-fold higher catalytic efficiency compared to that of the wild-type. Through this technology, a TyrRS variant with improved selectivity toward p-iodo-L-phenylalanine was also obtained [59]. Engineering of the aaRS-tRNA pairs has proven useful in incorporating ncAAs at multiple sites. In a recent study, Chatterjee *et al.* simultaneously incorporated three distinct bioconjugation handles onto the GFP protein by employing three orthogonal pairs to decode three different codons. These pairs include the *E. coli*-derived tryptophanyl pair, the archaeal tyrosyl and pyrrolysyl pairs. This work allows facile multiple-site labeling of proteins without an exogenous chemical catalyst. However, the incorporation efficiency is only 2% of the wild-type yield [60, 61].

Besides improving the specificity of the aaRS-tRNA pair, evolving other translation machinery components has also been proven to be important for the efficient incorporation of ncAAs into multiple-sites of the POIs. To facilitate the incorporation of ncAAs with bulky side-chains or altered backbones, ribosome engineering has been attempted. There are two challenges in ribosome engineering. First, the activity and fidelity of ribosomes are essential for cellular survival. Second, the 23S rRNA subunit can freely exchange between the native and the orthogonal 16S rRNAs in the cell. To overcome the former issue, an orthogonal ribosome containing a modified anti Shine-Dalgarno (SD) sequence was introduced in *E. coli* (**Error! Reference source not found.**a). This modification allows the translation of the orthogonal genes containing the SD sequence that matches with the anti-SD of the orthogonal ribosome, which reduces the ability of the orthogonal ribosome to translate native mRNA transcripts [62, 63]. However, full orthogonality cannot be achieved due to the exchange between the native and orthogonal ribosomal subunits. To address this issue, Chin *et al.* reported an orthogonal ribosome with subunits linked through an optimized orthogonal RNA staple that can maintain activity comparable to that of the parent orthogonal ribosome with minimized endogenous subunit association (**Error! Reference source not found.**a). The peptidyltransferase center (PTC) of the O-d2d8 variant was further engineered to translate proline-rich sequences, a novel intrinsic ability that is lacking in native ribosomes [64].

Other exciting technologies for genetic code expansion extend beyond the canonical 64 codons. The first glimpse of quadruplet codon usage was reported from *Salmonella typhimurium,* in which a $tRNA^{Gly}$ containing the frameshifted CCCC anticodon was observed [65]. However, the efficiency of natural ribosomes for decoding quadruplet codons is extremely low. To

overcome this issue, Chin *et al.* engineered translation components that make use of quadruplet codons, instead of the traditional triplet ones (**Error! Reference source not found.**b) [66]. In this work, Chin and co-workers evolved an orthogonal ribosome (ribo-Q1) that can translate a quadruplet codon. They have also obtained a variant Seryl-tRNA synthetase/tRNA pair that can incorporate ncAAs at the quadruplet codon. A combination of ribo-Q1 and SerRS was then used to incorporate azide and alkyne-containing amino acids to a calcium-binding messenger protein calmodulin. A combination of these efforts (amber codon, quadruplet codon, and translational machinery engineering) led to the incorporation of nearly 200 ncAAs into POIs. Another important piece of work is a semisynthetic bacteria with expanded genetic codes [67]. In this study, Romesberg *et al.* introduced an unnatural base pair (UBP) comprised of two deoxynucleoside triphosphates (dNaM and dTPT3). Theoretically, the addition of these two UBPs enables the incorporation of 152 new unnatural codons. In this study, the authors identified nine unnatural codons that can be used to incorporate ncAAs efficiently. Among these, three codons are orthogonal and could be used for three distinct ncAAs, which results in the first 67-codon semisynthetic organism. This work opens a new avenue for multiple-site incorporation of ncAAs into POIs.

1.2.3 Chemical modifications of the incorporated ncAAs

The genetic code expansion allows for the site-specific incorporation of a variety of ncAAs into a POI, providing useful chemical handles for biorthogonal reactions [1, 68]. Several functional groups including azides, alkynes, alkenes, and tetrazines, can be incorporated through the amber codon suppression system, which are subsequently modified by appropriate reagents. For instance, the azides or alkynes of incorporated ncAAs can undergo a copper-catalyzed azide-alkyne cycloaddition (CuAAC), which is termed the "CuAAC Click Chemistry reaction" (**Error! Reference source not found.**a) [69-71]. Due to its high specificity and reasonable reaction rate, the CuAAC reaction has been applied in the analysis of the cellular proteome via biorthogonal non-canonical amino acid tagging (BONCAT) [72, 73]. By replacing Met in the cell culture media with an azide- or alkyne-containing ncAA such as azidohomoalanine (AHA) or homopropargylglycine (HPG), the newly synthesized proteins are marked with azide or alkyne functionality, and are distinguishable from the pool of preexisting proteins (**Error! Reference source not found.**b-c). Subsequently, the AHA- or HPG-containing proteins are covalently attached to an affinity tag via Click Chemistry. The newly synthesized and affinity-tagged proteins can be analyzed by conventional biochemical studies or high-resolution mass spectrometry (**Error! Reference source not found.**c). In a recent study, an alkyne N_{ε}-(propargyloxycarbonyl)-L-lysine (AlkK) was incorporated into the proteome of murine tissue slices and the brains of live mice expressing an orthogonal pyrrolysyl pair introduced by viral injection [73]. The alkyne-bearing proteins are covalently linked to the streptavidin beads via diazobenzene linkers, which are then cleaved to provide the labeled proteome for analysis by

tandem mass spectrometry. This method has been applied to characterize newly synthesized proteins after the pharmacological treatments of a particular type of cell.

Despite the promising applications, the CuAAC reaction involves the use of toxic metal catalysts and exogenous ligands. The strain-promoted azide-alkyne cycloaddition (SPAAC) has emerged as a powerful copper-free Click Chemistry for the modification of POIs (**Error! Reference source not found.**d) [1, 74]. Several strained alkynes have been developed, including difluorinated cyclooctyne, dibenzoccyclooctynol, and biarylazacyclooctynone. Some of these functional groups can already be incorporated into POIs via the amber codon suppression system [75-78]. The copper-free Click Chemistry has been implemented in a variety of applications, which are well summarized in a recent review [79]. Another emerging strategy for ncAA modifications is the photo-click chemistry, which involves a photo-inducible dipolar [3+2]-cycloaddition reaction between an alkenyl-ncAA and a nitrile imine. This reaction results in the formation of fluorescent adducts, allowing not only for the *in vitro* modification of isolated proteins, but also the visualization of proteins in living cells [80-82] . These examples illuminate the general applicability of the biorthogonal reactions in the site-specific modifications of POIs under complex biological environments, including living cells and multicellular organisms with minimal disruption of the biological functions and activities of the POIs.

1.3. Ligand-directed modifications of proteins

In some cases, proteins possess high binding affinities for their respective binding partners. The interactions between proteins and their binders allow for affinity-driven site-specific modifications. A variety of recognition-driven chemical modification strategies have been developed, including ligand-directed (LD) modifications [1, 2, 83, 84]. The LD chemical modifications of proteins have been developed for endogenous protein labeling by harnessing the specific interactions between the POIs and small molecules [85]. In this approach, the labeling reagent contains an affinity ligand, a reporter tag, and a reactive moiety (**Error! Reference source not found.**a). The binding of the ligand to the POI places the reactive unit of the labeling reagent in proximity to residues close to the ligand binding site, which enhances the selective modification. The LD approach has been applied in the labeling carbonic anhydrase (CA) [85]. In this study, the specific ligand, benzenesulfonaide, is conjugated to a synthetic probe, which includes a fluorophore via the phenylsulfonate linkage (**Error! Reference source not found.**a). Upon binding to the protein, a nucleophilic side chain in the POI reacts with the electrophilic phenylsulfonate ester group. This reaction leads to the labeling of the POI, while at the same time, the ligand may also dissociate. Therefore, the functions of the POI are not disrupted after the chemical modification. In this study, it was also demonstrated that the LD approach used in the labeling of CA could be applied to both *in vitro* and *in vivo* studies.

Non-covalent and reversible ligand binders are often useful in many applications. However, in some cases, the ability to form a stable covalent adduct might be important, such as in the development of inhibitors. In a recent study, Hamachi *et al.* reported a new ligand-directed *N*-acyl-*N*-alkyl sulfonamide (LDNASA)-based covalent inhibitor, which targets chaperone Hsp90 (**Error! Reference source not found.**b)[86]. LDNASA enabled the site-specific modification of Lys58, which is a residue near the ligand binding site of the chaperone Hsp90. This reaction leads to a covalent adduct between Hsp90 and its inhibitor (Figure 9b). Treatment of cancer cells with an LDNASA-linked inhibitor results in irreversible covalent modification of Hsp90 at Lys58 by the inhibitor and reduces Hsp90's molecular chaperone activity in cancer cells.

2. Utilizing protein modifications in the studies of metalloenzymes

Chemical modifications of proteins enable the study of their biological functions and regulation in cells. Particularly, chemical modifications and the expansion of the genetic code enables the incorporation of structurally and chemically diverse ncAAs into POIs. These technologies have also opened new avenues for investigating enzymatic reaction mechanisms and the engineering of enzymes with new functions. One-third of all naturally occurring enzymes are metalloenzymes [87, 88]. Among these metalloenzymes, non-heme iron (NHFe) and heme enzymes catalyze a remarkable range of chemical transformations, including hydroxylation, endoperoxidation, and desaturation at the expense of molecular oxygen as the final electron acceptor [87, 89, 90]. This section presents selected case studies, from recent years, of NHFes, heme, and Cu-containing enzymes where ncAAs are used as novel tools. The cases discussed in this section are summarized in **Error! Reference source not found.**.

2.1. Probing the role of residues in NHFe enzymes' active sites

Post-translational modifications (PTMs) are required for proper protein functionalities, including subcellular localization, protein-protein interactions, and catalysis. For instance, a protein-derived thioether Cys-Tyr crosslink enhances the cysteine oxidation activity of the NHFe cysteine dioxygenase enzyme (CDO, **Error! Reference source not found.**a-10b) [91]. In CDO, the octahedral ferrous iron is coordinated by a 3-His motif (**Error! Reference source not found.**b). The Cys-Tyr crosslink biogenesis in CDO is an autocatalytic oxidation reaction, in which the oxygen-activated iron center oxidizes the residues at the active site (Cys and Tyr) rather than the substrate, L-Cys. Depending on the reaction conditions, it takes hundreds to thousands of turnovers to obtain fully mature CDOs, that can efficiently oxidize the L-Cys substrate [92]. Despite the important role that the Cys-Tyr crosslink plays in the catalytic activity, the mechanistic details of the biogenesis of this crosslink remain unclear, partly due to the challenges of obtaining a homogeneous population of either cross- or uncross-linked CDO. In a recent study, Liu and co-workers employed the amber codon suppression method to replace

Tyr157 with a halogen-substituted Tyr in human CDO [93]. The fluorine- and chlorine-containing CDO variants remained active, albeit with a lower activity (2-10% of the wild-type activity). Intriguingly, the Cys93-Tyr157 crosslink is also observed among these variants. Given the high bond dissociation energy, C-F bond cleavage is unexpected in the Cys-Tyr crosslink biogenesis (**Error! Reference source not found.**c). Subsequent mass spectrometric analysis and biochemical reactions of the difluoro-substituted variant confirmed the oxidative C-F bond cleavage, which might take place during the oxidation of Tyr157 by the iron center or the Cys93 radical. This study represents the first reported case of an oxidative C-F bond cleavage mediated by an NHFe enzyme.

To further understand the mechanism of the crosslinking reaction, the uncross-linked structure of F_2-Tyr CDO has been obtained, revealing that Cys93 assumes two conformations prior to the Cys-Tyr cofactor formation (**Error! Reference source not found.**d). The authors proposed that during the crosslink biogenesis, the sulfur atom of Cys93 rotates toward the iron center. This hypothesis is supported by a recently reported structure of the uncross-linked F_2-TyrCDO•L-Cys•NO ternary complex (**Error! Reference source not found.**e) [94]. This structure revealed an interaction between NO and one conformer of Cys93, suggesting that cysteine oxidation may occur prior to the crosslink formation. The oxidation of Cys93 by the iron center is further supported by a computational study which suggested that an iron-bound oxygen species oxidizes Cys93 as the first step of the crosslink biogenesis, instead of the Tyr157 oxidation, as previously proposed [92, 95]. The studies of the CDO cofactor biogenesis highlight the general applicability of site-specific modifications of POIs in mechanistic studies. Due to the important roles of Tyr in enzymatic catalysis, the ability to control the structural features and chemical properties (reduction potential and the pK_a) of this residue has been sought after [49, 88, 96, 97]. The non-canonical Tyrs, including o-substituted tyrosine and fluoro-substituted tyrosines, are structurally similar to natural tyrosine, but differ greatly in terms of their reduction potentials and pK_a (**Error! Reference source not found.**a). Several non-canonical Tyrs can be site-selectively incorporated into POIs via the amber-codon suppression approach [98-100]. This toolkit has proven useful in studying the bioactivity and functions of other metalloenzymes, including sulfoxide synthases in the biosynthesis of ergothioneine **(5)**, (**Error! Reference source not found.**b).

Ergothioneine is a potent antioxidant and has been proposed as a longevity vitamin [101, 102]. Due to its potential benefits, the biosynthesis of ergothioneine has received considerable interest. In the mycobacterial ergothioneine biosynthetic pathway, a NHFe sulfoxide synthase, EgtB, mediates the oxidative coupling between hercynine **(2)** and γ-glutamyl-cysteine (γ-Glu-Cys, **Error! Reference source not found.**b) [103]. Recently, the fungal biosynthesis of ergothioneine was also reported, in which a sulfoxide synthase, Egt1, catalyzes the C-S bond formation between hercynine **(2)** and L-Cys (**Error! Reference source not found.**b) [104, 105]. A similar reaction has also been observed in the biosynthesis of another potent antioxidant,

ovothiol A **(9)**. In the ovothiol biosynthesis, a sulfoxide synthase, OvoA, catalyzes the oxidative coupling between L-His **(1)** and L-Cys (**Error! Reference source not found.**b) [103, 106, 107]. These ergothioneine and ovothiol sulfoxide synthase reactions differ in terms of their substrate selectivity and their product's C-S bond regioselectivity. In EgtB- and Egt1-catalysis, the C-S bond is formed at the ε-carbon of hercynine's imidazole ring, while OvoA-catalysis incorporates the C-S bond at the δ-carbon of L-His' imidazole ring (**Error! Reference source not found.**b).

The structure of *Mycobacterium thermoresistible*'s EgtB has been reported, revealing a 3-His iron coordination site similar to that of CDO (**Error! Reference source not found.**a) [108]. When the active site Tyr377 was mutated to Phe, the sulfoxide synthase activity was significantly suppressed, and in this mutant, the oxidation of γ-Glu-Cys was the dominant activity. Based on this EgtB crystal structure, three different mechanistic models have been proposed based on calculations using the density function theory or the quantum mechanics/molecular mechanics (QM/MM) methods (**Error! Reference source not found.**c) [109-111]. These mechanistic models have led to questions regarding whether the reaction involves a sulfenic acid intermediate and what role the active site Tyr377 play in EgtB-catalysis. EgtB-catalysis is a four-electron oxidation process and molecular oxygen is the final electron acceptor. In **Error! Reference source not found.**c, two independent theoretical studies reported by Tian and Wei suggested that the sulfur's oxidation to sulfenic acid is the first half of the reaction (Path IA-IC), while another study by Faponle *et al.* proposed that the oxidative C-S bond formation is the first half of EgtB-catalysis (Path II) [109-111]. In these two mechanistic models (pathway I vs pathway II), different functions have been proposed for the active site Tyr377. In the Tian model, Tyr377 functions as a Lewis acid/base (Path I), (**Error! Reference source not found.**c). In the Faponle model, Tyr377 plays a key role in a proton-coupled electron transfer (PCET) process (Path II), **Error! Reference source not found.**c). Due to the similarities between the EgtB- and OvoA-catalyzed reactions, all of the mechanistic discussions regarding EgtB-catalysis are also applicable to OvoA-studies [96, 112, 113].

To unravel the mechanism of OvoA-catalysis, Liu and co-workers harnessed the unique chemical properties of the non-canonical Tyrs [96, 112]. Notably, the biochemical characterizations also revealed the cysteine oxidation activity in OvoA-catalysis [113]. The mechanistic models from the computational studies coupled with the biochemical results suggest that the sulfoxidation and cysteine oxidation activities might be two pathways branching out from a common intermediate in OvoA-catalysis. Due to this reason, the isotopically sensitive branching method was used to measure the kinetic isotope effect (KIE) of the OvoA reaction with the structural model of OvoA serving as a guide (Figure 12b). In this model, Tyr417 in OvoA is the EgtB's Tyr377 counterpart. The Tyr417 of OvoA was replaced with 2-amino-3-(4-hydroxy-3-(methylthio)-phenyl)propanoic acid (MtTyr) using the amber codon suppression system (**Error! Reference source not found.**a) [96]. Using [U–$^{2}H_5$]-histidine, the substrate KIE

of the $OvoA_{Y417MtTyr}$ variant is close to unity (1.08±0.01), which is comparable to that of the wild-type (substrate KIE of 1.01±0.02). This experimental result contradicts the model suggested by Wei *et al.*, in which they predicted a primary substrate deuterium KIE as high as 5.7 [110].

By replacing Tyr417 with MtTyr, the two OvoA activities (oxidative coupling between cysteine and histidine, and cysteine dioxygenase activity) can be modulated. The biochemical results reported by Liu and co-workers support the supposition that the two OvoA activities branch out from a common intermediate and that the active site Tyr417 plays some key roles in controlling the partitioning between these two pathways. However, MtTyr differs from the canonical Tyr in both reduction potential and pK_a. Whether Tyr417 plays a role in redox chemistry or functions as a Lewis acid/base in the OvoA reaction remains unclear. To overcome this issue, Liu and co-workers replaced Tyr417 with 3-methoxytyrosine (MeOTyr), which possesses a similar pK_a but a much lower reduction potential than that of canonical Tyr [112]. Using the kinetic isotope sensitive branching method, the deuterium substrate KIE of the $OvoA_{Y417MeOTyr}$ variant was found to be 0.86±0.03. This inverse KIE is significantly different from that observed in the wild-type OvoA enzyme, in which the KIE is close to unity. This difference might be due to a change in the rate-determining step of OvoA-catalysis as a result of replacing Tyr with MeOTyr. The inverse substrate deuterium KIE of $OvoA_{Y417MeOTyr}$ was further confirmed using an intermolecular competition method and the substrate KIE determined was 0.94±0.02. The inverse KIE implies that the rate-limiting step in MeOTyr-containing OvoA reaction is a step involving an sp^2 to sp^3 hybridization transition.

Among the proposed mechanisms, the model proposed by Faponle *et al.* is more consistent with the experimental results previously described [109]. In this computational model, an inverse deuterium KIE was predicated if the δ-hydrogen of the imidazole ring is replaced with deuterium. This inverse KIE was, indeed, observed in the $OvoA_{Y417MeOTyr}$ variant. Given the similarities between OvoA- and EgtB-catalysis, the OvoA mechanism will be discussed in the context of EgtB mechanistic models (**Error! Reference source not found.**c). In the model proposed by Faponle and co-worker, the oxygen binding and activation results in an Fe^{III}-superoxo species **(1-f)**. Then, the Tyr377 in EgtB participates in the PCET process through the active site's water network, resulting in the Fe^{III}-hydroperoxo species **(1-g)** along with a Tyr377-based radical species (**Error! Reference source not found.**c). Subsequent nucleophilic or radical attack by the sulfur atom of γ-Glu-Cys on the imidazole's sidechain of hercynine affords a thioether product. In this step, the hydrogen atom from the hydroperoxo species relays back to the Tyr377-based radical, regenerating Tyr377 and producing an Fe^{II}-superoxo species **(1-h)**. The Fe^{II}-superoxo species then abstract a hydrogen atom from the imidazole ring to regenerate the aromaticity and results in an Fe^{II}-hydroperoxo intermediate **(1-i)**, which undergoes oxidation to afford the sulfoxide product (**Error! Reference source not found.**c). Our MeOTyr-containing OvoA studies are consistent with the EgtB model suggested by Faponle *et al.* [109]. However, the OvoA structure is not available and it is possible that OvoA and EgtB might follow different

mechanistic models. Additional kinetic and spectroscopic studies are needed in order to uncover the mechanistic details of these sulfoxide synthases.

The incorporation of non-canonical Tyrs has been proven to be useful for the mechanistic characterization of the sulfoxide synthases in the ergothioneine and ovothiol biosyntheses. There exists another enzymatic system where the role of Tyr in catalysis is under heated debate. In the biosynthesis of a tremorgenic mycotoxin verruculogen **(12)**, an α-ketoglutarate-dependent (αKG) NHFe enzyme, FtmOx1, mediates the endoperoxidation of fumitremorgin B **(11),** (**Error! Reference source not found.**a) [114, 115]. Biochemical characterization under single-turnover reaction conditions in the absence of extra reductants (e.g., ascorbate) revealed that each cycle of FtmOx1-catalysis consumes two molecules of molecular oxygen and involves two distinct reactions: the endoperoxidation reaction to afford verruculogen **(12)** and the oxidation of verruculogen's C_{13}-hydroxyl group to form a keto-product **(13)** (**Error! Reference source not found.**a), [97]. In the reported crystal structure of the FtmOx1•αKG complex binary complex, αKG coordinates to the metallocenter through a distal-type configuration, and Tyr224 is adjacent to the putative oxygen binding site (**Error! Reference source not found.**b) [97]. Structural, biochemical, and spectroscopic characterizations led to a mechanistic model involving the Tyr224 radical as one of the key species in FtmOx1-catalysis (**Error! Reference source not found.**c).

However, in a recent report, Bollinger and co-workers proposed that an alternative tyrosine residue (Tyr68) might be the tyrosyl radical site (**Error! Reference source not found.**) [116]. The authors reported that in the reaction catalyzed by the $FtmOx1_{Y68F}$ variant, a new major, uncharacterized product is formed. In FtmOx1-catalysis, under single-turnover conditions and in the absence of other reductants, the keto-product **(13)** is the dominant product. In this Tyr68 radical model (Figure 14), there is no discussion of how Tyr68 might be involved in the formation of keto-product **(13)**. In our hands, the dominant product formed from the $FtmOx1_{Y68F}$ mutant is the exact same as the wild type FtmOx1 (Liu *et al.*, unpublished results). Therefore, Tyr224 remains the best candidate for the proposed tyrosyl radical in FtmOx1-catalysis. To provide further evidence to differentiate between the two proposed mechanistic models, the use of non-canonical Tyr will be beneficial in characterizing the catalytic roles of the two Tyr residues in FtmOx1-catalysis; These are on-going efforts in our laboratory.

2.2. Heme enzymes engineering

Another common type of iron-dependent enzymes are the heme-containing enzymes, in which the iron center is coordinated to a porphyrin as a tetradentate ligand. In addition to heme cofactors, these enzymes often possess additional metals that catalyze a diverse range of reactions, such as copper (Cu_B) in heme copper oxidase (HCO) and iron (Fe_B) in nitric oxide reductase (NOR) [117-121]. The mechanistic studies of heme enzymes are limited by the difficulties of working with these membrane-bound proteins. Designing artificial heme enzymes

is a powerful tool not only for facilitating in-depth mechanistic studies, but also for developing novel biocatalysts. Both rational enzyme design and directed evolution methods have been applied for heme-enzymes engineering where metal-ligands and the secondary coordination shell residues are modulated [122-125]. The cytochrome *c* oxidase (C*c*O) is a membrane-bound protein from the HCO family. It catalyzes the reduction of O_2 to water without the generation of toxic reactive oxygen species (ROS). In this reaction, the catalytic functions of the additional Cu_B site remain unclear [126]. Due to the difficulty of obtaining large amounts of membrane-bound C*c*O, an artificial model for this enzyme has been generated through enzyme engineering of the heme-containing oxygen carrier protein, myoglobin, from sperm whale *Physeter macrocephalus* (swMb), (**Error! Reference source not found.**a). Based on the structural comparison between swMb and the native bovine C*c*O, residues His64, Leu29, and Phe34 from the active site of swMb were selected to create the Cu_B binding site [126]. Both Leu29 and Phe34 were mutated to His to create the swMb L29H F43H variant, which is referred to as Cu_BMb. This engineered Cu_BMb variant exhibited biophysical properties similar to those of a heme-copper center from the HCO family. In addition, spectroscopic studies of the engineered Cu_BMb indicated that the Cu_B center enhances the molecular oxygen binding affinity to the heme center in the HCO reaction.

The creation of the Cu_BMb protein has enabled subsequent efforts to improve its natural activities and create non-natural activities. In Cu_BMb, ROS formation was inhibited while H_2O production did not change significantly relative to the wild-type swMb. HCOs contain a conserved Tyr-His crosslink at the active site and this crosslink has been proposed to be important for water production (Figure 15b). However, the function of this protein modification is not fully understood. Given the potential role of the active site Tyr in the HCO reaction, Lu and co-workers replaced Gly65, a residue adjacent to the copper-binding site, with a Tyr residue in Cu_BMb. The water production rate of the Cu_BMb_{G65Y} variant was dramatically enhanced compared to that of native Cu_BMb [127]. The importance of Tyr in Cu_BMb_{G65Y}-catalysis motivated the authors to further investigate the contribution of a Tyr-His crosslink in HCO-catalysis using a ncAA incorporated by the amber-codon suppression method. In this study, Lu and co-workers chose (S)-2-Amino-3-(4-hydroxy-3-(1H-imidazol-1-yl)phenyl)propanoic acid (imiTyr) to mimic the Tyr-His crosslink [100]. The resulting imiTyr-Cu_BMb variant not only mediated the O_2 reduction reaction with more than 1000 turnovers, but also minimized the ROS formation relative to the Cu_BMb_{G65Y} variant. This study highlights the crucial role of the Tyr-His crosslink in HCO for the selective H_2O production with minimized ROS formation. The enhanced catalytic activity and selectivity of the engineered HCO reveals the value of protein modifications in developing robust biocatalysts.

NORs are similar to HCOs in that they both contain an additional iron-binding site near the heme-center. However, NORs catalyze the two-electron reduction of NO to N_2O. Additionally, the creation of the NOR activity from a heme template (i.e. swMb) is challenging

because of the lack of an NOR crystal structure. Based on their success in the Cu_BMb studies, Lu and co-workers created a swMb variant, Fe_BMb, with the NO reduction activity [121]. Starting from the Cu_BMb (swMb L29H F43H variant) scaffold, an additional Val68 to Glu mutation was introduced to mimic the Glu residue in NOR (**Error! Reference source not found.**a). The Cu_BMb_{V68E} variant had the desired Fe_B site and also exhibited the NO reduction activity at a level comparable to that of the native NOR.

In these heme enzymes engineering efforts, besides mutating the metal-binding site, the heme cofactors can also be replaced by heme analogs. For instance, native heme *b* in an HCO has been replaced with monoformyl-, diformyl- and diacetyl hemes (**Error! Reference source not found.**a). By increasing the number of electron withdrawing groups in these heme analogs, the reduction potential of the HCO increases proportionally [128]. The enhanced reduction potential resulted in an HCO with an increased dioxygen reduction activity. In another study, a manganese salen complex (Mn(salen) was introduced to swMb through a two-point covalent attachment strategy where Leu72 and Tyr103 were mutated to Cys. These two Cys residues of the $swMb_{L72C\_Y103C}$ variant then reacted with the methane thiosulfate groups of the artificial cofactor to covalently attach Mn(salen) to the protein (**Error! Reference source not found.**a) [120]. The Mn(salen)-bearing swMb catalyzed the thioanisole sulfoxidation with a remarkably improved enantioselectivity (51%) relative to that of the apo swMb (<5%), (**Error! Reference source not found.**b).

In these studies, the heme analogs are introduced into the purified apo-enzymes by *in vitro* reconstitution. This method often results in a low yield of the reconstituted holo-enzyme. To overcome this issue, an orthogonal enzyme/heme pair consisting of a heme-like cofactor and a complementary heme-binding protein has been developed (**Error! Reference source not found.**c) [129]. In this system, the native heme-transport channel, ChuA, is evolved to import the iron deuteroporphyrin IX (Fe-DPIX) cofactor, which binds selectively to an engineered cytochrome P450 scaffold from *Bacillus megaterium*. The non-natural cofactor containing variant does not catalyze the native P450 monooxygenation reactions. Instead, this artificial enzyme catalyzes cyclopropanation reactions (**Error! Reference source not found.**d).

2.3. Copper enzymes engineering

Copper-containing metalloenzymes play essential roles in multiple biological processes, especially as electron transfer mediators due to the wide range of redox potentials of their metal centers [130]. Of this family, Azurin (Az) from *Pseudomonas aeruginosa*, is a type I copper protein which has been extensively investigated and re-engineered. The trigonal bipyrimidal Cu^{2+} is bound by 2-His and 1-Cys in a plane with a Met121 and the carboxyl moiety of Gly45 binding at the axial positions (**Error! Reference source not found.**). To determine the role of the axial Met ligand in Az, Lu and co-workers replaced Met121 with multiple ncAAs through an expressed protein ligation method [131, 132]. A linear correlation was observed between the

hydrophobicity of the ncAAs and the redox potential of the copper center, with higher hydrophobicity resulting in higher reduction potentials. Other than the engineering of the copper binding ligands, Az has also served as a protein scaffold for incorporating non-native cofactors because of the stability of the apo-protein. Ferrocene is attractive organometallic catalyst but its applications are limited because of its poor solubility in aqueous solutions. To resolve this issue, an Azurin-Ferrocene complex was generated via covalent attachment between Cys112 and 2-[(Methylsulfonyl)thio]ethylferrocene [133]. The reduction potential of the ferrocene complex-containing Az variant was increased compared to the wild-type. Engineering of the secondary coordination shell residues further enhanced the reduction potential of the Az variant. The tunability of Az's reduction potential makes it a good electron transfer mediator in biological systems and an excellent target of enzyme engineering.

3. Conclusions

This review summarizes some of the recent advances in protein modifications through both chemical and biological approaches, which have wide applications in biochemistry, chemical biology, cell biology, and drug discovery. In recent years, there have been more and more cases of the application of these tools in enzyme mechanistic studies or enzyme functional engineering. We summarized a few case studies where introducing ncAAs or non-native cofactors either served as a new mechanistic study tools or introduced novel catalytic activities that do not exist in the wild-type enzyme.

Author Contributions: Conceptualization, P.L.; writing the original draft, N.N. and R.H.; review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by National Science Foundation (CHE-2004109 to P. L.)

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Sakamoto, S.,Hamachi, I. Recent progress in chemical modification of proteins. *Anal Sci* 2018; 35:18R003.

[2] Shadish, J. A.,DeForest, C. A. Site-Selective Protein Modification: From Functionalized Proteins to Functional Biomaterials. *Matter* 2020; 2:50-77.

[3] Nischan, N.,Hackenberger, C. P. Site-specific PEGylation of proteins: recent developments. *J Org Chem* 2014; 79:10727-33.

[4] Wurm, F.,Steinbach, T.,Klok, H. A. One-pot squaric acid diester mediated aqueous protein conjugation. *Chem Commun* 2013; 49:7815-7.

[5] Spicer, C. D.,Davis, B. G. Selective chemical protein modification. *Nat Commun* 2014; 5:4740.

[6] DeForest, C. A.,Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat Mat* 2015; 14:523-31.
[7] Matos, M. J., et al. Chemo- and Regioselective Lysine Modification on Native Proteins. *J Am Chem Soc* 2018; 140:4004-17.
[8] Nanna, A. R., et al. Harnessing a catalytic lysine residue for the one-step preparation of homogeneous antibody-drug conjugates. *Nat Commun* 2017; 8:1112.
[9] Schnolzer, M.,Kent, S. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* 1992; 256:221-5.
[10] Dawson, P. E.,Muir, T. W.,Clark-Lewis, I.,Kent, S. Synthesis of proteins by native chemical ligation. *Science* 1994; 266:776-9.
[11] Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annu Rev Biochem* 2003; 72:249-89.
[12] Vila-Perelló, M.,Liu, Z.,Shah, N. H.,Willis, J. A.,Idoyaga, J.,Muir, T. W. Streamlined expressed protein ligation using split inteins. *J Am Chem Soc* 2013; 135:286-92.
[13] Minnihan, E. C.,Yokoyama, K.,Stubbe, J. Unnatural amino acids: better than the real things? *F1000 biology reports* 2009; 1:88.
[14] David, Y.,Vila-Perelló, M.,Verma, S.,Muir, T. W. Chemical tagging and customizing of cellular chromatin states using ultrafast trans-splicing inteins. *Nat Chem* 2015; 7:394-402.
[15] Gilmore, J. M.,Scheck, R. A.,Esser-Kahn, A. P.,Joshi, N. S.,Francis, M. B. N-terminal protein modification through a biomimetic transamination reaction. *Angew Chem Int Ed* 2006; 45:5307-11.
[16] Scheck, R. A.,Francis, M. B. Regioselective labeling of antibodies through N-terminal transamination. *ACS Chem Bio* 2007; 2:247-51.
[17] Scheck, R. A.,Dedeo, M. T.,Iavarone, A. T.,Francis, M. B. Optimization of a biomimetic transamination reaction. *J Am Chem Soc* 2008; 130:11762-70.
[18] Christman, K. L.,Broyer, R. M.,Tolstyka, Z. P.,Maynard, H. D. Site-specific protein immobilization through N-terminal oxime linkages. *J Mater Chem* 2007; 17:2021-7.
[19] Jie, W.,Peng, C. Development and applications of bioorthogonal cleavage reactions. *Acta Chim Sinica* 2017; 75:1173-82.
[20] Chalker, J. M.,Bernardes, G. J.,Davis, B. G. A "tag-and-modify" approach to site-selective protein modification. *Acc Chem Res* 2011; 44:730-41.
[21] Crankshaw, M. W.,Grant, G. A. Modification of cysteine. *Current protocols in protein science* 1996; 3:15.1. 1-.1. 8.
[22] Goddard, D. R.,Michaelis, L. Derivatives of keratin. *J Biol Chem* 1935; 112:361-71.
[23] Lundell, N.,Schreitmüller, T. Sample preparation for peptide mapping—a pharmaceutical quality-control perspective. *Anal Biochem* 1999; 266:31-47.
[24] Moore, J.,Ward, W. Cross-linking of bovine plasma albumin and wool keratin. *J Am Chem Soc* 1956; 78:2414-8.
[25] Tsao, T.-C.,Bailey, K. The extraction, purification and some chemical properties of actin. *Biochim Biophys Acta* 1953; 11:102-13.
[26] Masri, M. S.,Friedman, M. Protein reactions with methyl and ethyl vinyl sulfones. *J Protein Chem* 1988; 7:49-54.
[27] Fontaine, S. D.,Reid, R.,Robinson, L.,Ashley, G. W.,Santi, D. V. Long-term stabilization of maleimide–thiol conjugates. *Bioconjugate Chem* 2015; 26:145-52.

[28] Lyon, R. P., et al. Self-hydrolyzing maleimides improve the stability and pharmacological properties of antibody-drug conjugates. *Nat Biotechnol* 2014; 32:1059-62.
[29] Shiu, H.-Y.,Chong, H.-C.,Leung, Y.-C.,Zou, T.,Che, C.-M. Phosphorescent proteins for bio-imaging and site selective bio-conjugation of peptides and proteins with luminescent cyclometalated iridium (III) complexes. *Chem Commun* 2014; 50:4375-8.
[30] Koniev, O., et al. Selective irreversible chemical tagging of cysteine with 3-arylpropiolonitriles. *Bioconjugate Chem* 2014; 25:202-6.
[31] Kolodych, S., et al. CBTF: New amine-to-thiol coupling reagent for preparation of antibody conjugates with increased plasma stability. *Bioconjugate Chem* 2015; 26:197-200.
[32] Abbas, A.,Xing, B.,Loh, T. P. Allenamides as orthogonal handles for selective modification of cysteine in peptides and proteins. *Angew Chem Int Ed* 2014; 53:7491-4.
[33] Conte, M. L.,Staderini, S.,Marra, A.,Sanchez-Navarro, M.,Davis, B. G.,Dondoni, A. Multi-molecule reaction of serum albumin can occur through thiol-yne coupling. *Chem Commun* 2011; 47:11086-8.
[34] Kuan, S. L.,Wang, T.,Weil, T. Site-Selective Disulfide Modification of Proteins: Expanding Diversity beyond the Proteome. *Chem Eur J* 2016; 22:17112-29.
[35] Forte, N.,Chudasama, V.,Baker, J. R. Homogeneous antibody-drug conjugates via site-selective disulfide bridging. *Drug Discovery Today: Technologies* 2018; 30:11-20.
[36] Martínez-Sáez, N., et al. Oxetane Grafts Installed Site-Selectively on Native Disulfides to Enhance Protein Stability and Activity In Vivo. *Angew Chem Int Ed* 2017; 56:14963-7.
[37] Lee, M. T.,Maruani, A.,Baker, J. R.,Caddick, S.,Chudasama, V. Next-generation disulfide stapling: reduction and functional re-bridging all in one. *Chem Sci* 2016; 7:799-802.
[38] Jones, M. W., et al. Direct peptide bioconjugation/PEGylation at tyrosine with linear and branched polymeric diazonium salts. *J Am Chem Soc* 2012; 134:7406-13.
[39] Chen, S.,Li, X.,Ma, H. New approach for local structure analysis of the tyrosine domain in proteins by using a site-specific and polarity-sensitive fluorescent probe. *ChemBioChem* 2009; 10:1200-7.
[40] Seim, K. L.,Obermeyer, A. C.,Francis, M. B. Oxidative modification of native protein residues using cerium (IV) ammonium nitrate. *J Am Chem Soc* 2011; 133:16970-6.
[41] Tilley, S. D.,Francis, M. B. Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J Am Chem Soc* 2006; 128:1080-1.
[42] Ruan, Z.,Sauermann, N.,Manoni, E.,Ackermann, L. Manganese-Catalyzed C− H Alkynylation: Expedient Peptide Synthesis and Modification. *Angew Chem Int Ed* 2017; 56:3172-6.
[43] Schischko, A.,Ren, H.,Kaplaneris, N.,Ackermann, L. Bioorthogonal Diversification of Peptides through Selective Ruthenium (II)-Catalyzed C–H Activation. *Angew Chem Int Ed* 2017; 56:1576-80.
[44] Siti, W.,Khan, A. K.,de Hoog, H.-P. M.,Liedberg, B.,Nallani, M. Photo-induced conjugation of tetrazoles to modified and native proteins. *Org Biomol Chem* 2015; 13:3202-6.
[45] Seki, Y., et al. Transition metal-free tryptophan-selective bioconjugation of proteins. *J Am Chem Soc* 2016; 138:10798-801.

[46] deGruyter, J. N.,Malins, L. R.,Baran, P. S. Residue-specific peptide modification: a chemist's guide. *Biochemistry* 2017; 56:3863-73.
[47] Drienovská, I.,Alonso-Cotchico, L.,Vidossich, P.,Lledós, A.,Maréchal, J.-D.,Roelfes, G. Design of an enantioselective artificial metallo-hydratase enzyme containing an unnatural metal-binding amino acid. *Chem Sci* 2017; 8:7228-35.
[48] Drienovská, I.,Roelfes, G. Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nat Catal* 2020:1-10.
[49] Won, Y.,Pagar, A. D.,Patil, M. D.,Dawson, P. E.,Yun, H. Recent advances in enzyme engineering through incorporation of unnatural amino acids. *Biotechnol Bioprocess Eng* 2019; 24:1-13.
[50] Srinivasan, G.,James, C. M.,Krzycki, J. A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 2002; 296:1459-62.
[51] Böck, A., et al. Selenocysteine: the 21st amino acid. *Mol Microbiol* 1991; 5:515-20.
[52] Munier, R.,Cohen, G. Incorporation of structural analogues of amino acids into bacterial proteins during their synthesis in vivo. *Biochim Biophys Acta* 1959; 31:378.
[53] Yang, W.,Hendrickson, W. A.,Crouch, R. J.,Satow, Y. Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. *Science* 1990; 249:1398-405.
[54] Liu, C. C.,Schultz, P. G. Adding new chemistries to the genetic code. *Annu Rev Biochem* 2010; 79:413-44.
[55] Chin, J. W. Expanding and reprogramming the genetic code. *Nature* 2017; 550:53-60.
[56] Lajoie, M. J., et al. Genomically recoded organisms expand biological functions. *Science* 2013; 342:357-60.
[57] Mukai, T., et al. Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. *Sci Rep* 2015; 5:1-9.
[58] Dumas, A.,Lercher, L.,Spicer, C. D.,Davis, B. G. Designing logical codon reassignment – Expanding the chemistry in biology. *Chem Sci* 2015; 6:50-69.
[59] Bryson, D. I.,Fan, C.,Guo, L.-T.,Miller, C.,Söll, D.,Liu, D. R. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Bio* 2017; 13:1253.
[60] Italia, J. S., et al. An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nat Chem Bio* 2017; 13:446-50.
[61] Italia, J. S.,Addy, P. S.,Erickson, S. B.,Peeler, J. C.,Weerapana, E.,Chatterjee, A. Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J Am Chem Soc* 2019; 141:6204-12.
[62] Rackham, O.,Chin, J. W. A network of orthogonal ribosome· mRNA pairs. *Nat Chem Bio* 2005; 1:159-66.
[63] Hui, A.,De Boer, H. A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. *PNAS* 1987; 84:4762-6.
[64] Schmied, W. H.,Tnimov, Z.,Uttamapinant, C.,Rae, C. D.,Fried, S. D.,Chin, J. W. Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* 2018; 564:444-8.
[65] RIDDLE, D. L.,CARBON, J. Frameshift suppression: a nucleotide addition in the anticodon of a glycine transfer RNA. *Nat-New Biol* 1973; 242:230-4.

[66] Neumann, H.,Wang, K.,Davis, L.,Garcia-Alai, M.,Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 2010; 464:441-4.
[67] Fischer, E. C., et al. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat Chem Bio* 2020:1-7.
[68] Lang, K.,Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem Rev* 2014; 114:4764-806.
[69] Rostovtsev, V. V.,Green, L. G.,Fokin, V. V.,Sharpless, K. B. A stepwise huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew Chem Int Ed* 2002; 41:2596-9.
[70] Tornøe, C. W.,Christensen, C.,Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J Org Chem* 2002; 67:3057-64.
[71] Chin, J. W.,Santoro, S. W.,Martin, A. B.,King, D. S.,Wang, L.,Schultz, P. G. Addition of p-Azido-l-phenylalanine to the Genetic Code of Escherichia c oli. *J Am Chem Soc* 2002; 124:9026-7.
[72] Dieterich, D. C.,Link, A. J.,Graumann, J.,Tirrell, D. A.,Schuman, E. M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *PNAS* 2006; 103:9482-7.
[73] Krogager, T. P., et al. Labeling and identifying cell-specific proteomes in the mouse brain. *Nat Biotechnol* 2018; 36:156-9.
[74] Agard, N. J.,Prescher, J. A.,Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J Am Chem Soc* 2004; 126:15046-7.
[75] Codelli, J. A.,Baskin, J. M.,Agard, N. J.,Bertozzi, C. R. Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. *J Am Chem Soc* 2008; 130:11486-93.
[76] Ning, X.,Guo, J.,Wolfert, M. A.,Boons, G.-J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angew Chem Int Ed* 2008; 47:2253-5.
[77] Jewett, J. C.,Sletten, E. M.,Bertozzi, C. R. Rapid Cu-Free Click Chemistry with Readily Synthesized Biarylazacyclooctynones. *J Am Chem Soc* 2010; 132:3688-90.
[78] Nikić, I., et al. Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy. *Angew Chem Int Ed* 2016; 55:16172-6.
[79] Kim, E.,Koo, H. Biomedical applications of copper-free click chemistry: in vitro, in vivo, and ex vivo. *Chem Sci* 2019; 10:7835-51.
[80] Song, W.,Wang, Y.,Qu, J.,Lin, Q. Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via "Photoclick Chemistry" in Bacterial Cells. *J Am Chem Soc* 2008; 130:9654-5.
[81] Yu, Z.,Pan, Y.,Wang, Z.,Wang, J.,Lin, Q. Genetically Encoded Cyclopropene Directs Rapid, Photoclick-Chemistry-Mediated Protein Labeling in Mammalian Cells. *Angew Chem Int Ed* 2012; 51:10600-4.
[82] Ramil, C. P.,Lin, Q. Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. *Curr Opin Chem Biol* 2014; 21:89-95.
[83] Lata, S.,Piehler, J. Stable and Functional Immobilization of Histidine-Tagged Proteins via Multivalent Chelator Headgroups on a Molecular Poly(ethylene glycol) Brush. *Anal Chem* 2005; 77:1096-105.

[84] Baneyx, F.,Schwartz, D. T. Selection and analysis of solid-binding peptides. *Curr Opin Biotechnol* 2007; 18:312-7.
[85] Tsukiji, S.,Miyagawa, M.,Takaoka, Y.,Tamura, T.,Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat Chem Bio* 2009; 5:341.
[86] Tamura, T.,Hamachi, I. Chemistry for covalent modification of endogenous/native proteins: from test tubes to complex biological systems. *J Am Chem Soc* 2018; 141:2782-99.
[87] Gao, S.-S.,Naowarojna, N.,Cheng, R.,Liu, X.,Liu, P. Recent examples of α-ketoglutarate-dependent mononuclear non-haem iron enzymes in natural product biosyntheses. *Nat Prod Rep* 2018; 35:792-837.
[88] Poulos, T. L. Heme Enzyme Structure and Function. *Chem Rev* 2014; 114:3919-62.
[89] Solomon, E. I.,Decker, A.,Lehnert, N. Non-heme iron enzymes: Contrasts to heme catalysis. *PNAS* 2003; 100:3589-94.
[90] Song, H.,Naowarojna, N.,Cheng, R.,Lopez, J.,Liu, P., Non-heme iron enzyme-catalyzed complex transformations: endoperoxidation, cyclopropanation, orthoester, oxidative CC and CS bond formation reactions in natural product biosynthesis. In *Advances in Protein Chemistry and Structural Biology*, Elsevier: 2019; Vol. 117, pp 1-61.
[91] Siakkou, E.,Rutledge, M. T.,Wilbanks, S. M.,Jameson, G. N. Correlating crosslink formation with enzymatic activity in cysteine dioxygenase. *BBA-Proteins Proteom* 2011; 1814:2003-9.
[92] Dominy, J. E.,Hwang, J.,Guo, S.,Hirschberger, L. L.,Zhang, S.,Stipanuk, M. H. Synthesis of amino acid cofactor in cysteine dioxygenase is regulated by substrate and represents a novel post-translational regulation of activity. *J Biol Chem* 2008; 283:12188-201.
[93] Li, J., et al. Cleavage of a carbon–fluorine bond by an engineered cysteine dioxygenase. *Nat Chem Bio* 2018; 14:853-60.
[94] Li, J.,Koto, T.,Davis, I.,Liu, A. Probing the Cys-Tyr Cofactor Biogenesis in Cysteine Dioxygenase by the Genetic Incorporation of Fluorotyrosine. *Biochemistry* 2019; 58:2218-27.
[95] Simmons, C. R., et al. Crystal structure of mammalian cysteine dioxygenase a novel mononuclear iron center for cysteine thiol oxidation. *J Biol Chem* 2006; 281:18723-33.
[96] Chen, L., et al. Use of a Tyrosine Analogue To Modulate the Two Activities of a Nonheme Iron Enzyme OvoA in Ovothiol Biosynthesis, Cysteine Oxidation versus Oxidative C–S Bond Formation. *J Am Chem Soc* 2018; 140:4604-12.
[97] Yan, W., et al. Endoperoxide formation by an α-ketoglutarate-dependent mononuclear non-haem iron enzyme. *Nature* 2015; 527:539-43.
[98] Yu, Y., et al. Defining the role of tyrosine and rational tuning of oxidase activity by genetic incorporation of unnatural tyrosine analogs. *J Am Chem Soc* 2015; 137:4594-7.
[99] Zhou, Q., et al. Probing the function of the Tyr-Cys cross-link in metalloenzymes by the genetic incorporation of 3-methylthiotyrosine. *Angew Chem Int Ed* 2013; 52:1203-7.
[100] Liu, X.,Yu, Y.,Hu, C.,Zhang, W.,Lu, Y.,Wang, J. Significant Increase of Oxidase Activity through the Genetic Incorporation of a Tyrosine–Histidine Cross-Link in a Myoglobin Model of Heme–Copper Oxidase. *Angew Chem Int Ed* 2012; 51:4312-6.
[101] Cheah, I. K.,Halliwell, B. Ergothioneine; antioxidant potential, physiological function and role in disease. *BBA-Mol Basis Dis* 2012; 1822:784-93.
[102] Ames, B. N. Prolonging healthy aging: Longevity vitamins and proteins. *PNAS* 2018; 115:10836-44.

[103] Seebeck, F. P. *In vitro* reconstitution of mycobacterial ergothioneine biosynthesis. *J Am Chem Soc* 2010; 132:6632-3.
[104] Hu, W., et al. Bioinformatic and biochemical characterizations of C-S bond formation and cleavage enzymes in the fungus *Neurospora crassa* ergothioneine biosynthetic pathway. *Org Lett* 2014; 16:5382-5.
[105] Irani, S., et al. Snapshots of C-S cleavage in Egt2 reveals substrate specificity and reaction mechanism. *Cell Chem Biol* 2018; in press.
[106] Naowarojna, N., et al. Mini-review: Ergothioneine and ovothiol biosyntheses, an unprecedented trans-sulfur strategy in natural product biosynthesis. *Biochemistry* 2018; 57:3309-25.
[107] Braunshausen, A.,Seebeck, F. P. Identification and characterization of the first ovothiol biosynthetic enzyme. *J Am Chem Soc* 2011; 133:1757-9.
[108] Goncharenko, K. V.,Vit, A.,Blankenfeldt, W.,Seebeck, F. P. Structure of the sulfoxide synthase EgtB from the ergothioneine biosynthetic pathway. *Angew Chem Int Edit* 2015; 54:2821-4.
[109] Faponle, A. S.,Seebeck, F. P.,de Visser, S. P. Sulfoxide synthase versus cysteine dioxygenase reactivity in a nonheme iron enzyme. *J Am Chem Soc* 2017; 139:9259-70.
[110] Wei, W. J.,Siegbahn, P. E.,Liao, R. Z. Theoretical study of the mechanism of the non-heme iron enzyme EgtB. *Inorg Chem* 2017; 56:3589-99.
[111] Tian, G.,Su, H.,Liu, Y. Mechanism of Sulfoxidation and C–S Bond Formation Involved in the Biosynthesis of Ergothioneine Catalyzed by Ergothioneine Synthase (EgtB). *ACS Catal* 2018; 8:5875-89.
[112] Chen, L., et al. Mechanistic studies of a nonheme iron enzyme OvoA in ovothiol biosynthesis using a tyrosine analogue, 2-Amino-3-(4-hydroxy-3-(methoxyl) phenyl) propanoic acid (MeOTyr). *ACS Catal* 2018; 9:253-8.
[113] Song, H.,Her, A. S.,Raso, F.,Zhen, Z.,Huo, Y.,Liu, P. Cysteine oxidation reactions catalyzed by a mononuclear non-heme iron enzyme (OvoA) in ovothiol biosynthesis. *Org Lett* 2014; 16:2122-5.
[114] Kato, N.,Suzuki, H.,Takagi, H.,Uramoto, M.,Takahashi, S.,Osada, H. Gene disruption and biochemical characterization of verruculogen synthase of Aspergillus fumigatus. *ChemBiochem* 2011; 12:711-4.
[115] Steffan, N.,Grundmann, A.,Afiyatullov, S.,Ruan, H.,Li, S. M. FtmOx1, a non-heme Fe(II) and α-ketoglutarate-dependent dioxygenase, catalyses the endoperoxide formation of verruculogen in *Aspergillus fumigatus*. *Org Biomol Chem* 2009; 7:4082-7.
[116] Dunham, N. P., et al. Hydrogen Donation but not Abstraction by a Tyrosine (Y68) during Endoperoxide Installation by Verruculogen Synthase (FtmOx1). *J Am Chem Soc* 2019; 141:9964-79.
[117] Chan, S. I.,Li, P. M. Cytochrome c oxidase: understanding nature's design of a proton pump. *Biochemistry* 1990; 29:1-12.
[118] Collman, J. P., et al. A cytochrome c oxidase model catalyzes oxygen to water reduction under rate-limiting electron flux. *Science* 2007; 315:1565-8.
[119] Kaila, V. R.,Verkhovsky, M. I.,Wikström, M. Proton-coupled electron transfer in cytochrome oxidase. *Chem Rev* 2010; 110:7062-81.
[120] Carey, J. R., et al. A site-selective dual anchoring strategy for artificial metalloprotein design. *J Am Chem Soc* 2004; 126:10812-3.

[121] Yeung, N., et al. Rational design of a structural and functional nitric oxide reductase. *Nature* 2009; 462:1079-82.
[122] Lin, Y. Rational Design of Artificial Metalloenzymes: Case Studies in Myoglobin. *Prog Chem* 2018; 30:1464-74.
[123] Lin, Y.-W. Rational design of metalloenzymes: From single to multiple active sites. *Coord Chem Rev* 2017; 336:1-27.
[124] Nastri, F.,Chino, M.,Maglio, O.,Bhagi-Damodaran, A.,Lu, Y.,Lombardi, A. Design and engineering of artificial oxygen-activating metalloenzymes. *Chem Soc Rev* 2016; 45:5020-54.
[125] Reetz, M. T. Directed evolution of artificial metalloenzymes: a universal means to tune the selectivity of transition metal catalysts? *Acc Chem Res* 2019; 52:336-44.
[126] Sigman, J. A.,Kwok, B. C.,Lu, Y. From Myoglobin to Heme-Copper Oxidase: Design and Engineering of a CuB Center into Sperm Whale Myoglobin. *J Am Chem Soc* 2000; 122:8192-6.
[127] Miner, K. D., et al. A Designed Functional Metalloenzyme that Reduces O2 to H2O with Over One Thousand Turnovers. *Angew Chem Int Ed* 2012; 51:5589-92.
[128] Bhagi-Damodaran, A.,Petrik, I. D.,Marshall, N. M.,Robinson, H.,Lu, Y. Systematic Tuning of Heme Redox Potentials and Its Effects on O2 Reduction Rates in a Designed Oxidase in Myoglobin. *J Am Chem Soc* 2014; 136:11882-5.
[129] Reynolds, E. W.,McHenry, M. W.,Cannac, F.,Gober, J. G.,Snow, C. D.,Brustad, E. M. An evolved orthogonal enzyme/cofactor pair. *J Am Chem Soc* 2016; 138:12451-8.
[130] Liu, J., et al. Metalloproteins containing cytochrome, iron–sulfur, or copper redox centers. *Chem Rev* 2014; 114:4366-469.
[131] Berry, S. M.,Ralle, M.,Low, D. W.,Blackburn, N. J.,Lu, Y. Probing the role of axial methionine in the blue copper center of azurin with unnatural amino acids. *J Am Chem Soc* 2003; 125:8760-8.
[132] Garner, D. K., et al. Reduction potential tuning of the blue copper center in Pseudomonas aeruginosa azurin by the axial methionine as probed by unnatural amino acids. *J Am Chem Soc* 2006; 128:15608-17.
[133] Hwang, H. J.,Carey, J. R.,Brower, E. T.,Gengenbach, A. J.,Abramite, J. A.,Lu, Y. Blue ferrocenium azurin: An organometalloprotein with tunable redox properties. *J Am Chem Soc* 2005; 127:15356-7.

Appendix A: Abbreviation List

2-PCA	2-pyridinecarbaldehyde
AA	Amino acid
aaRS	tRNA synthetase
ADC	Antibody-drug conjugates
AHA	Azidohomoalanine
AlkK	Alkyne Nε-(propargyloxycarbonyl)-L-lysine
BONCAT	Biorthogonal non-canonical amino acid tagging
cAA	Canonical amino acid
CBTF	4-(cyanoethynyl)benzoyl)oxy)-2,3,5,6-tetrafluorobenzenesulfonate
C*c*O	Cytochrome *c* oxidase
CDO	Cysteine dioxygenase
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
Cu_B	Copper-binding site
Cu_BMb	swMb L29H F43H variant
EPL	Expressed protein ligation
Fe-DPIX	Iron-deuteroporphyrin IX
H2B-K20Ub	Ubiquitinated H2B
HAT	Hydrogen-atom transfer
HCO	Heme-copper oxidase
HPG	Homopropargylglycine
imiTyr	(S)-2-amino-3-(4-hydroxy-3-(1H-imidazol-1-yl)phenyl)propanoic acid
keto-ABNO	9-azabicyclo[3.3.1]nonane-3-one-*N*-oxy
KIE	Kinetic isotope effect
LD	Ligand-directed
LDNASA	Ligand-directe d *N*-acyl-*N*-alkyl sulfonamide
MeOTyr	3-methoxytyrosine
Mn(salen)	Manganese-salen complex
MtTyr	2-amino-3-(4-hydroxy-3-(methylthio)-phenyl)propanoic acid
ncAA	Noncanonical amino acid
NCL	Native chemical ligation
NHFe	Non-heme iron enzyme
NHS	*N*-hydroxysuccinimide
NOR	Nitric oxide reductase
PACE	Phage-assisted continuous evolution
PCET	Proton-coupled electron transfer
PCS	Primary coordination shell
PEG	Polyethylene glycol

PLP	Pyridoxal-5'-phosphate
POI	Protein of interest
PTC	Peptidyltransferase center
PTM	Post-translational modification
PTM	Post-translational modification
PylRS	Pyrrolysyl-tRNA synthetase
QM/MM	Quantum mechanics/molecular mechanics
RF-1	Release factor 1
ROS	Reactive oxygen species
SCS	Secondary coordination shell
SD	Shine-Dalgarno
SOD	Superoxide dismutase
SPAAC	Strain-promoted azide-alkene cycloaddition
swMb	*Physeter microcephalus* Myoglobin
tRNA	Transfer RNA
UBP	Unnatural base pair
αKG	α-ketoglutarate