Chemical Conjugation to Less Targeted Proteinogenic Amino Acids

Nanna L. Kjærsgaard,[a, b] Thorbjørn B. Nielsen,[b] and Kurt V. Gothelf*[a, b]

Protein bioconjugates are in high demand for applications in biomedicine, diagnostics, chemical biology and bionanotechnology. Proteins are large and sensitive molecules containing multiple different functional groups and in particular nucleophilic groups. In bioconjugation reactions it can therefore be challenging to obtain a homogeneous product in high yield. Numerous strategies for protein conjugation have been developed, of which a vast majority target lysine, cysteine and to a lesser extend tyrosine. Likewise, several methods that involve recombinantly engineered protein tags have been reported. In recent years a number of methods have emerged for chemical bioconjugation to other amino acids and in this review, we present the progress in this area.

1. Introduction

The need for high-quality protein conjugates is continually rising.[1–5] Protein conjugation allows the preparation of multifunctional constructs where the function of the protein is combined with a desired molecule. Examples of this include the attachment of a drug for drug delivery,[6–9] a fluorophore for visualization[10,11] or a polyethylene glycol (PEG) to improve pharmacokinetics and pharmacodynamics of a protein therapeutic.[12–14] Advanced constructs, where a protein is conjugated to another protein[15,16] or oligonucleotides,[17–19] have also emerged. This demonstrates the opportunities selective protein conjugation strategies can provide. Such constructs have found applications in research,[18] diagnostic imaging[20,21] and therapeutics.[6] For all applications, it is essential that the function of the protein is not hampered by the added functionality. Furthermore, homogeneity can be important to generate consistent results.[22–24] In this review, we distinguish between residue-specific conjugation, where only one amino acid residue is modified, and site-selective conjugation where the conjugation reaction can occur at different residues within a limited area of the protein. This means that residue-specific methods yield a fully homogeneous product whereas conjugates produced using site-selective methods have some degree of heterogeneity.

Among the most successful protein conjugation methods is the incorporation of a specific amino acid (AA), either proteinogenic or non-proteogenic, into the protein sequence.[25–29] This is often done to obtain a reactive handle for further protein modification. Other strategies utilize a ligand[30,31] or another directing group[32] to limit the reaction to the proximity of a specific site. Furthermore, some reactions rely solely on the chemical properties and diversities in the AA sidechains and their microenvironments to ensure selectivity.[33] Lastly, some methods combine features from these categories to improve selectivity even further.[34] In most cases, this requires aqueous conditions, without elevated temperature and pressure.

Furthermore, the acceptable pH range is limited for many proteins and must often be close to physiological pH. By far, the most commonly targeted AAs are cysteine and lysine. These have been widely studied and conjugation to these residues has become standard procedure in many laboratories. Typically, classical reagents such as maleimide and NHS-esters are employed, however more specialized reagents have also been developed.[33,35–38] Selective conjugation to tyrosine has also been thoroughly studied.[39,40] This residue can be targeted by different methods including Mannich-type reactions[41] and ene-type reactions with 4-phenyl-3*H*-1,2,4-triazoline-3,5(4*H*)-diones (PTADs)[42] Although not as widely used, a few reagents for selective tyrosine conjugation such as PTAD-azides and alkynes are commercially available. This review seeks to expand upon the remaining proteogenic AAs that have been less prominent targets for chemical modification. In the following, the amino acids are grouped according to their chemical properties. While most proteins contain multiple lysines and few or no available cysteines, the expansion of the toolbox of conjugation methods to other amino acids, provides improved opportunities to select the optimal conjugation strategy, both in terms of stoichiometry and selectivity, to the individual protein. Strategies for chemical modification of the amino acids are discussed and relevant examples of conjugation to individual residues are highlighted to demonstrate different reaction types.

[a] *N. L. Kjærsgaard, Prof. K. V. Gothelf*
Center for Multifunctional Biomolecular Drug Design
Interdisciplinary Nanoscience Center, Aarhus University
Gustav Wieds Vej 14, 8000 Aarhus C (Denmark)
E-mail: kvg@chem.au.dk
[b] *N. L. Kjærsgaard, Dr. T. B. Nielsen, Prof. K. V. Gothelf*
Department of Chemistry, Aarhus University
Langelandsgade 140, 8000 Aarhus C (Denmark)

2. Alkaline Amino Acid Residues

The three alkaline amino acids residues include arginine (Arg, R, pK_a 12.5), histidine (His, H, pK_a 6.0) and lysine (Lys, K, pK_a 10.8).[43] As the pK_a values indicate, Lys and Arg are both positively charged at physiological pH, while His is partially protonated depending on the environment. Here, we will focus on conjugation to His and Arg.

2.1. Histidine

His residues are frequently found in the active site of enzymes. Here, they are essential for the enzymatic function.[44] Furthermore, they have a unique ability to function as metal chelators in metalloproteins. Histidine is a relatively low-abundancy AA in native proteins as indicated in Table 1.[45] However, recombinant proteins are often expressed with a polyhistidine tag containing four or more His residues (His-tag) that facilitates purification by immobilized metal affinity chromatography. The reaction of His with various electrophiles has been a choice for His-selective conjugation. However, this type of reaction suffers from cross-reactivity. As Cys and Lys have more nucleophilic sidechains, they will compete with His in the reaction. Advanced reagents have been used for selective conjugation to His in the proximity of a ligand binding site.[46,47]

In general, reagents for ligand directed conjugation contain a ligand and a reacting group. When the ligand is bound, the reaction group will be in high local concentration and react with a reaction partner in close proximity to the binding site of the ligand. Ligand directed methods often result in less heterogeneous conjugates, however, specialized reagents are required and the reagents are often specific for a single protein.

An alternative to this approach was demonstrated in 2018 where Rai and co-workers developed a His-selective reagent using an aldehyde as chemical directing moiety (Figure 1A).[48] The reagent contained an epoxide capable of covalent irreversible conjugation to His and a 2-hydroxybenzaldehyde for fast and reversible reaction with surface Lys. The aldehyde reacts with all available Lys residues to generate imines. This reaction is fast and reversible under physiological conditions and allows the epoxide to react with a His in close proximity of an imine-forming Lys. The aldehyde can then be released and used for late-stage modification of the protein conjugate. Using this method, Rai and co-workers achieved chemo- and site-selective labeling of several proteins, including an antibody Fab fragment, under mild conditions. The conversion was very variable (from 0% to 99%), with 30% to 41% conversion for the two-step modification of trastuzumab-Fab. This strategy has recently been expanded to be used for labeling of a single Lys on the protein surface.[49] An alternative technique targeting two His-residues in close proximity, was shown to limit cross reactivity and increase chemoselectivity in His labelling. This was exploited by Brocchini and co-workers, who used a PEG-bis-sulfone reagent to selectively PEGylate proteins at a His-tag.[50] In a later study, they further expanded this method to selectively label proteins at a His_2-tag containing two His and 0–2 Gly positioned either in the N or C-terminal or internally in the protein sequence.[51]

A few electrophiles have shown to be chemoselective for His without the need for a directing group. Joshi and Rai found that 2-cyclohexenone is able to selectively label a single His residue in the presence of other surface-exposed Lys and His residues (Figure 1B).[52] The monolabeled product was obtained for several proteins using 250 equiv. of the reagent at pH 7 resulting in 25% to 80% conversion. The installed ketone could be used for subsequent introduction of a variety of tags by reaction with aminooxy-derivatives to form the oxime product. Interestingly, all labeled His residues were surface exposed. The His in the active site of the model protein, RNase A, was not modified, thereby ensuring maintained protein activity. Another example was published by Chang and co-workers, who developed a His-selective conjugation method inspired by the reversible post-translational phosphorylation of His residues in

Nanna L. Kjærsgaard received her MSc in Medicinal Chemistry in 2019 from Aarhus University and is currently finishing her PhD studies in the group of Prof. K. V. Gothelf, Center for Multifunctional Biomolecular Drug Design, Interdisciplinary Nanoscience Center and Department of Chemistry at Aarhus University, Denmark. Her research focusses on protein modification using small molecules and modified DNA for preparation of multifunctional biomolecular constructs.

Thorbjørn B. Nielsen obtained his MSc in Organic Chemistry in 2016 from Aarhus University. In 2020, he received his PhD degree in nanoscience, following a collaboration between Novo Nordisk and Aarhus University. His academic career concluded as a postdoc in the group of Kurt V. Gothelf. As of 2022, he is employed as a research scientist, focusing on functional protein discovery at International Flavors & Fragrances.

Kurt V. Gothelf graduated from Aarhus University, Denmark in 1995. After completing his PhD he conducted postdoctoral studies at Duke University. Since 2001 he has been associated with Aarhus University, as full professor since 2007. He has headed the Center for DNA Nanotechnology and is currently in charge of the Novo Nordisk Foundation: Center for Multifunctional Biomolecular Drug Design.

Table 1. Mean frequencies of amino acids (p(a)) in the genomes of living organisms and reported chemoselective protein conjugation methods.

Amino acid	p(a) %[a]	References
Ala	7.80 (2.38)	–
Arg	5.23 (1.43)	[68, 69, 70, 71, 72, 73, 74, 75]
Asp	5.19 (0.81)	[82, 83, 89, 90, 91]; N-terminal: [92, 93]
Glu	6.72 (1.24)	
Asn	4.37 (1.73)	–
Cys	1.10 (0.44)	Not covered here; Reviews: [37, 38]
Gln	3.45 (1.19)	–
Gly	6.77 (1.32)	N-terminal: [107, 108]
His	2.03 (0.41)	[48, 51, 52, 53, 55, 59, 60]
Ile	6.95 (2.16)	–
Leu	10.15 (0.86)	–
Lys	6.32 (2.53)	Not covered here; Reviews: [33, 36]
Met	2.28 (0.39)	[136, 143, 144, 147]
Phe	4.39 (0.89)	–
Pro	4.26 (1.01)	N-terminal: [109, 110, 112]
Ser	6.46 (1.17)	101. N-terminal: [94, 95, 96, 97, 98]
Thr	5.12 (0.69)	
Trp	1.09 (0.25)	[117, 118, 119, 121, 123, 124, 126, 129, 131, 132]
Tyr	3.30 (0.63)	Not covered here. Review: [39, 40]
Val	7.01 (1.18)	–

[a] Average frequencies of individual amino acids as stated in the literature.[45] The standard deviation of the distributions are shown in parentheses. –: No methods suitable for chemoselective protein conjugation reported.

prokaryotes and eukaryotes.[53] They found that thiophosphoro alkyne dichloridate (Figure 1C) could selectively label His residues under mild aqueous conditions at slightly alkaline pH (pH 8.5) in 20 min. The alkyne on the reagent allows for later attachment of a label using copper catalyzed azide alkyne cycloaddition. Approximately 75% of the product was intact after 60 h incubation in serum at 37 °C. However, it is efficiently cleaved in cell lysates at the same temperature where $< 20\%$ remains intact after 100 min. These two reactions seems to have overcome the challenge of Lys cross reactivity, however, free thiols from unpaired Cys residues still poses a great challenge for His-selective electrophiles.[54]

All the His-selective methods described so far proceeds though nucleophilic attack from His. Nakamura and co-workers have recently reported the first His-selective nucleophilic protein labeling reagent targeting histidine oxidized by $^1O_2$ (Figure 1D).[55] Using 1-methyl-4-arylurazole (MAUra) and Rose Bengal as $^1O_2$ generating reagent, they obtained His selective labeling of several peptides in 50% CH_3CN solution at pH 7.4 with white LED irradiation for 10 min on ice. No methionine, tryptophan, tyrosine or cysteine labeling was observed, however, methionine and tryptophan are oxidized under these conditions. By using a solid support modified with ruthenium photo-catalyst and an Fc-binding ligand they were able to obtain Fc-selective labeling of the antibody trastuzumab in aqueous buffer at pH 7.4. As $^1O_2$ is a very reactive species with a possible diffusion distance of ~10 nm, the reaction can be controlled around the photo-catalyst, providing the site-selectivity. However, this is only possible for very large proteins such as antibodies that generally is around 10–15 nm in size. MAUra is known to react with tyrosine under single electron transfer (SET) conditions, and as ruthenium photo-catalyst catalyzes both SET and $^1O_2$ generation, both Tyr and His can be labeled.[56–58]

In recent years, a new type of C–H functionalization for the selective reaction with His has emerged. The C2 position of the imidazole sidechain is electrophilic. This provides a unique opportunity for development of selective conjugation reactions where cross reactivity with other nucleophilic AAs is avoided. This was exploited by two research groups, who simultaneously reported the use of radical mediated Minisci-type C–H functionalization of His in peptides. Gopalakrishnan and co-workers used sulfinate salts as radical precursors to obtain selective peptide modification under mild, aqueous and metal free conditions (Figure 1E).[59] Furthermore, the method was used for site-selective introduction of a ketone as a reactive handle for oxime/hydrazone conjugation at the C2 position on His. Wang and co-workers used a visible light promoted reaction with C_4-alkyl-1,4-dihydropyridine reagents (Figure 1F).[60] Reactions were performed under acidic conditions as protonation of imidazole facilitates the reaction. Using this method, they demonstrated the selective incorporation of various functional groups, including an azide for later modification of the conjugate. However, as the reaction is performed in organic solvent under an argon atmosphere, it may not be applicable in many cases. Furthermore, this type of reaction is not compatible with free cysteine residues and more research is needed to evaluate if the method can be optimized for a broader scope. His-based metal chelation has been exploited for direct metal complexation to generate non-covalent conjugation products.[61,62] However, these are unstable under physiological conditions. His-tags and natural metal binding sites on protein surfaces has been used

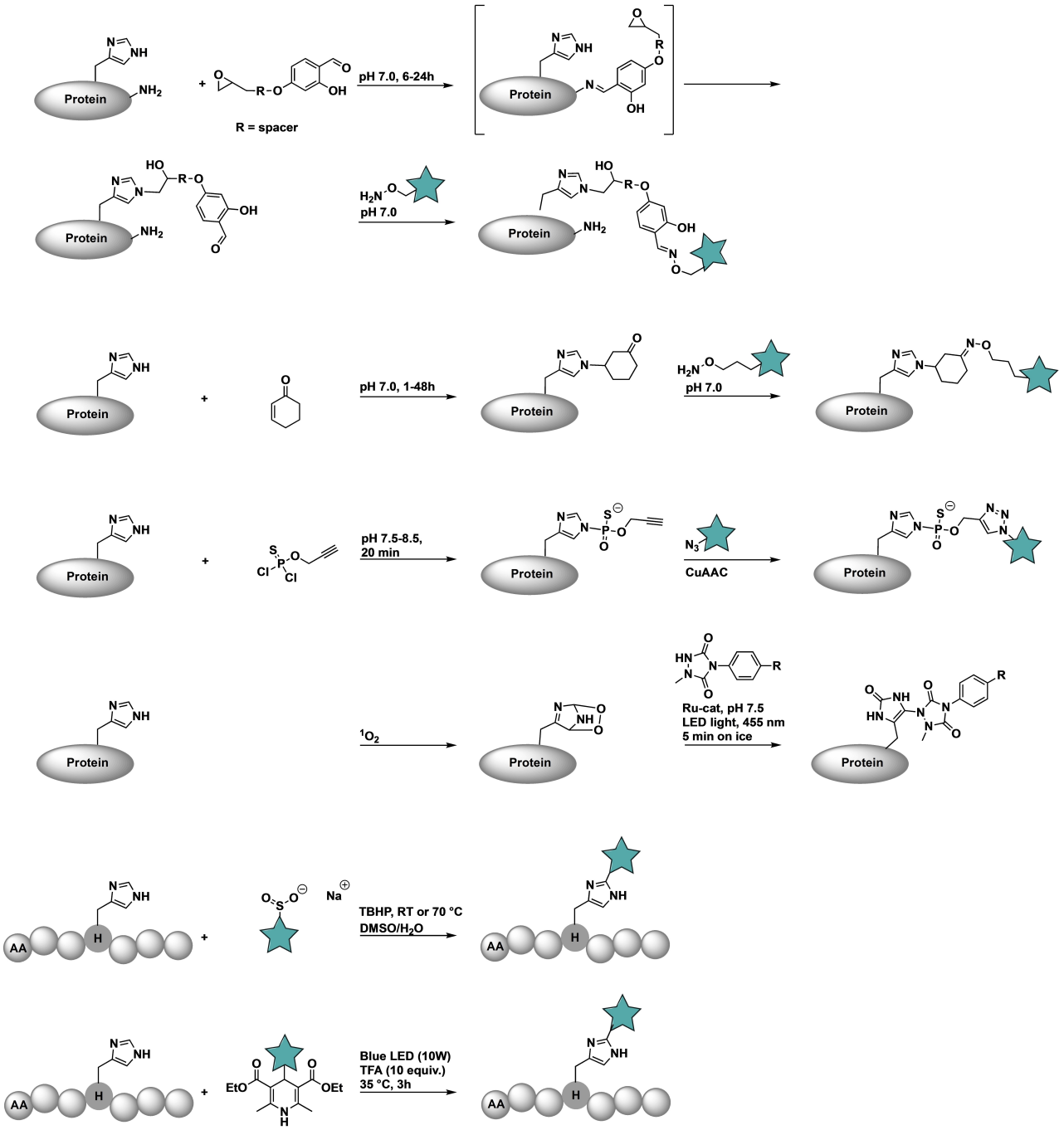

Figure 1. Reactions for selective labeling of His residues. TBHP: tert-butyl hydroperoxide. Ru-cat: ruthenium photocatalyst.

to direct reaction at other residues using metal cheaters.[63,64] Finally, His has been used as directing group for backbone N–H modification.[65]

2.2. Arginine

With a pK_a of 12.5,[43] Arg carries a positive charge under physiological conditions. Due to the charge and ability to form hydrogen bonds, it is important for both protein structure and solubility.[66] In addition to the structural properties, it is important for function and cell signaling, and is susceptible to several types of posttranslational modifications.[67] Most strategies for Arg selective conjugation involve the reaction between the guanidinium group and a modified glyoxal or other dicarbonyl compound. Early examples of this include reaction with glyoxal,[68] phenylglyoxal (illustrated in Figure 2),[69] 2,3-butanedione,[70] and later methylglyoxal.[71,72] The glyoxal is in equilibrium with the hydrated form under aqueous conditions. Hence, the reaction has been described using either the

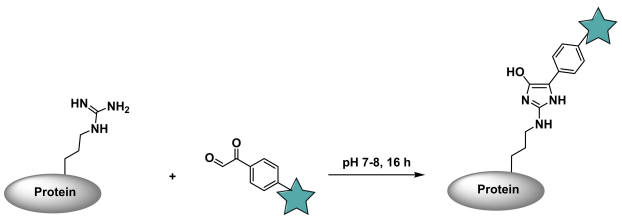

Figure 2. An example of an arginine selective reaction using dicarbonyl compounds. Reaction conditions can vary, however 3 h at 37 °C in PBS pH 7.4[76] and 16 h at 25 °C in PBS pH 7.5[77] has been reported.

unhydrated or the monohydrated form. The reaction proceeds through formation of an imidazolidine diol, which is susceptible to hydrolysis. The intermediate can be stabilized towards hydrolysis by complexation of the 1,2-diol with borate buffer or reaction with an additional glyoxal reagent, forming stable reaction products.[69,73] Furthermore, the diol can undergo elimination of water to form an heteroaromatic enol.[74] Methylglyoxal can also react with Lys or free thiols from unpaired Cys residues through the formation of imines, hemiaminals and hemithioacetals. However, these unwanted byproducts can often be removed under the purification process as the conjugation is reversible.[73] Furthermore, phenylglyoxal, glyoxal and methylglyoxal have been shown to react with α-amino groups, resulting in the formation of an α-keto acyl group at the N-terminal.[69] This side reaction can however be avoided by using suitable reaction conditions.[75]

More recently, Dawson and co-workers demonstrated the use of *p*-azidophenylglyoxal hydrate as a suitable reagent for Arg-selective protein conjugation.[75] Using optimized reaction conditions, they were able to achieve Arg-selective conjugation to the two proteins lysozyme and RNAse A with no unmodified protein observed after the reaction. Furthermore, no transamination of the N-terminus was observed. Inspired by this, Wagner and co-workers utilized *p*-azidophenylglyoxal monohydrate to modify native antibodies.[77] The use of Arg-selective conjugation to antibodies was further demonstrated by Rader and co-workers.[76] They used phenylglyoxal-triazole probes to selectively conjugate to a reactive Arg on a dual variable domain (DVD)-IgG1 forming highly stable products. Furthermore, the reaction was shown to be orthogonal to lysine labeling using a β-lactam probe. This was utilized to prepare heterodimeric DVD-IgG1 s with two different cargos using a one-pot protocol, demonstrating the strength of this method. In a very recent study, Hocek and co-workers enzymatically incorporated phenylglyoxal modified nucleoside triphosphates into oligonucleotides for conjugation of the oligonucleotides to arginine-containing peptides and proteins.[78] Arginine conjugation has also been used for identification of highly reactive Arg residues on proteins using a azide-cyclohexanedione probe for reaction and enrichment.[79] Furthermore, Arg-selective cross-linkers have been used to improve structural mass spectrometry analysis of proteins containing Lys-deficient regions.[80]

3. Acidic Amino Acid Residues

The sidechains of aspartic acid (Asp, D) and glutamic acid (Glu, E) contain a carboxylic acid functionality with a pK_a value of approximately 4,[43] depending on the local environment. Therefore, they carry a negative charge at physiological pH. These AAs possess various functions in the protein with regards to solubility and activity.[44,81] The carboxylate reactivity is limited in water, making Asp- and Glu-selective reactions a challenging prospect. Furthermore, the C-terminus of the protein also contains a carboxylate group.

Activation of protein carboxylic acids with activating agents such as a carbodiimides, followed by reaction with a desired nucleophile, is well established.[82,83] This incoming nucleophile will be competing with multiple other primary amines on the protein, and considerable amounts of by-products are to be expected. The low nucleophilicity of the carboxylate requires the probe to be highly reactive. This requirement increases the risk of cross reactions with more nucleophilic amino acids. Therefore, it is often necessary to use a reagent with a directing group. Affinity directed tetrazole based reagents have been used to react with carboxylic residues by a photo induced reaction (Figure 3A).[84,85] Upon irradiation with UV light, the tetrazole is converted to a reactive carboxy-nitrile imine that can react with various nucleophiles.

Reactions with Woodward's reagent K based reagents containing a directing ligand have also been reported.[86,87] These reagents can covalently modify the active site carboxylic residues by ester formation.[88]

Recently, Rai and co-workers reported a method to obtain single-site protein conjugation at Asp or His in proximity to a Lys using chemical guiding group.[89] Here, a chemoselective reversible reaction between an amine and an aldehyde was used to guide the reaction between the electrophile aryl sulphonate ester and a nucleophile on the protein.

A study by Zhengqiu Li and co-workers demonstrate the use of 2H-azirine-based reagents for chemoselective protein conjugation at carboxyl residues (Figure 3B).[90] The reaction is performed at room temperature, pH 7.3, and was used to label proteins both in vitro and in live cells. The reagent selectively labels Asp and Glu with a preference towards Glu because the steric hindrance is lower. Only a few off-target reactions were observed when used for protein profiling in living cells with $>95\%$ selectivity towards Asp/Glu. Furthermore, the probe showed selectivity towards buried residues in active sites compared to surface exposed residues. This property makes this type of probe ideal for reactivity profiling and proteomics. However, it is a disadvantage when used for development of protein constructs where the activity of the protein must be intact.

A third approach was demonstrated by Chaubet and co-workers who investigated multicomponent reactions for Asp/Glu conjugation on native proteins.[91] By reacting the model protein trastuzumab with a combination of suitable aldehyde and isocyanide reagents, they envisioned that protein conjugation could occur by Ugi four-center three-component reaction (U-4C-3CR) at Asp or Glu in vicinity of a Lys.

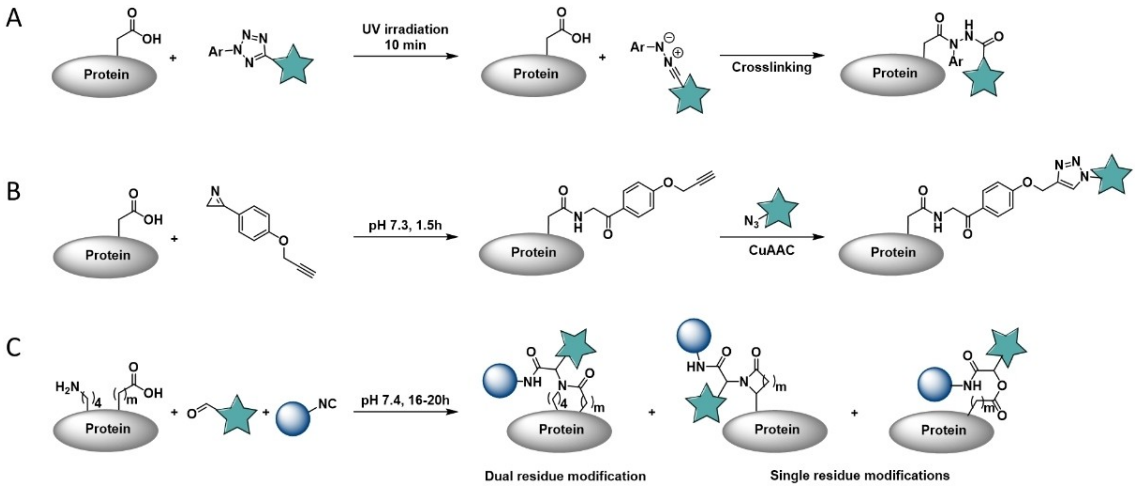

Figure 3. Examples of reactions used for conjugation to carboxylic acids. A: Conjugation using tetrazole based probes. B: Conjugation using 2*H*-azirine-based reagents. C: Three-component reaction, resulting in either dual or single residue modification.

Bioconjugation reactions were performed in PBS buffer, pH 7.4 at 25 °C for 16–20 h. Conjugation was observed with varying conversions between 0 and 97% depending on the reagent structure, the amount and the ratio between the reagents. The average number of modifications for the product mixtures were ≤ 4. Multiple functional groups could be incorporated, including azide, alkyne, isocyanide and a fluorophore. Analysis of the conjugates by LC-MS/MS validated that the reaction is specific for Asp/Glu. However, both the U-4C-3CR dual residue modification product, the U-4C-3CR single residue modification product from reaction with an N-terminal Asp/Glu and the product from single residue modification through Passerini reaction was observed (Figure 3C). The system is modular and dual labeling is easily obtained. Furthermore, selectivity towards N-terminal Asp and Glu residues was observed. Selective conjugation to N-terminal Glu has also been reported using transamination with N-methylpyridinium-4-carboxaldehyde, generating an N-terminal ketone, available for further modification with aminooxy reagents.[92] This reaction was later optimized resulting in up to 98% conversion for a protein carrying an EES tag. However, other N-terminal residues could also be modified using this procedure.[93]

4. Polar Uncharged Amino Acid Residues

The two AAs Serine (Ser, S) and threonine (Thr, T) possess a primary and secondary hydroxyl group, respectively. They are highly abundant at the surface of proteins and serve as sites for post-translational modifications such as glycosylations and phosphorylations. Furthermore, they perform important functions in the active site of several enzymes, including serine and threonine proteases. The low nucleophilicity of the Ser/Thr hydroxyl group compared to the amino group of Lys and the thiol group of Cys, makes chemoselective reactions with incoming electrophiles a challenging task. Therefore, conjugation to in-chain Thr and Ser remains a great challenge. However, several strategies have been developed for selective reaction with N-terminal Ser/Thr residues,[94,95] including introduction of an aldehyde handle through mild oxidation using periodate.[96–98] However, periodate treatment can result in cleavage of the vicinal diols in the carbohydrates or oxidation of sugar moieties in glycosylated protein.[99,100] The first method for chemoselective serine conjugation on proteins were reported by Phil S. Baran and co-workers.[101] Inspired by the phosphorylation of primary alcohols in protein performed by kinases and a new class of phosphorus reagents, they envisioned that regents based on phosphorus(V) oxidation state could be used to target primary alcohols in peptides and proteins. The P(V) based bioconjugation reaction (Figure 4) showed great selectivity for Ser, and remarkably, Ser is favored over Thr. Under optimized reaction conditions, only Ser labeling was observed for the protein conjugation reactions with 20–40% conversion. The selectivity was explained by a difference in energy barriers for the rate limiting step, found in density functional theory studies. Excess EtSH was added when the reaction was performed on peptides containing Cys to avoid side reactions with the thiirane byproduct from the reagent. Furthermore, dithiothreitol (DTT) was added to reduce any disulfide bonds formed during the reaction. Conjugation reactions were performed at 25 °C in a 4:1 DMF/water

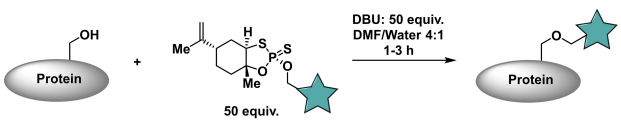

Figure 4. Serine selective reaction using phosphorus(V) reagent. DBU: 1,8-diazabicyclo[5.4.0]undec-7-ene.

mixture using 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) as base. These conditions are far from physiological conditions and might not be applicable for many proteins. Optimization of the reagent for use under aqueous and cysteine-compatible conditions would greatly enhance the protein scope for the reaction.

Similar to Ser and Thr, asparagine (Asn, N) and glutamine (Gln, Q) have proven to be difficult targets for protein conjugation. The only reported reaction with the primary amide sidechain involves proximity driven reactions and metal catalysis.[102] Here, an Fc-binding peptide modified with three dirhodium complexes was used to guide and catalyze the reaction between the alkyne-functionalized diazo reagent and an Asn sidechain of the antibody trastuzumab. However, the selectivity is due to the targeting from the peptide rather than the resection chemistry, as several amino acids including Ser, Arg, Asp, Glu, Phe and more can participate in the reaction.[34] The development of selective reactions towards these polar uncharged amino acids would be of great value to the field of protein conjugation. A starting point for gathering inspiration could be to investigate how Nature post-translationally modifies these amino acids *via* enzymes.

5. Hydrophobic Amino Acid Residues

This group can be divided into the aliphatic AAs, aromatic AAs and Methionine (Met, M) as a special case. Here, the aliphatic AAs include alanine (Ala, A), valine (Val, V), isoleucine (Ile, I), leucine (Leu, L), proline (Pro, P) and glycine (Gly, G). The aromatic amino acids are phenylalanine (Phe, F), tyrosine (Tyr, Y) and tryptophan (Trp, W). The sidechain of His is also aromatic, however, as His can be ionized at physiological pH, it will not be included in this category.

5.1. Aliphatic amino acid residues

The aliphatic amino acids either have an aliphatic sidechain (Ala, Val, Ile, Leu, Pro) or no sidechain (Gly). Due to the lack of functional groups, they are difficult to target using chemoselective regents. Therefore, very few methods for conjugation to these amino acids are available today. Although not broadly applicable, advances in C–H functionalization has led to new opportunities for the selective modification of these residues in peptides.[103–105] Gly and Pro are special cases with features that allows for development of selective reactions. These are described below.

5.1.1. Glycine

Conjugation reactions for N-terminal Gly have been developed. However, these methods often require a poly Gly-tag[106] or can be used for conjugation with other N-terminal residues.[107] Rai and co-workers published a strategy for site-selective conjugation at N-terminal Gly residues under the formation of a C–C bond as shown in Figure 5.[108] They discovered that an appropriately designed aldehyde probe could react selectively with an N-terminal Gly under the formation of an amino alcohol. The reactions were performed at room temperature, pH 7.8 for a duration of 24 to 48 h with a conversion of 40% for the model protein myoglobin and 71% for insulin. The N-terminal Gly was selectively targeted over all other proteinogenic AAs. The proposed reaction mechanism is given in Figure 5B. First, an imine is formed by reaction between the amine and the aldehyde. The following enol formation will only occur when no side chain is present. Therefore, the internal nucleophile is only generated when the N-terminal residue is Gly, ensuring residue specific reaction. The formed internal nucleophile then reacts with an additional equivalent of the aldehyde, followed by hydrolysis of the imine to form the product and regenerate one equivalent of the aldehyde reagent. A model experiment where phenylalanine amide was vortexed under the reaction conditions confirmed the presence of a high barrier for enolization, as no scrambling of the stereochemistry was observed. The utility of this method was demonstrated by selectively labeling an N-terminal Gly containing protein in a cell lysate. For most proteins, that do not naturally have an N-terminal Gly, recombinant expression would need to precede the use of this method.

5.1.2. Proline

Proline is the only proteinogenic AA containing a secondary amine, as the α-amino group is attached to the sidechain forming a pyrrolidine ring. This provides a unique feature that

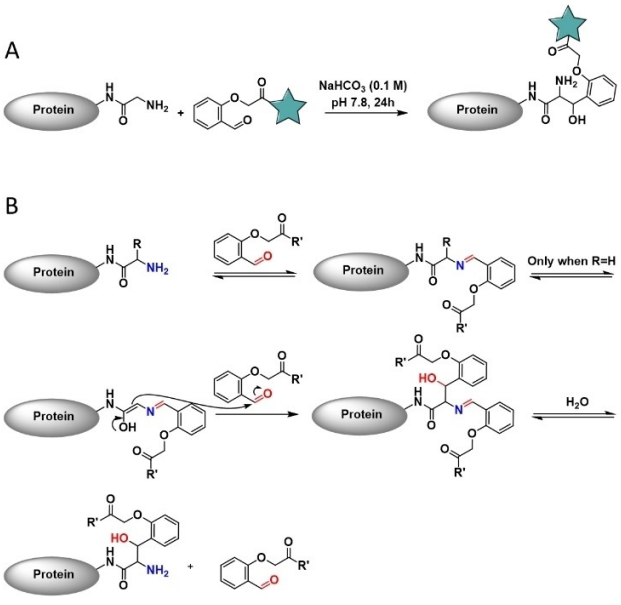

Figure 5. A: Conjugation to an N-terminal glycine residue forming an amino alcohol product. B: Proposed reaction mechanism for the selective conjugation reaction to N-terminal Gly.[108] The enol is only formed when R=H under the used reaction conditions. This ensures residue specific reaction on N-terminal Gly residues.

can be used to selectively target Pro residues at the N-terminal. Francis and co-workers found that o-amino phenols and catechols could selectively react with an N-terminal Pro under oxidative conditions (Figure 6A).[109] The reaction is performed under mild conditions at room temperature and has fast second order kinetics. Even though other N-terminal residues could also react, high preference for N-terminal Pro was observed. Furthermore, aniline could also be targeted. A limitation of this reaction however, is that the thiol groups of free cysteine residues react with both *o*-amino phenols and catechols under the same conditions and must be protected prior to the reaction. Later, the reaction was modified to a two-step reaction, where the ketone generated in the oxidative coupling to N-terminal Pro could be used for attaching an additional cargo with carrying an alkoxyamine or hydrazine group.[110] The same research group later developed an enzymatic method to selectively label N-terminal Pro using phenol derivatives and a tyrosinase (Figure 6B).[111] An alternative strategy was published by Raj and co-workers, utilizing the Petasis reaction for selective conjugation to secondary amines (Figure 6C).[112] The reaction is a multicomponent reaction involving the secondary amine, an aldehyde and a nucleophilic organoboronate. This allows two different groups to be introduced in one-pot fashion. However, the reaction rate is slower compared to the previous methods. Conversion of $>80\%$ was observed for the modification of a peptide (chain A of insulin) and the reaction could also be used to modify proteins with an N-terminal Pro.

5.2. Aromatic amino acid residues

Out of the amino acids included in this group (Phe, Tyr and Trp), Tyr is the most targeted residue for bioconjugation. Some strategies takes advantage of the fact that it can be readily converted to a reactive phenolate,[30,113] whereas other strategies utilize the reactivity of the electron-rich aromatic ring.[114,115] These reactions will not be covered in this review.

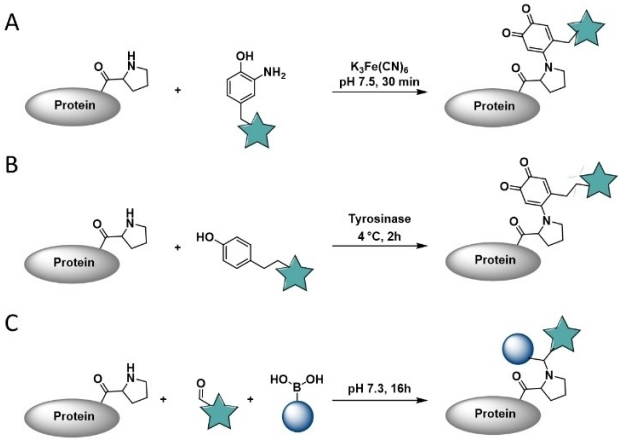

Figure 6. Bioconjugation reactions targeting an N-terminal proline residue. A: Oxidative coupling with 2-amino phenols. B: Enzyme mediated reaction with a phenol. C: A three component reaction involving the N-terminal proline, and aldehyde and an organoboronate reagent.

5.2.1. Tryptophan

Trp is, together with cysteine, one of the least abundant amino acids in proteins (Table 1).[45] It contains an indole group substituted on the C3-position. Despite the low abundancy, 90% of proteins contain few, but at least one tryptophan.[116] This makes it particularly interesting for residue specific bioconjugation as high homogeneity of the product is expected. Selective conjugation to tryptophan has however been a great challenge, and despite great efforts in the field, only a few reactions suitable for protein conjugation have been developed. One of the first tryptophan-selective bioconjugation reactions was developed by Francis and co-workers in 2004.[117] Here, they used rhodium carbenoids to react with tryptophan in aqueous solution at low pH to generate a mixture of the N- and the C2-substituted product. Later, they improved the reaction conditions by adding *t*BuNHOH, which allowed the reaction to proceed at pH 6 (Figure 7A).[118] However, reactions were performed at elevated temperature (75 to 95 °C) and a small amount of product with multiple additions of rhodium carbenoids to single indole side chains was observed. Soon after these improved conditions were reported, Ball and co-workers reported the use of dirodium metallopeptides for directed conjugation to polypeptides.[119] By combining the directing properties of the peptide with the selectivity of the chemical reaction, they were able to obtain selective conjugation at room temperature at pH 6.2 in aqueous buffer. In this method, the reactive metallocarbenoide is generated *in situ* by addition of the diazo-reagent directly to the buffer. An excess of the diazo reagent is needed to obtain good yields, as the metallocarbenoide is unstable in water. Under these conditions, both tryptophan as well as tyrosine and phenylalanine can be labeled. This cross-reactivity would in other cases generate inhomogeneous conjugates. However, this is not the case with this method as the reaction will occur in close proximity to the rhodium catalyst at the site where the directing peptide is bound. Later, the scope of the reaction was expanded to target a large variety of amino acid sidechains, including the first examples of conjugation to Gln and Asn as mentioned in section 4.[102] Another tryptophan-selective conjugation reaction was developed by Hoeg-Jensen and co-workers, using Waser's reagent[120] under gold catalysis (Figure 7B).[121] This method provided both chemo- and regioselective labeling, as only the C2-position of the indole was modified with a terminal alkyne as a reactive handle for later modification. The model protein apomyoglobin was modified using this method, resulting in a conversion of 25% and 67% to the mono- and di-functionalized product, respectively. However, organic co-solvent was necessary, and the reactions were performed in MeCN/water (3:1) with 2% TFA. Therefore, this method is only applicable to peptides and very robust proteins that tolerate organic solvent and low pH. Gold catalysis has also been used to obtain direct trifluoromethylation of tryptophan containing oligopeptides using readily available reagents.[122] However, as the reactions occur in DMSO at 40 °C, this method is not directly

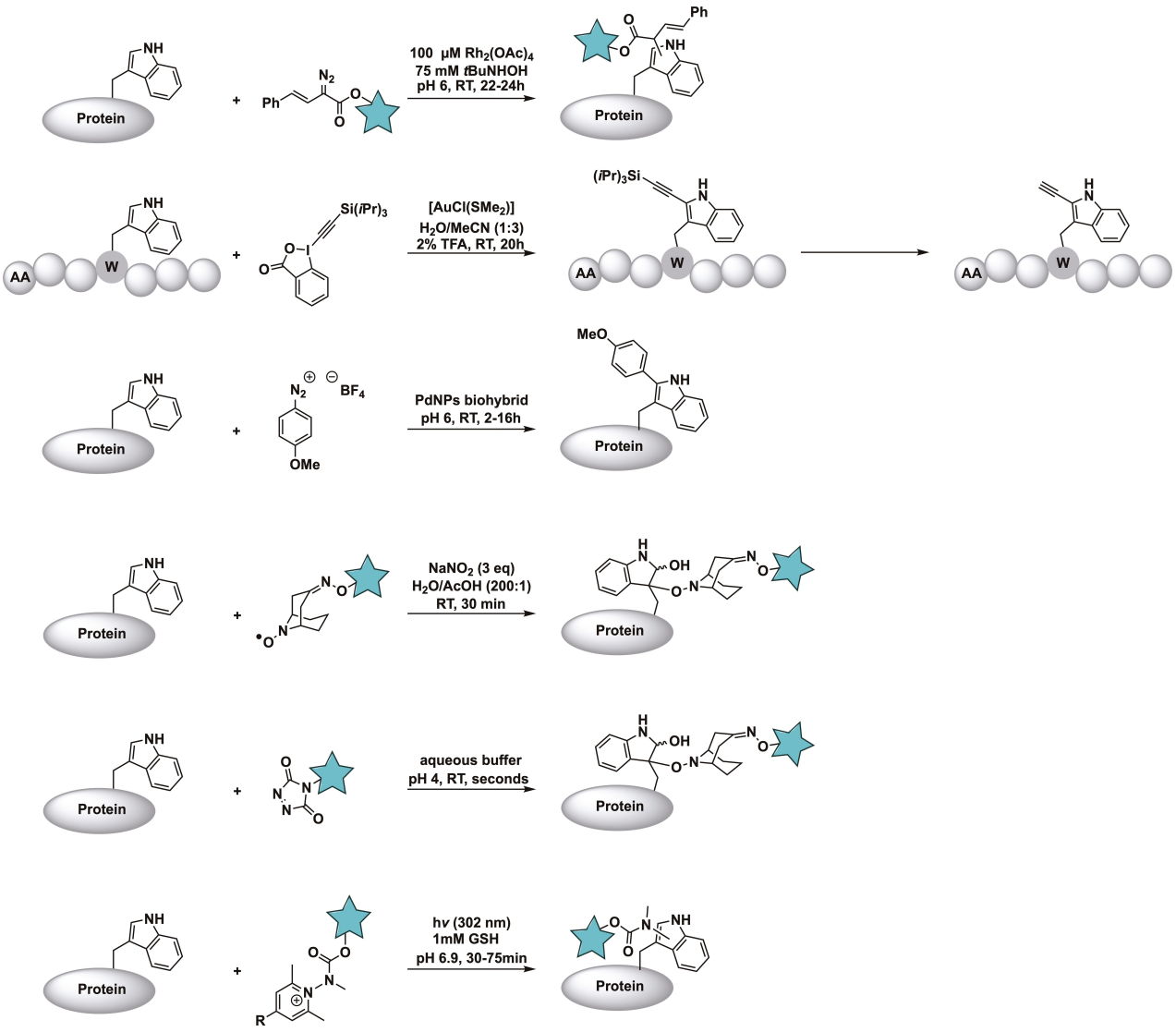

Figure 7. Tryptophan selective bioconjugation reactions. A–C: Reactions using transition metal catalysis. D: Reaction with 9-azabicyclo[3.3.1]nonane-3-one-N-oxyl (keto-ABNO). E: Selective labeling of Trp using a triazolinedione reagent at pH 4. F: Labeling of Trp using photocatalysis. R: Me or P.

compatible with most proteins. An approach, published in 2019, by Palomo and co-workers, employed a Pd nanoparticle (PdNPs) biohybrid as catalyst (Figure 7C).[123] The biohybrid consisted of Pd(0), dispersed in a protein network, forming heterogeneous nanoparticles with the size of 5 nm. Successful protein conjugation was conducted in water at RT for 48 h using 4-methoxybenzene-diazonium tetrafluoroborate as an electrophilic arylating coupling partner. Furthermore, they were able to tune the number of modifications by altering the catalyst loading.

Tryptophan can also be targeted by transition metal-free chemistry. This was demonstrated by Kanai and co-workers, who developed a mild and selective method for Trp conjugation without the use of transition metals (Figure 7D).[124] The reagent 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxyl (keto-ABNO) was added in stoichiometric amounts, together with 0.6 equiv. $NaNO_2$ and the reactions were performed in aqueous solution containing 0.1% acetic acid for 30 min. The reaction proceeds by a nucleophilic attack of the oxoammonium by the indole.

Hereafter, ether the hydrated or the dehydrated product can be formed. The utility of the reaction was demonstrated as the reaction was performed on a variety of proteins including an antibody, and the yield was dependent on the surface availability of the Trp residues. Furthermore, cross-reactivity to other AAs was not observed, and the method was orthogonal to Tyr conjugation using PTAD[125] chemistry. Davis and co-workers developed another example of using a radical reaction for trifluoromethylation of Trp in proteins.[126] They were able to obtain direct trifluoromethylation of tryptophan using 200 equiv. sodium trifluoromethanesulfinate, 25 equiv. *tert*-butyl hydroperoxide and 25 equiv. Met. The reactions were performed at 0 °C, with pH 6 for 5–10 min obtaining >50% conversion. Using this method, the model

protein myoglobin was labeled, with primarily one or two CF_3 groups.

In addition, peptides containing an N-terminal Trp can be modified using aldehyde reagents in a Pictet-Spengler reaction.[127,128] Near full conversion was observed for the reaction, however reactions were performed in organic solvents with TFA or in glacial acetic acid which is not compatible with protein-modification.

Furthermore, the reaction was observed for otherwise inaccessible Trp residues, as this method is less sensitive for the surface exposition of the residues, compared to other methods. Another recent example was published in 2022 by Madder and coworkers who report that an overlooked side reaction with Trp during triazolinedione based tyrosine labeling, can be used for selective Trp modification at lower pH (pH 4) (Figure 7E).[129] The reaction is fast (occurs within a few seconds), is performed at rt and showed 50–86% conjugation for the recombinant proteins tested in this study. It therefore constitutes a new strategy for selective Trp conjugation, however, it is only useful for proteins that can withstand reaction at pH 4. Additionally, the Trp should be positioned at the protein surfaces or loops to ensure good conversions.

Finally, Trp can also be selectively labeled using photocatalysis.[130] This was demonstrated by Taylor and co-workers, who used *N*-carbamoylpyridinium salts and UV-B light to obtain conjugation (Figure 7F).[131] The reaction proceeds through photoinduced electron transfer (PET) between Trp and a pyridinium salt, leading to cleavage of the N–N bond. The generated $Trp^{\bullet+}$ and N-centered radical can then recombine resulting in Trp conjugation.

Several groups including biotin and an alkyne were successfully transferred to Trp containing peptides. The strategy was used to label peptides and small proteins in $>85\%$ conversion in all cases. The reaction proceeds in 30—75 min without the need for organic solvents. Glutathione enhances the reaction, likely by acting as a reactive oxygen species scavenger. A similar strategy was published by Melchiorre and co-workers, using pyridinium salts for Trp conjugation on short peptides.[132] Here, single electron transfer from Trp to the pyridinium salt results in the formation of $Trp^{\bullet+}$ and a C-centered radical. The radicals can then combine, resulting in C2 alkylation of Trp. There are, to date, no phenylalanine-specific reactions that can be applied on proteins. The peptide directed rhodium catalyzed reaction is one of the only examples of selective phenylalanine conjugation.[119] However, the selectivity of this reaction is due to a combination of the reaction chemistry and the directing peptide, as the reaction can occur on several amino acids. A number of metal-catalyzed reactions for reaction with Phe containing peptides has been reported.[133,134] However, the need for organic solvents and/or elevated reaction temperatures limit the scope for these reactions with regards to protein conjugation. These trends in C–H activation for peptide conjugation may also result in the development of reactions suitable for protein Phe conjugation in the future.

5.3. Methionine

The Met side chain contains a thioether, which is unique among all proteinogenic Aas. This provides a unique chemical reactivity profile, making Met an interesting target for residue-specific protein conjugation. Furthermore, it is relatively low abundant in proteins with around 2.3% (Table 1),[45,135] and most proteins only have a single residue, if any, making site selective conjugation possible. Despite these features, relatively few Met selective reactions have been described. However, in recent years, new reactions have been developed, demonstrating the potential for methionine selective protein conjugation.

Bioconjugation to Met can either proceed through alkylation, under acidic conditions, or by redox based reactions with oxaziridine reagents (Figure 8).[136] Alkylation reagents such as alkyl halides, alkyl triflates[137—139] and epoxides[140] have been used for Met labeling of polypeptides yielding the positively charged sulfonium products. The alkylation products can be selectively dealkylated using nucleophiles or demethylated[141] to generate thioether derivatives. As Met exhibits relativity low nucleophilicity compared to other nucleophilic amino acids, it can be challenging to obtain selectivity by alkylation. Therefore, reactions are carried out at low pH ($pH \leq 3$), as this causes the other nucleophilic amino acids to be protonated, which diminishes their reactivity.[142] However, this limits the potential of this method, as not all proteins tolerate these conditions.

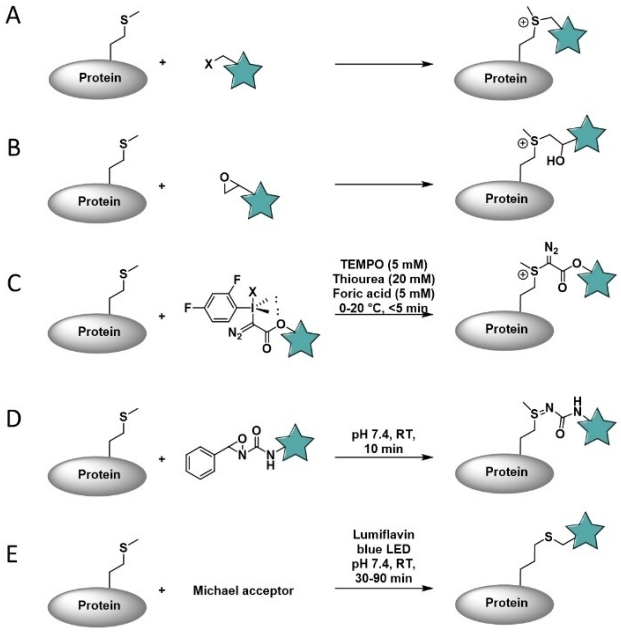

Figure 8. Methionine selective bioconjugation reactions. A: Alkylation of methionine using alkyl halide or alkyl triflate. X=I, Br, OTf. B: Alkylation of methionine using an epoxide. C: Alkylation of methionine using a hypervalent iodine reagent. A–C: All reactions are performed at $pH \leq 3$. C: X=OTf or BF_4. D: Labeling of methionine using an oxaziridine reagent by redox-activated chemical tagging. E: Labeling of Met using a photoredox reaction.

Gaunt and co-workers developed a method for Met selective conjugation at neutral pH using a hypervalent iodine reagent as shown in Figure 8C.[143] The reaction proceeds in five minutes with reported conversions of $\geq$84% for all model proteins. Following the reaction, further modification of the protein or polypeptide can be achieved through photocatalyzed reaction with a diazo group on the alkylation reagent. The diazo group further contributes to increased stability of the product, compared to other alkylation protocols. The sulfonium conjugates can be cleaved by reaction with tris(2-carboxyethyl)phosphine (TCEP). This method is highly modular and allows for fast modification, however reactions are performed at low pH (approximately pH 3) and with additives such as TEMPO and thiourea.

Chang, Toste and co-workers development a redox-activated chemical tagging strategy for Met selective bioconjugation at pH 7.4 in aqueous buffer (Figure 8D).[144] They discovered that oxaziridine reagents could react selectively with Met in a strain-driven sulfur imidation reaction. By tuning the reactivity of the reagent, they were able to promote the formation of the desired nitrogen-transfer product over the unwanted oxygen-transfer product. Yields of up to 95% were obtained for protein conjugation. Furthermore, the reaction proceeds in 10 min. The sulfimide bond was stable for 1 h in 1 N HCl and 1 N NaOH with 3% and 18% conversion after 18 h respectively. A conversion of 9% was observed after 1 h at 80 °C. The versatility of the strategy was demonstrated, as the reaction was used to label both native proteins and antibody Fab fragments with a recombinantly introduced Met.

The reaction was also used for chemoproteomic identification of functional Met residues in HeLa cells. Further work has been performed to increase the stability of the product by altering the structure of the reagent.[145] A study shows that it is not only the structure of the reagent but also the location of the Met that determines the stability of the product.[146] Finally, MacMillan and co-workers demonstrated the use of a photoredox reaction for Met modification at the carbon in the methyl group (Figure E).[147] Using lumiflavin as photocatalyst and blue LED light they were able to selectively alkylate Met over on different proteins in 43% to full conversion. Several Michael acceptors with different functional groups were employed, demonstrating the versatility of the reaction. Furthermore, they were able to conjugate to green fluorescent protein in 45% conversion with 95% retention of fluorescence. The reaction proceeds through single electron transfer between the triplet excited state of lumiflavin and Met, followed by α-deprotonation of Met to generate an α-thio radical. The radical can then react with a Michael acceptor. Subsequent hydrogen atom transfer results in a stable Met conjugation product. It should be noted that mixtures of mono-, bis-, and trisalkylation products were obtained for some acceptors. However, all were positioned at the same Met residue. The method is not compatible with unpaired Cys residues as the generated thiol radical can react with the acceptor. Cysteine residues involved in disulfide bridges are not modified.

6. Perspectives

In this review, we focused on chemical protein conjugation reactions targeting residues that has been less prominent targets for chemical modification. Conjugation to nucleophilic residues such as Lys and Cys is well established and Tyr has also become a relatively common target for protein modification. Therefore, this review is focusing on conjugation to residues that are emerging as targets for bioconjugation.

Chemoselective protein conjugation provides a great tool for modification of both natural and genetically altered proteins. Furthermore, they benefit from using relatively simple reagents compared to most guided bioconjugation methods. However, reactions can still be improved in terms of conversion, selectivity and milder reaction conditions. While chemoselective reactions for highly abundant residues such as Lys often provides a global labeling, targeting less abundant residues may provide a more site selective method as only one or a few residues are available for reaction.

As the field of organic chemistry develops, potential new protein conjugation reactions are investigated. Reactions targeting His, N-terminal Pro, Trp and Met residues are becoming more accessible, and new strategies continue to emerge. Reactions using umpolung strategies,[55] single electron transfer,[132,113] multicomponent reactions[90,112] and metal catalysis[118] are now used to obtain selective conjugation and avoid side products from reaction with nucleophilic amino acid residues. In comparison, none or very few selective reactions have been reported ft he aliphatic amino acids, Phe, Asn/Gln and Ser/Thr, in particular when disregarding terminal residues. Here, C–H activation and metal-catalyzed reactions are promising new strategies, however, more research is needed to develop methods that are compatible with aqueous conditions and lower temperatures. Only one reaction for chemoselective conjugation to internal Ser on proteins has emerged.[101] Inspired by phosphorylation performed by kinases, Baran and co-workers developed a new phosphorus (V) reagent for selective Ser conjugation, demonstrating how nature can help inspire new reactions.

As high-quality protein conjugates are required for many research and industry purposes, development of suitable protein conjugation methods are essential. New reactions should ideally provide high conversion ft he target protein forming a single product with one or multiple labels. Reaction conditions should be compatible with the protein of interest to ft hehe protein structurally and functionally intact after modification. Therefore, reaction should ideally be performed at physiological pH without the need for elevated temperature or organic solvent. Furthermore, short reaction times and simple procedures are important ft he method to become broadly applicable. This also constitutes a great advantage over conjugation methods requiring genetic modification ft he protein. We believe that the field of chemical protein conjugation will continue to contribute with new reactions and provide new means to generate high quality protein conjugates.

Acknowledgements

The work is funded by the Novo Nordisk foundation to Center for Multifunctional Biomolecular Drug Design (CEMBID) (Grant Number NNF17OC0028070).

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bioconjugation · chemical protein labeling · protein conjugation · protein modifications · site-specific conjugation

[1] J. A. Shadish, C. A. Deforest, *Matter* **2020**, *2*, 50—77.
[2] E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nat. Chem. Rev.* **2019**, *2*, 147–171.
[3] W. H. So, Y. Zhang, W. Kang, C. T. T. Wong, H. Sun, J. Xia, *Curr. Opin. Biotechnol.* **2017**, *48*, 220–227.
[4] J. N. DeGruyter, L. R. Malins, P. S. Baran, *Biochemistry* **2017**, *56*, 3863–3873.
[5] O. Koniev, A. Wagner, *Chem. Soc. Rev.* **2015**, *44*, 5495–5551.
[6] A. Beck, L. Goetsch, C. Dumontet, N. Corvaïa, *Nat. Rev. Drug Discovery* **2017**, *16*, 315–337.
[7] K. Yamada, Y. Ito, *ChemBioChem* **2019**, *20*, 2729–2737.
[8] S. J. Walsh, J. D. Bargh, F. M. Dannheim, A. R. Hanby, H. Seki, A. J. Counsell, X. Ou, E. Fowler, N. Ashman, Y. Takada, A. Isidro-Llobet, J. S. Parker, J. S. Carroll, D. R. Spring, *Chem. Soc. Rev.* **2021**, *50*, 1305–1353.
[9] P. M. LoRusso, D. Weiss, E. Guardino, S. Girish, M. X. Sliwkowski, *Clin. Cancer Res.* **2011**, *17*, 6437–6448.
[10] J. Liu, J. Hanne, B. M. Britton, M. Shoffner, A. E. Albers, J. Bennett, R. Zatezalo, R. Barfield, D. Rabuka, J. Lee, R. Fishel, *Sci. Rep.* **2015**, *5*, 16883.
[11] G. Zhang, S. Zheng, H. Liu, P. R. Chen, *Chem. Soc. Rev.* **2015**, *44*, 3405–3417.
[12] T. Majtan, E. M. Bublil, I. Park, E. Arning, T. Bottiglieri, F. Glavin, J. P. Kraus, *Life Sci.* **2018**, *200*, 15–25.
[13] Y. Hou, Y. Zhou, H. Wang, J. Sun, R. Wang, K. Sheng, J. Yuan, Y. Hu, Y. Chao, Z. Liu, H. Lu, *ACS Cent. Sci.* **2019**, *5*, 229–236.
[14] N. Nischan, C. P. R. Hackenberger, *J. Org. Chem.* **2014**, *79*, 10727–10733.
[15] S. Schulte, *Thromb. Res.* **2009**, *124*, 6–8.
[16] M. B. Geeson, G. J. L. Bernardes, *ACS Cent. Sci.* **2020**, *6*, 1473–1475.
[17] J. B. Trads, T. Tørring, K. V. Gothelf, *Acc. Chem. Res.* **2017**, *50*, 1367–1374.
[18] J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, R. Jungmann, *Nat. Publ. Gr.* **2017**, *12*, 1198–1228.
[19] A. Mullard, *Nat. Rev. Drug Discovery* **2022**, *21*, 6–8.
[20] C. H. Nielsen, T. E. Jeppesen, L. K. Kristensen, M. M. Jensen, H. H. El Ali, J. Madsen, B. Wiinberg, L. C. Petersen, A. Kjaer, *J. Nucl. Med.* **2016**, *57*, 1112–1119.
[21] G. W. Severin, J. Fonslet, L. K. Kristensen, C. H. Nielsen, A. I. Jensen, A. Kjær, A. P. Mazar, K. Johnston, U. Köster, *Sci. Rep.* **2022**, *12*, 3863.
[22] J. R. Junutula, H. Raab, S. Clark, S. Bhakta, D. D. Leipold, S. Weir, Y. Chen, M. Simpson, S. P. Tsai, M. S. Dennis, Y. Lu, Y. G. Meng, C. Ng, J. Yang, C. C. Lee, E. Duenas, J. Gorrell, V. Katta, A. Kim, K. Mcdorman, K. Flagella, R. Venook, S. Ross, S. D. Spencer, W. L. Wong, H. B. Lowman, R. Vandlen, M. X. Sliwkowski, R. H. Scheller, P. Polakis, W. Mallet, *Nat. Biotechnol.* **2008**, *26*, 925–932.
[23] Y. T. Adem, K. A. Schwarz, E. Duenas, T. W. Patapo, W. J. Galush, O. Esue, *Bioconjugate Chem.* **2014**, *25*, 656–664.
[24] C. A. Boswell, E. E. Mundo, C. Zhang, D. Bumbaca, N. R. Valle, K. R. Kozak, A. Fourie, J. Chuh, N. Koppada, O. Saad, H. Gill, B. Shen, B. Rubinfeld, J. Tibbitts, S. Kaur, F. Theil, P. J. Fielder, L. A. Khawli, K. Lin, *Bioconjugate Chem.* **2011**, *22*, 1994–2004.
[25] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.
[26] H. Neumann, K. Wang, L. Davis, M. Garcia-alai, J. W. Chin, *Nature* **2010**, *464*, 441–444.
[27] C. S. Diercks, D. A. Dik, P. G. Schultz, *Chem* **2021**, *7*, 2883–2895.
[28] D. D. Young, P. G. Schultz, *ACS Chem. Biol.* **2018**, *13*, 854–870.
[29] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764–4806.
[30] Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa, I. Hamachi, *J. Am. Chem. Soc.* **2008**, *130*, 245–251.
[31] T. Tamura, I. Hamachi, *J. Am. Chem. Soc.* **2019**, *141*, 2782–2799.
[32] M. R. Mortensen, M. B. Skovsgaard, A. H. Okholm, C. Scavenius, D. M. Dupont, C. B. Rosen, J. J. Enghild, J. Kjems, K. V. Gothelf, *Bioconjugate Chem.* **2018**, *29*, 3016–3025.
[33] M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés, G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, *140*, 4004–4017.
[34] J. Ohata, Z. T. Ball, *J. Am. Chem. Soc.* **2017**, *139*, 12617–12622.
[35] A. L. Baumann, S. Schwagerus, K. Broi, K. Kemnitz-hassanin, C. E. Stieger, N. Trielo, P. Schmieder, C. P. R. Hackenberger, *J. Org. Chem.* **2020**, *142*, 9544–9552.
[36] G. H. Pham, W. Ou, B. Bursulaya, M. DiDonato, A. Herath, Y. Jin, X. Hao, J. Loren, G. Spraggon, A. Brock, T. Uno, B. H. Geierstanger, S. E. Cellitti, *ChemBioChem* **2018**, *19*, 799–804.
[37] P. Ochtrop, C. P. R. Hackenberger, *Curr. Opin. Chem. Biol.* **2020**, *58*, 28–36.
[38] J. You, J. Zhang, J. Wang, M. Jin, *Bioconjugate Chem.* **2021**, *32*, 1525–1534.
[39] D. A. Dorta, D. Deniaud, M. Møvel, S. G. Gouin, *Chem. Eur. J.* **2020**, *26*, 14257–14269.
[40] P. A. Szijj, K. A. Kostadinova, R. J. Spears, V. Chudasama, *Org. Biomol. Chem.* **2020**, *18*, 9018–9028.
[41] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.
[42] H. Ban, J. Gavrilyuk, C. F. Barbas, *J. Am. Chem. Soc.* **2010**, *132*, 1523–1525.
[43] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, Freeman and Company **2012**, pp. 25–59.
[44] X. A. J. M. Ribeiro, J. D. Tyzack, N. Borkakoti, X. G. L. Holliday, J. M. Thornton, *J. Biol. Chem.* **2020**, *295*, 314–324.
[45] D. Gilis, S. Massar, N. J. Cerf, M. Rooman, *Genome Biol.* **2001**, *2*, research0049.1.
[46] Y. Takaoka, H. Tsutsumi, N. Kasagi, E. Nakata, *J. Am. Chem. Soc.* **2006**, *128*, 3273–3280.
[47] G. Chen, A. Heim, D. Riether, D. Yee, Y. Milgrom, M. A. Gawinowicz, D. Sames, *J. Am. Chem. Soc.* **2003**, *125*, 8130–8133.
[48] S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla, V. Rai, *J. Am. Chem. Soc.* **2018**, *140*, 15114–15123.
[49] S. R. Adusumalli, D. G. Rawale, K. Thakur, L. Purushottam, N. C. Reddy, N. Kalra, S. Shukla, V. Rai, *Angew. Chem. Int. Ed.* **2020**, *59*, 10332–10336; *Angew. Chem.* **2020**, *132*, 10418–10422.
[50] Y. Cong, E. Pawlisz, P. Bryant, S. Balan, E. Laurine, R. Tommasi, R. Singh, S. Dubey, K. Peciak, M. Bird, A. Sivasankar, J. Swierkosz, M. Muroni, S. Heidelberger, M. Farys, F. Khayrzad, J. Edwards, G. Badescu, I. Hodgson, C. Heise, S. Somavarapu, J. Liddell, K. Powell, M. Zloh, J. Choi, A. Godwin, S. Brocchini, *Bioconjugate Chem.* **2012**, *23*, 248–263.
[51] K. Peciak, E. Laurine, R. Tommasi, J. Choib, S. Brocchini, *Chem. Sci.* **2019**, *10*, 427–439.
[52] P. N. Joshi, V. Rai, *Chem. Commun.* **2019**, *55*, 1100–1103.
[53] S. Jia, D. He, C. J. Chang, *J. Am. Chem. Soc.* **2019**, *141*, 7294–7301.
[54] D. G. Rawale, K. Thakur, R. Adusumalli, V. Rai, *Eur. J. Org. Chem.* **2019**, 6749–6763.
[55] K. Nakane, S. Sato, T. Niwa, M. Tsushima, S. Tomoshige, H. Taguchi, M. Ishikawa, H. Nakamura, *J. Am. Chem. Soc.* **2021**, *143*, 7726–7731.
[56] S. Sato, K. Hatano, M. Tsushima, H. Nakamura, *Chem. Commun.* **2018**, *54*, 5871–5874.
[57] S. Sato, K. Nakamura, H. Nakamura, *ACS Chem. Biol.* **2015**, *10*, 2633–2640.
[58] S. Depienne, D. Alvarez-Dorta, M. Croyal, R. C. T. Temgoua, C. Charlier, D. Deniaud, M. Mével, M. Boujtita, S. G. Gouin, *Chem. Sci.* **2021**, *12*, 15374–15381.
[59] F. M. Noisier, M. J. Johansson, L. Knerr, M. A. Hayes, W. J. Drury, E. Valeur, L. R. Malins, R. Gopalakrishnan, *Angew. Chem. Int. Ed.* **2019**, *58*, 19096–19102; *Angew. Chem.* **2019**, *131*, 19272–19278.
[60] X. Chen, F. Ye, X. Luo, X. Liu, J. Zhao, S. Wang, Q. Zhou, G. Chen, P. Wang, *J. Am. Chem. Soc.* **2019**, *141*, 18230–18237.
[61] C.-Y. Wong, L.-H. Chung, S. Lin, D. S.-H. Chan, C.-H. Leung, D.-L. Ma, *Sci. Rep.* **2014**, *4*, 7136.

[62] A. I. Solomatina, P. S. Chelushkin, D. V. Krupenya, I. S. Podkorytov, T. O. Artamonova, V. V. Sizov, A. S. Melnikov, V. V. Gurzhiy, E. I. Koshel, V. I. Shcheslavskiy, S. P. Tunik, *Bioconjugate Chem.* **2017**, *28*, 426–437.
[63] S. Uchinomiya, H. Nonaka, S. Fujishima, S. Tsukiji, A. Ojida, I. Hamachi, *Chem. Commun.* **2009**, *39*, 5880–5882.
[64] C. B. Rosen, A. L. B. Kodal, J. S. Nielsen, D. H. Schaffert, C. Scavenius, A. H. Okholm, N. V. Voigt, J. J. Enghild, J. Kjems, T. Tørring, K. V. Gothelf, *Nat. Chem.* **2014**, *6*, 804–809.
[65] J. Ohata, M. B. Minus, M. E. Abernathy, Z. T. Ball, *J. Am. Chem. Soc.* **2016**, *138*, 7472–7475.
[66] J. Warwicker, S. Charonis, R. A. Curtis, *Mol. Pharm.* **2014**, *11*, 294–303.
[67] D. J. Slade, V. Subramanian, J. Fuhrmann, P. R. Thompson, *Biopolymers* **2013**, *101*, 133–143.
[68] K. Nakaya, H. Horinishi, K. Shibata, *J. Biochem.* **1966**, *61*, 345–351.
[69] K. Takahashi, *J. Biol. Chem.* **1968**, *243*, 6171–6179.
[70] J. A. Yankeelov, C. D. Mitchell, T. H. Crawford, *J. Am. Chem. Soc.* **1968**, *90*, 1664–1666.
[71] T. W. C. Lo, M. E. Westwood, A. C. Mclellan, T. Selwood, P. J. Thornalleys, *J. Biol. Chem.* **1994**, *269*, 32299–32305.
[72] T. Oya, N. Hattori, Y. Mizuno, S. Miyata, S. Maeda, T. Osawa, K. Uchida, *J. Biol. Chem.* **1999**, *274*, 18492–18502.
[73] M. A. Gauthier, H. Klok, *Biomacromolecules* **2011**, *12*, 482–493.
[74] Y. Gong, D. Andina, S. Nahar, J.-C. Leroux, M. A. Gauthier, *Chem. Sci.* **2017**, *8*, 4082–4086.
[75] D. A. Thompson, R. Ng, P. E. Dawson, *J. Pept. Sci.* **2016**, *22*, 311–319.
[76] D. Hwang, N. Nilchan, A. R. Nanna, W. R. Roush, H. Park, C. Rader, D. Hwang, N. Nilchan, A. R. Nanna, X. Li, M. D. Cameron, W. R. Roush, H. Park, C. Rader, *Cell Chem. Biol.* **2019**, *26*, 1229–1239.
[77] I. Dovgan, S. Erb, S. Hessmann, S. Ursuegui, C. Michel, C. Muller, G. Chaubet, S. Cianférani, A. Wagner, *Org. Biomol. Chem.* **2018**, *16*, 1305–1311.
[78] D. L. Leone, R. Pohl, M. Hubálek, M. Kadeřábková, M. Krömer, V. Sýkorová, M. Hocek, *Chem. Eur. J.* **2022**, *28*, e202104208.
[79] M. S. K. Wanigasekara, X. Huang, J. K. Chakrabarty, A. Bugarin, S. M. Chowdhury, *ACS Omega* **2018**, *3*, 14229–14235.
[80] A. X. Jones, Y. Cao, Y. Tang, J. Wang, Y. Ding, H. Tan, Z. Chen, R. Fang, J. Yin, R. Chen, X. Zhu, Y. She, N. Huang, F. Shao, K. Ye, R. Sun, S. He, X. Lei, M. Dong, *Nat. Commun.* **2019**, *10*, 3911–3921.
[81] S. R. Trevino, J. M. Scholtz, C. N. Pace, *J. Mol. Biol.* **2007**, *366*, 449–460.
[82] D. G. Hoare, D. E. J. Koshland, *J. Am. Chem. Soc.* **1966**, *88*, 2057–2058.
[83] D. G. Hoare, D. E. Koshland, *J. Biol. Chem.* **1967**, *242*, 2447–2453.
[84] K. Cheng, J. Lee, P. Hao, S. Q. Yao, K. Ding, Z. Li, *Angew. Chem. Int. Ed.* **2017**, *56*, 15044–15048; *Angew. Chem.* **2017**, *129*, 15240–15244.
[85] J. Marjanovic, T. M. Lewandowski, V. Marin, M. Patterson, L. Miesbauer, D. Ready, J. Williams, A. Vasudevan, Q. Lin, *J. Am. Chem. Soc.* **2016**, *138*, 14609–14615.
[86] R. B. Woodward, R. A. Olofson, *J. Am. Chem. Soc.* **1961**, *83*, 1007–1009.
[87] R. B. Woodward, R. A. Olofson, H. Mayer, *J. Am. Chem. Soc.* **1961**, *83*, 1010–1012.
[88] E. K. Fansa, P. Martı, M. Winzker, M. Baumann, A. Wittinghofer, H. Waldmann, E. K. Fansa, M. Winzker, S. Murarka, P. Janning, C. Schultz-fademrecht, M. Baumann, A. Wittinghofer, H. Waldmann, *Cell Chem. Biol.* **2017**, *24*, 589–597.
[89] D. G. Rawale, K. Thakur, P. Sreekumar, T. K. Sajeev, A. Ramesh, S. R. Adusumalli, R. K. Mishra, V. Rai, *Chem. Sci.* **2021**, *12*, 6732–6736.
[90] N. Ma, J. Hu, Z. Zhang, W. Liu, M. Huang, Y. Fan, X. Yin, J. Wang, K. Ding, W. Ye, Z. Li, *J. Am. Chem. Soc.* **2020**, *142*, 6051–6059.
[91] C. Sornay, S. Hessmann, S. Erb, I. Dovgan, A. Ehkirch, T. Botzanowski, S. Cianférani, A. Wagner, G. Chaubet, *Chem. Eur. J.* **2020**, *26*, 13797–13805.
[92] L. S. Witus, C. Netirojjanakul, K. S. Palla, E. M. Muehl, C. Weng, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2013**, *135*, 17223–17229.
[93] K. S. Palla, L. S. Witus, K. J. Mackenzie, C. Netirojjanakul, M. B. Francis, *J. Am. Chem. Soc.* **2015**, *137*, 1123–1129.
[94] P. M. Levine, T. W. Craven, R. Bonneau, K. Kirshenbaum, *Chem. Commun.* **2014**, *50*, 6909–6912.
[95] C. B. Rosen, M. B. Francis, *Nat. Publ. Gr.* **2017**, *13*, 697–705.
[96] K. F. Geoghegan, J. G. Stroh, *Bioconjugate Chem.* **1992**, *3*, 138–146.
[97] L. De Rosa, R. Di Stasi, L. Longhitano, L. D. D'Andrea, *Bioorg. Chem.* **2019**, *91*, 103160.
[98] D. Chelius, T. A. Shaler, *Bioconjugate Chem.* **2003**, *14*, 205–211.
[99] K. E. Norgard, H. Han, L. Powell, M. Kriegler, A. Varki, N. M. Varki, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1068–1072.
[100] R. Abraham, D. Moller, D. Gabel, P. Senter, I. Hellström, K. E. Hellström, *J. Immunol. Methods* **1991**, *144*, 77–86.
[101] J. C. Vantourout, S. R. Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padial, A. Istrate, K. Maziarz, N. Justine, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes, P. S. Baran, *J. Am. Chem. Soc.* **2020**, *142*, 17236–17242.
[102] C. Sci, B. V. Popp, Z. T. Ball, *Chem. Sci.* **2011**, *2*, 690–695.
[103] S. Mondal, S. Chowdhury, *Adv. Synth. Catal.* **2018**, *360*, 1884–1912.
[104] F. M. Noisier, M. A. Brimble, *Chem. Rev.* **2014**, *114*, 8775–8806.
[105] W. Wang, M. M. Lorion, J. Shah, A. R. Kapdi, L. Ackermann, *Angew. Chem. Int. Ed.* **2018**, *57*, 14700–14717; *Angew. Chem.* **2018**, *130*, 14912–14930.
[106] T. Tanaka, N. Kamiya, T. Nagamune, *FEBS Lett.* **2005**, *579*, 2092–2096.
[107] J. M. Gilmore, R. A. Scheck, A. P. Esser-kahn, N. S. Joshi, M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, *45*, 5307–5311; *Angew. Chem.* **2006**, *118*, 5433–5437.
[108] L. Purushottam, S. R. Adusumalli, U. Singh, V. B. Unnikrishnan, D. G. Rawale, M. Gujrati, R. K. Mishra, V. Rai, *Nat. Commun.* **2019**, *10*, 2539.
[109] A. C. Obermeyer, J. B. Jarman, M. B. Francis, *J. Am. Chem. Soc.* **2014**, *136*, 9572–9579.
[110] J. C. Maza, A. V. Ramsey, M. Mehare, S. W. Krska, C. A. Parish, M. B. Francis, *Org. Biomol. Chem.* **2020**, *18*, 1881–1885.
[111] J. C. Maza, D. L. V. Bader, L. Xiao, A. M. Marmelstein, D. D. Brauer, A. M. Elsohly, M. J. Smith, S. W. Krska, C. A. Parish, M. B. Francis, *J. Am. Chem. Soc.* **2019**, *141*, 3885–3892.
[112] Y. E. Sim, O. Nwajiobi, S. Mahesh, R. D. Cohen, M. Y. Reibarkh, M. Raj, *Chem. Sci.* **2020**, *11*, 53–61.
[113] E. J. Choi, D. Jung, J. S. Kim, Y. Lee, B. M. Kim, *Chem. A Eur. J.* **2018**, *24*, 10948–10952.
[114] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.
[115] J. Zhang, J. Xu, D. Cheng, C. Shi, X. Liu, B. Tan, *Nat. Commun.* **2016**, *7*, 10677.
[116] K. Gevaert, P. Van Damme, L. Martens, J. Vandekerckhove, *Anal. Biochem.* **2005**, *345*, 18–29.
[117] J. M. Antos, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 10256–10257.
[118] J. M. Antos, J. M. McFarland, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2009**, *131*, 6301–6308.
[119] B. V. Popp, Z. T. Ball, *J. Am. Chem. Soc.* **2010**, *132*, 6660–6662.
[120] J. P. Brand, J. Charpentier, J. Waser, *Angew. Chem. Int. Ed.* **2009**, *48*, 9346–9349; *Angew. Chem.* **2009**, *121*, 9510–9513.
[121] M. B. Hansen, F. Hubàlek, T. Skrydstrup, T. Hoeg-Jensen, *Chem. Eur. J.* **2016**, *22*, 1572–1576.
[122] I. Guerrero, A. Correa, *Org. Lett.* **2020**, *22*, 1754–1759.
[123] C. Perez-rizquez, O. Abian, J. M. Palomo, *Chem. Commun.* **2019**, *55*, 12928–12931.
[124] Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki, M. Kanai, *J. Am. Chem. Soc.* **2016**, *138*, 10798–10801.
[125] H. Ban, M. Nagano, J. Gavrilyuk, W. Hakamata, T. Inokuma, C. F. Barbas, *Bioconjugate Chem.* **2013**, *24*, 520–532.
[126] M. Imio, G. Karunanithy, W. Ng, A. J. Baldwin, B. G. Davis, *J. Am. Chem. Soc.* **2018**, *140*, 1568–1571.
[127] K. Pulka, M. Slupska, A. Puszko, M. Misiak, M. Wilczek, W. Kozminski, A. Misicka, *J. Pept. Sci.* **2013**, *19*, 433–440.
[128] X. Li, L. Zhang, S. E. Hall, J. P. Tam, *Tetrahedron Lett.* **2000**, *41*, 4069–4073.
[129] K. W. Decoene, K. Unal, A. Staes, O. Zwaenepoel, J. Gettemans, K. Gevaert, J. M. Winne, A. Madder, *Chem. Sci.* **2022**, *13*, 5390–5397.
[130] Y. Weng, C. Song, C. Chiang, A. Lei, *Commun. Chem.* **2020**, *3*, 171.
[131] S. J. Tower, W. J. Hetcher, T. E. Myers, N. J. Kuehl, M. T. Taylor, *J. Am. Chem. Soc.* **2020**, *142*, 9112–9118.
[132] B. Laroche, X. Tang, G. Archer, R. Di Sanza, P. Melchiorre, *Org. Lett.* **2021**, *23*, 285–289.
[133] M. S. Segundo, A. Correa, *Chem. Sci.* **2019**, *10*, 8872–8879.
[134] M. J. Terrey, C. C. Perry, W. B. Cross, *Org. Lett.* **2019**, *21*, 104–108.
[135] M. N. Fodje, S. Al-Karadaghi, *Protein Eng.* **2002**, *15*, 353–358.
[136] J. Zang, Y. Chen, W. Zhu, S. Lin, *Biochemistry* **2020**, *59*, 132–138.
[137] J. R. Kramer, T. J. Deming, *Biomacromolecules* **2012**, *13*, 1719–1723.
[138] J. R. Kramer, T. J. Deming, *Chem. Commun.* **2013**, *49*, 5144–5146.
[139] J. R. Kramer, R. Petitdemange, L. Bataille, K. Bathany, A. Wirotius, B. Garbay, T. J. Deming, E. Garanger, S. Lecommandoux, *ACS Macro Lett.* **2015**, *4*, 1283–1286.
[140] E. G. Gharakhanian, T. J. Deming, *Biomacromolecules* **2015**, *16*, 1802–1806.
[141] E. G. Gharakhanian, T. J. Deming, *Chem. Commun.* **2016**, *52*, 5336–5339.
[142] K. Lindorff-larsen, J. R. Winther, *Anal. Biochem.* **2000**, *310*, 308–310.

[143] M. T. Taylor, J. E. Nelson, M. G. Suero, M. J. Gaunt, *Nature* **2018**, *562*, 563–568.
[144] S. Lin, X. Yang, S. Jia, A. M. Weeks, M. Hornsby, P. S. Lee, R. V. Nichiporuk, A. T. Iavarone, J. A. Wells, F. D. Toste, C. J. Chang, *Science.* **2017**, *355*, 597–602.
[145] A. H. Christian, S. Jia, W. Cao, P. Zhang, A. T. Meza, M. S. Sigman, C. J. Chang, F. D. Toste, *J. Am. Chem. Soc.* **2019**, *141*, 12657–12662.
[146] S. K. Elledge, H. L. Tran, A. H. Christian, V. Steri, B. Hann, F. D. Toste, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 5733–5740.
[147] J. Kim, B. X. Li, R. Y. Huang, J. X. Qiao, W. R. Ewing, D. W. C. Macmillan, *J. Am. Chem. Soc.* **2020**, *142*, 21260–21266.

Manuscript received: April 29, 2022
Revised manuscript received: July 1, 2022
Accepted manuscript online: July 3, 2022
Version of record online: July 20, 2022