Author for correspondence:
Guilhem Chaubet
e-mail: chaubet@unistra.fr

An overview of chemo- and site-selectivity aspects in the chemical conjugation of proteins

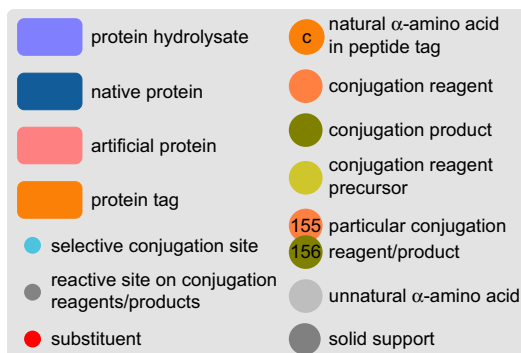Charlotte Sornay, Valentine Vaur, Alain Wagner and Guilhem Chaubet

Bio-Functional Chemistry (UMR 7199), LabEx Medalis, University of Strasbourg, 74 Route du Rhin, Illkirch-Graffenstaden 67400, France

CS, 0000-0003-0293-0423; VV, 0000-0001-8198-159X; AW, 0000-0003-3125-601X; GC, 0000-0001-6462-5112

The bioconjugation of proteins—that is, the creation of a covalent link between a protein and any other molecule—has been studied for decades, partly because of the numerous applications of protein conjugates, but also due to the technical challenge it represents. Indeed, proteins possess inner physico-chemical properties—they are sensitive and polynucleophilic macromolecules—that make them complex substrates in conjugation reactions. This complexity arises from the mild conditions imposed by their sensitivity but also from selectivity issues, *viz* the precise control of the conjugation site on the protein. After decades of research, strategies and reagents have been developed to address two aspects of this selectivity: chemoselectivity—harnessing the reacting chemical functionality—and site-selectivity—controlling the reacting amino acid residue—most notably thanks to the participation of synthetic chemistry in this effort. This review offers an overview of these chemical bioconjugation strategies, insisting on those employing native proteins as substrates, and shows that the field is active and exciting, especially for synthetic chemists seeking new challenges.

1. Introduction

Bioconjugation can be defined as the covalent coupling between a biomolecule—in this case, a protein—and any other molecule, including another biomolecule. The resulting bioconjugates can display either the combined properties of its separate constituents—as in the case of fluorophore-conjugated proteins for example—or a new set of properties that cannot be explained by simply considering the two conjugated entities

2

protein hydrolysate

native protein

artificial protein

protein tag

selective conjugation site

reactive site on conjugation reagents/products

substituent

c natural α-amino acid in peptide tag

conjugation reagent

conjugation product

conjugation reagent precursor

155 particular conjugation
156 reagent/product

unnatural α-amino acid

solid support

Figure 1. Legend of all graphical elements used in the manuscript.

separately—as is the case with toxins leading to non-toxic but still immunogenic toxoids upon chemical conjugation. Because of this modularity, applications of protein conjugation are plentiful, and too vast to be cited exhaustively. One can, however, mention the therapeutic field, which has become the object of intense scrutiny over the past decades, with the emergence of immunoconjugates—that is, bioconjugates of immunoglobulins, and more specifically antibodies—for the treatment of several types of cancers. This increased attention toward protein conjugates went hand-in-hand with improved understanding of proteins reactivity and more advanced bioconjugation tools, currently dominated by enzymatic and chemical strategies, the latter being the central topic of this review. Because of proteins' polynucleophilic nature—with the possible exception of asparagine and glutamine, all reactive groups found on the side-chains of proteinogenic amino acids are nucleophiles to some extent, large in size compared to classical small molecules and sensitivity—limiting the use of organic solvents; narrowing the range of applicable temperatures; forcing us to ban strong oxidizing/reducing agents, bases and acids, etc.—their chemical conjugation represents a challenge on many aspects. Similar to the French literary movement Oulipo using constrained writing techniques to foster inventiveness and innovation, the rigid framework imposed by proteins' physico-chemical properties has been seen by many synthetic chemists as an exciting playground for the development of new chemical tools. In this regard, chemoselectivity—targeting selectively only one family of amino acid residues; and site-selectivity—targeting a single copy of a precise amino acid—represent the central issues to be addressed, in order to develop techniques that would be reproducible and lead to homogeneous, well-defined protein conjugates. After decades of work and investigation, myriad solutions to these two challenges have been offered, which will be summarized in this review. A large part of this document will be devoted to site-selective strategies—sometimes also referred to as 'site-specific', see §4 for a semantic discussion—and, more specifically, to that employing native proteins, as we believe this is where the pinnacle of chemical innovation is. In this particular section, we limited ourselves to publications in which site-selectivity was clearly demonstrated, evidenced and supported by analytical data—essentially peptide mapping studies by LC/MS–MS. While this review does not intend to be exhaustive, we tried to cover the field as extensively as possible in order to give the reader a good overview of the century spanning chemical bioconjugation, from its early development in the tanning industry to the most recent site-selective photoredox-catalysed techniques, thus providing a broader picture than the most recent reviews on a similar topic [1–5]. A legend explaining the various graphical elements used in this manuscript is provided in figure 1 to facilitate the reader's understanding.

2. Chemoselective strategies

2.1. Early developments

Initially based on empirical observations without a clear understanding of the processes taking place at the molecular level, the chemical modification of proteins had long been present and used for practical purposes before becoming a topic of interest in chemical biology. In the mid-nineteenth century, chromium(III) salts were thus employed in the tanning industry to produce leathers by cross-linking collagen proteins in animals' hides [6]. In the early 1920s, formaldehyde was used for the modification of toxins related to a certain number of bacterial diseases. In this chemical process, the
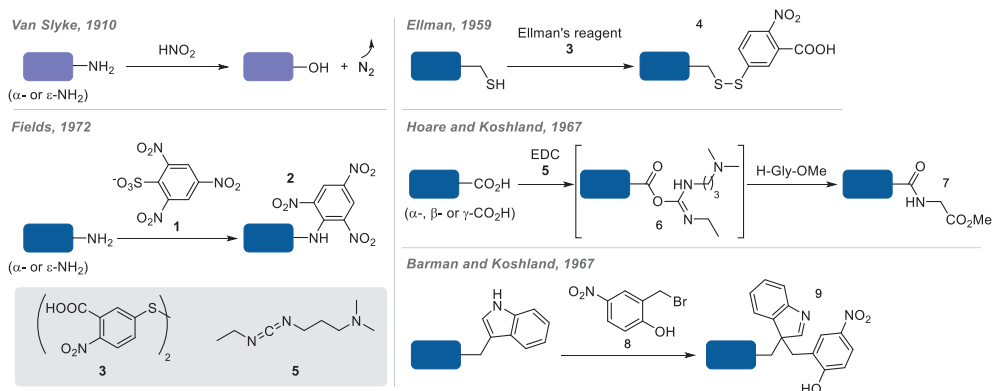
Figure 2. Early developments of chemoselective reagents for the conjugation of proteins.

bacterial proteins would lose their toxicity while retaining their ability to elicit an immune response, turning them into effective toxoid vaccines [7].

Over the twentieth century, the improvement of analytical techniques for the characterization of proteins (e.g. cation-exchanger amino acid analyser, ion exchange and gel exclusion chromatography) coupled with the definitive mapping of all proteinogenic amino acids (identification of methionine in 1922; threonine in 1925) have highly contributed to the development of new techniques for the chemical modification of proteins [6]. Several functional group-selective—i.e. *chemoselective, vide infra*— methods were thus described between the late 1950s and early 1970s, helping with the determination of proteins' primary structures (figure 2). Variations of the original Van Slyke's procedure, consisting of measuring the development of gaseous dinitrogen from the reaction of primary amines in proteins hydrolysates with nitrous acid, were developed for the quantification of amine groups in intact proteins [8,9]. For instance, Fields reported the use of trinitrobenzenesulfonate **1** (TNBS) for the formation of trinitroanilines **2** that could be quantified by measuring their absorbance at 420 nm [10].

Similar to sodium nitroprusside that gives a red colour in the presence of thiols in protein, Ellman's reagent (5,5′-dithiobis-(2-nitrobenzoic acid), DTNB; **3**) was developed in 1959 to quantify the concentration of thiols in various biological samples by measuring the absorbance of the highly chromogenic 5-thio-2-nitrobenzoic acid (TNB) **4** at 412 nm [11]. In 1967, Hoare & Koshland [12] proposed to modify all carboxylate groups (i.e. aspartate and glutamate side-chains as well as *C*-terminus) in insulin, lysozyme and ribonuclease with the water-soluble 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC; **5**). The resulting *O*-acylisoureas **6** were then reacted with the nucleophilic glycine methyl ester, leading to coupling products **7**. The protein was then hydrolysed and the extra-glycine content—as the direct reflection of carboxylate content—determined by an amino acid analyser. For the determination of tryptophan residues in proteins, Barman & Koshland [13] proposed to label those residues selectively with 2-hydroxyl-5-nitrobenzyl bromide **8**, forming coloured complexes **9** whose absorbance at 410 nm could be directly correlated to the number of indoles present in the protein.

The emergence of these chemoselective bioconjugation reagents also went hand-in-hand with a better understanding of proteins' functions. Following the idea that the residue(s) responsible for the activity of enzymes might be identified on the basis of their reactivity towards specific electrophiles [14,15], Balls & Jansen [16] showed in 1952 that diisopropyl fluorophosphate could inhibit several proteases such as α-chymotrypsin, trypsin or cholinesterase by phosphorylating selectively the serine residue present in their active site.

All these innovations facilitated the characterization of modified proteins and led to a better understanding of the reagents' chemoselectivity. In this context, chemoselectivity can be seen as the selective modification of a side-chain functional group borne by the amino acids constitutive of the protein—e.g. lysines' ε-amines—in the presence of others with similar reactivity—i.e. *N*-terminal α-amines, cysteines' thiols and nucleophiles in general.

However, early strategies tended to be hampered by limited conversion [6]. While employing more vigorous conditions (longer reaction time, larger excess of reagent, higher reaction temperature, etc.) could help overcome this limitation, it also came with the risk of eroding the chemoselectivity and affecting the conformation and activity of the protein. Consequently, this urged scientists to develop new families of reagents that could be employed under milder conditions—aqueous media, narrower range of temperature and pH, etc. —to preserve the integrity of the biomolecule. The following
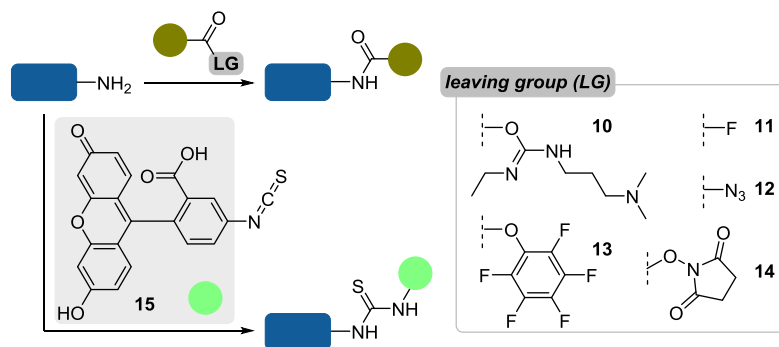
Figure 3. Classical bioconjugation methods for the chemoselective labelling of lysine residues.

section will thus present briefly the most classical chemoselective strategies that have been developed over the past decades.

2.2. Modern methods for the chemoselective conjugation of proteins

Alongside the improvement of analytical techniques for the characterization of proteins, the development of mild strategies and new reagents for their chemoselective modification arose. Those methods relied heavily and essentially on the modification of the two most nucleophilic amino acids found in proteins: lysine and cysteine.

2.2.1. Lysine residues

Lysine is one of the most abundant amino acids in proteins [17]. Composed of a linear four-carbon chain terminated with a primary amino group, lysine, along with arginine and histidine residues, contributes to the overall net positive charge of proteins. In comparison with the N-terminus α-amino group, the ε-amine of lysine possesses a slightly higher ionization point (pK_a of ε-amine is comprised between 9.3 and 9.5, α-amine between 7.6 and 8.0), thus favouring the formation of positively charged ammonium groups at physiological pH. Due to this dominant ionic character, lysines' side-chains are thus generally exposed to solvent, at the surface of proteins, making them easily accessible for chemical derivatization. Several chemoselective methods were hence described, generally conducted at alkaline pH between 8 and 10 in a wide range of buffers—including Tris [18], counterintuitively—and revolving mainly around three families of reagents: carbonyls, activated esters or iso(thio)cyanate [19–23].

As one of the fundamental reactions in organic chemistry, primary amines condense reversibly with carbonyls to form imines. Being disfavoured in aqueous medium and leading to unstable adducts, imine formation is rarely, if ever, used as a standalone bioconjugation technique but rather as a way to generate reactive intermediates that will participate in a subsequent irreversible step. Reductive amination is an example of such a two-step transformation, sodium cyanoborohydride acting generally as the reducing agent in the irreversible step leading to secondary amines [19,24], but many more bioconjugation reactions employing imines as intermediates have been developed, which will be discussed later in this review.

As an alternative to carbonyls, primary amines can also react with activated carboxylates to yield amides, the most-employed strategy in the field of protein bioconjugation (figure 3). Of the myriad reagents reported, one can cite O-acylisoureas—such as **10**, transiently generated from EDC **5**—acyl halides and azides (**11** and **12**, respectively), pentafluorophenyl esters **13** or N-hydrosuccinimidyl (NHS) esters **14** [25–30]. The latter are by far the number one reagents for lysine conjugation, thanks to their facile preparation and mild reaction conditions associated with their use—neutral to near-neutral pH, room temperature, various buffers tolerated, reaction times usually comprised between dozens of minutes and few hours—leading to exceedingly stable conjugates (the half-life of amide bonds in neutral aqueous solution has been estimated up to 600 years) [31] albeit suppressing a permanent charge at the surface of the protein and thus altering its isoelectric point. To address this limitation, imidoesters were developed, yielding positively charged amidine groups at neutral pH [32,33]. However, the amidine bond was shown to be unstable at elevated pH, thus limiting the use of imidoesters as conjugation reagents [34].

Even though NHS esters are seen as the gold standard of lysine conjugation reagents, an inevitable drawback of employing reactive electrophiles in aqueous media is their competitive hydrolytic decomposition, which affects the efficiency of the conjugation step. Increasing the pH of the reaction to work under more alkaline conditions has been proposed to facilitate amine acylation, but this was also accompanied with accelerated hydrolysis of NHS esters, with a typical half-life of just 10 min at pH 8.6 and 4°C [34]. While similar issues can be observed with isocyanates, making them disregarded reagents in bioconjugation, isothiocyanates are much more resistant towards hydrolysis and have been classically employed for the chemical transformation of amines to thioureas. The most famous member of this family of reagents is probably fluorescein isothiocyanate (FITC, **15**), frequently used for the fluorescent labelling of proteins [35–37]. Even though the rate of hydrolysis of isothiocyanates was found to be slower than that of NHS esters, thioureas tend to be less stable than amide bonds, explaining why activated esters—and reagents or strategies leading to amide conjugation product in general—remain unrivalled for the chemoselective modification of amines in proteins.

2.2.2. Cysteine residues

Because of their low abundance (1.5%) and the high reactivity of their sulfhydryl group, cysteine residues are one the most convenient targets in bioconjugation [34,38]. The ionization of cysteines' side-chain occurs at alkaline pH (thiol's pK_a 8.8–9.1) generating a negatively charged and more nucleophilic thiolate residue. Alkyl halides, disulfides and maleimides—and more generally electron-deficient alkenes and alkynes—are the three main classes of reagents used for the chemoselective conjugation of cysteines (figure 4) [20–22,39].

Alkyl halides and especially haloacetamides have been employed for the labelling of thiols in peptides and proteins for decades. Alongside 4-vinylpyridine and acrylamide, these reagents have traditionally been used to alkylate cysteines prior to Edman degradation and automated amino acid analysis [40,41], and are still a valuable tool for mass spectrometry analyses, where a preliminary conjugation step with iodoacetamide **16** is routinely performed before peptide mapping studies [42]. While alkyl halides are prone to side reactions with other nucleophilic amino acids (e.g. histidine, methionine, lysine and N-terminal residues) [43], the selectivity and reactivity of those reagents can be tuned by either varying the halide—order of reactivity: I > Br > Cl > F—or playing with reaction conditions. Cysteine selectivity has thus been demonstrated for iodoacetate derivatives when used as limiting reagents under slightly alkaline conditions (pH 8.2) for a few minutes. In the same vein, aryl halides were also described for the modification of thiols through aromatic nucleophilic substitution, generating aryl thioether species. Successful thiol conjugation when employing such reagents relies heavily on stereoelectronic effects favouring the rate-determining *ipso*-addition step; as such, aryl fluorides and chlorides bearing strongly electron-withdrawing groups in *ortho/para* positions—such as **18** and **20**—form the overwhelming majority of substrates reported to date [44].

As discussed previously, Ellman's reagent **3** (figure 2) and activated disulfide reagents in general have also proven to be highly cysteine-selective, reacting with this residue via disulfide exchange and yielding hydrolytically stable conjugates. However, their inherent reactivity makes them prone to reduction and disulfide exchange in thiol-rich, reductive environments, where their half-lives typically never exceed 1 h [45].

Of all cysteine-selective reagents reported to date, the most studied and most employed remains maleimide **22** (figure 4) [46]. Reacting with thiolates via a thio-Michael addition to form thioethers **23** [46,47], maleimides display best chemoselectivity at pH values comprised between 6.5 and 7.5 (competitive aza-Michael addition starts to occur at pH greater than 8.5) [48–50] and fast kinetics, with second-order rate constants oscillating between 100 and 1000 $M^{-1} s^{-1}$, depending on the structure of the thiol and the pH [46,51,52]. While extremely effective, maleimide conjugation is still accompanied by certain limitations: despite a good stability, maleimide adducts can undergo retro-Michael addition and thiol exchange under physiological conditions [45]; maleimides reagents can be prone to hydrolysis and yield ring-opened maleamic acids that are unreactive towards thiol addition; and maleimides are not amenable to multiple functionalization, due to their limited number of reactive sites.

With the intent of improving the latter point, Caddick and co-workers [53–55] reported in the early 2010s new families of structurally related reagents, based on maleimide (**24**; $n = 1$) and pyridazinedione (**24**, $n = 2$) skeletons, in which the incorporation of two nucleofuges, i.e. halides or thiophenolates, across the reactive C=C bond allowed two successive addition–elimination sequences, resulting in doubly functionalized conjugates **25**.

However, these were still found to suffer from *retro*-thio-Michael addition as thiol exchange was demonstrated in the presence of an excess of glutathione or 2-mercaptoethanol.
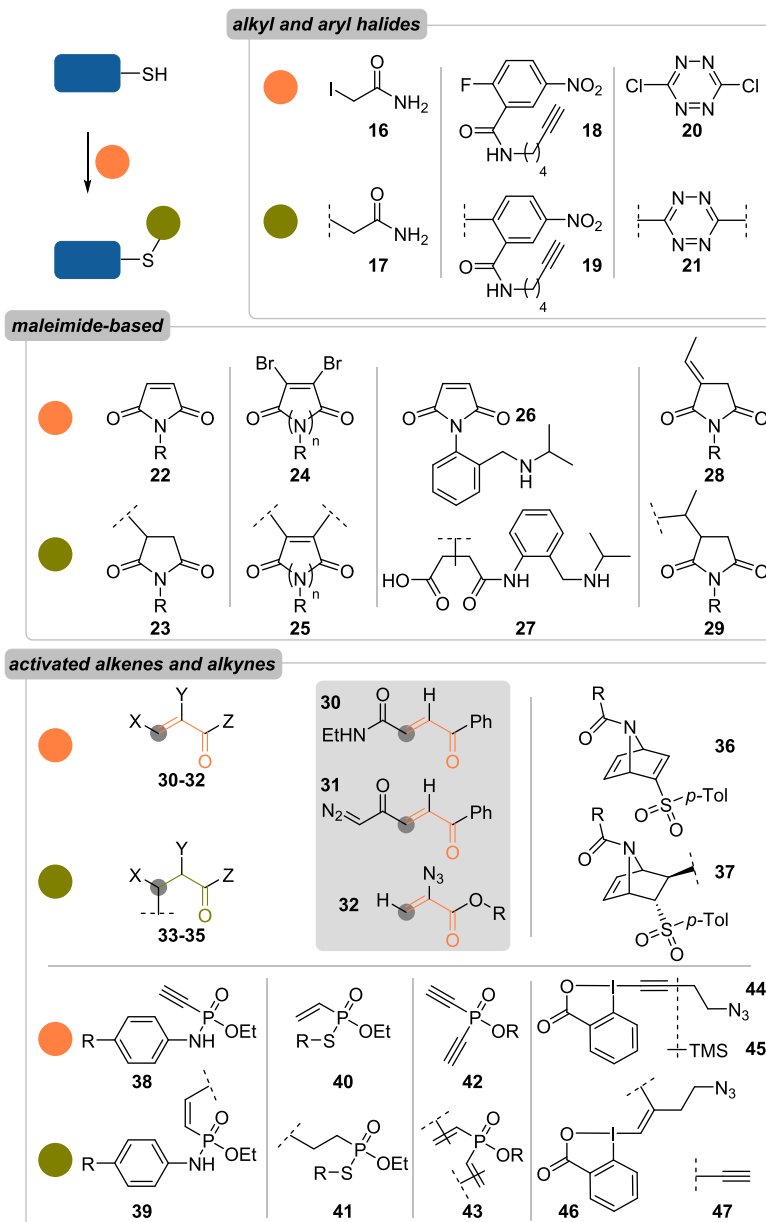
Figure 4. Overview of classical methods for the chemoselective conjugation of cysteine residues (R = alkyl chain; *p*-Tol = *para*-toluyl).

To minimize the reversibility of the *retro*-thio-Michael addition, several studies have tried to facilitate and accelerate ring hydrolysis after thiol addition, as the resulting thiomaleamic acids proved to be resistant to thiol exchange [45,56]. One notable example are *o*-CH₂NHⁱPr phenyl maleimides **26**, reported by Kalia *et al.* [57], whose corresponding thiol conjugates were rapidly hydrolysed (half-lives typically less than 1 h) to ring-opened derivatives **27** thanks to both the electron-withdrawing effect of the phenyl ring and the intramolecular maleimide activation from the appended ammonium group. In a previous study, they had also demonstrated that exocyclic olefinic pseudo-maleimides **28** led to more stable conjugates over time than classical maleimides, with negligible degrees of thiol exchange being detected on a bovine serum albumin-conjugate after 7 days of incubation at 37 °C in the presence of a 40-fold excess in *N*-acetyl cysteine [58].

Maleimides have thus been the object of intense scrutiny and can safely be considered as the mainstay of cysteine conjugation. But despite this dominant position, research towards the development of other thiol-selective reagents based on electron-deficient alkenes (and alkynes) has been quite active—and continues to be [59].

Acrylic derivatives **30** have thus been identified as effective Michael acceptors for the modification of cysteines residues (figure 4). These reagents demonstrated excellent chemoselectivity and fast kinetics for the conjugation of various proteins—including an engineered antibody; trastuzumab LC-V205C [60]—and paved the way towards the development of analogous compounds incorporating diazo or azido handles (**31** and **32**, respectively) allowing further functionalization of the protein conjugates [61,62].

Vinyl sulfones have also been known for decades as effective Michael acceptors for the modification of proteins, leading to exceedingly stable adducts (the adduct resulting from the reaction between cysteine and ethyl vinyl sulfone showed less than 15% hydrolysis after 24 h at 110 °C in 6 N HCl!) [63] despite poor chemoselectivity often being observed, with lysine and histidine residues competing with cysteines for alkylation, especially under alkaline pH (i.e. greater than or equal to 8.0) [63–65]. Interestingly, by embedding the vinyl sulfone motif into an azabicyclic system, as in reagent **36**, excellent cysteine selectivity was attained by the Bernardes group for the rapid conjugation of few mutant proteins with single solvent-exposed cysteine residue [66]. However, the resulting adducts **37** suffered from *retro*-Diels–Alder reaction, leading to payload fragmentation over prolonged incubation at 37 °C in plasma.

In a similar vein, Hackenberger and co-workers [67,68] reported the use of ethynylphosphonamidates **38** and vinylphosphonothiolates **40** for either azide-to-thiol or thiol-to-thiol conjugation, respectively. Even though kinetics were rather slow compared to maleimide (less than or equal to $0.6 \, \text{M}^{-1} \, \text{s}^{-1}$), both reagents showed excellent chemoselectivity at pH 8.5 and led to various protein conjugates, including protein hetero-dimers, all of which were stable for more than 4 days in different media (i.e. PBS, cell lysates, serum and highly reductive environment). Elaborating on this P(V) platform, the same group recently reported phosphinates **42**, incorporating two terminal alkynes and thus offering protein cross-linking possibilities [69]. Applied to the disulfide rebridging of trastuzumab, half-antibody was observed as the main conjugation product; however, a side-reaction invariably competing with the expected inter-chain conjugation in these strategies (see §3.2.2. for further details).

Finally, a more recent application of activated alkynes for the rapid and chemoselective conjugation of cysteine residues in a myriad of peptides and proteins was reported by Waser and co-workers, who described two structurally related ethynylbenziodoxolones (EBX) reagents. Azide-bearing EBX reagent **44** was used for the introduction of azides into proteins in the form of conjugates **46** that could be sequentially modified by SPAAC and Suzuki–Miyaura cross-coupling reactions [70,71]. Interestingly, a different reactivity was observed with trimethylsilyl derivative **45**: following a preliminary *in situ* desilylation, thiol-yne β-addition was followed by an α-elimination of the iodide moiety leading to a transient carbene that underwent a 1,2-sulfur shift, resulting in the generation of a terminal thioalkyne **47** that could be further modified by CuAAC [72].

This quick overview of the two main chemoselective strategies shows how the field of protein bioconjugation has flourished in the recent years, offering now dozens—if not hundreds—of diverse ways to label selectively amines and thiols. In response to this booming expansion, the scientific community took an increasing interest in the domain, which fostered creativity and promoted the development of alternative methods for the conjugation of less nucleophilic amino acid residues.

2.2.3. Other residues

Tyrosine emerged as an attractive target because of its limited abundance at the protein surface and its electron-rich phenol(ate) side-chain. Even though original strategies have recently appeared [73], tyrosine conjugation has essentially relied on nucleophilic aromatic *ortho*-substitution since its early days in the 1900s and Hermann Pauly's work on the use of diazobenzenesulfonic acid—which would later be remembered as Pauly's reagent—for the colorimetric determination of tyrosine (and histidine) residues in proteins [74]. While far from being optimal, this study laid the groundwork toward the development of more chemoselective diazonium reagents, a quest that is still being pursued nowadays, more than a hundred years after their first report [75]. Diazonium reagents only represent the tip of the tyrosine-conjugation iceberg, as dozens of other strategies and reagents have been developed, all summarized in two recent, comprehensive and independent reviews by Chudasama and co-workers [76] and Gouin and co-workers [77]. Another alternative target for chemoselective conjugation was brought to light by the study of advanced glycation end-products, which showed that glyoxal—and 1,2-dicarbonyls in general—were selectively reacting with arginines' guanidinium groups [78–80], a strategy recently used by our team for the production of immunoconjugates [81].

Pursuing this continuous effort toward mastering the chemoselective conjugation of protein, several strategies have recently focused on the functionalization of less nucleophilic—and thus more challenging—residues, such as methionine or serine. However, because these residues tend to be less abundant at the surface of proteins than those previously discussed, the border between chemoselectivity and site-selectivity becomes more porous and less pertinent. For this reason, several chemoselective strategies will be discussed in the section dedicated to site-selective conjugation (see §4).

Altogether, the development of all these reagents and methods led to the generation of countless protein conjugates and myriad applications: the incorporation of fluorophores helped with the understanding of cellular mechanisms by analysing the trafficking of labelled proteins [82] polyethylene glycol (PEG) chains functionalization, also known as PEGylation, yielded less-immunogenic and more plasma-stable proteins [83,84]; conjugation of cytotoxic drugs to a carrier protein led to diminished systemic toxicity and facilitated their targeted delivery to malignant cells [85,86].

In the latter case, and in particular for antibody-drug conjugates (ADCs), the conjugation method was found to be one of the determinant parameters of the efficacy of the resulting conjugates [86,87]. While chemoselective strategies led to highly potent therapeutics—notably the marketed ADCs Mylotarg, Kadcyla and Besponsa, whose cytotoxic payloads were incorporated via lysine conjugation—it also generated complex and highly heterogeneous samples. Composed of regioisomeric mixtures of thousands of species with various degrees of conjugation, all of which possess different pharmacokinetics, chemoselective strategies are not the best ally in insuring batch-to-batch reproducibility and consistency of immunoconjugates production. This is a direct consequence of the large size of antibodies (approx. 150 kDa) and therefore of the elevated number of potential conjugation sites. Gautier *et al.* [88] best epitomized this issue in a seminal work in 2015, where they demonstrated that 69 of the 88 possible reactive amine groups (84 lysines and 4*N*-termini) on immunoglobulins G (IgG) could be conjugated with NHS esters. All things being equal, based on this reactivity profile this means that an IgG conjugated with six payloads would actually correspond to a mixture of approximately 120 million regioisomers.

Therefore, the next logical step for researchers who wanted to solve the heterogeneity issues of chemoselective protein bioconjugation was to develop *regioselective*—i.e. *site-selective*—methods and reagents; that is, methods and reagents that would lead to the conjugation of one precise residue among its many copies. This was highly desired for therapeutic applications, for which homogeneous conjugates could widen the therapeutic window and potentially improve the efficacy of the treatment [86,87,89]. Because site-selectivity can be seen as the next logical step after chemoselectivity in the ranking of complexity in selectivity—which could roughly be summarized as chemoselectivity < site-selectivity < protein selectivity—the knowledge gained during the study of protein reactivity in the former field helped develop the latter. Thus, and despite isolated examples showing the opposite [90], it is not surprising to note that the vast majority of site-selective strategies employ chemoselective reagents. These can be structurally tuned and improved for the purpose of single-site conjugation, or used as they are but under particular conditions (e.g. lowered temperatures, acidic conditions to favour *N*-terminus conjugation) or on proteins for which the distinction between chemoselectivity and site-selectivity is no longer relevant (i.e. proteins with a single solvent-accessible cysteine residue). Looking closer at the types of proteins employed in site-selective conjugation studies, two main families can be distinguished: artificial proteins—in which specific amino acids or sequences that are naturally absent from the parent protein have been introduced—or native proteins, whose primary sequences have not been edited.

3. Site-selective strategies

Site-selectivity and regioselectivity can be considered as synonyms and defined as the preferential conjugation of a single residue on the protein of interest (POI). While site-specificity is also often used in the literature to describe such a process [91–93], selectivity and specificity possess distinct definitions in the context of organic chemistry that should make them not interchangeable [94]. However, for the sake of clarity and as the majority of strategies discussed hereafter can be considered site-selective, we will *only* use this term whenever describing a reaction favouring one or a limited number of conjugation sites on the POI, irrespective of the structure of the reagent and of the reaction mechanism.

Currently, site-selective strategies are dominated by artificial proteins, in which an unnatural amino acid (UAA)—which should be understood in this case as a residue not naturally present in the parent protein—residue or peptide sequence has been incorporated via genetic engineering, and whose
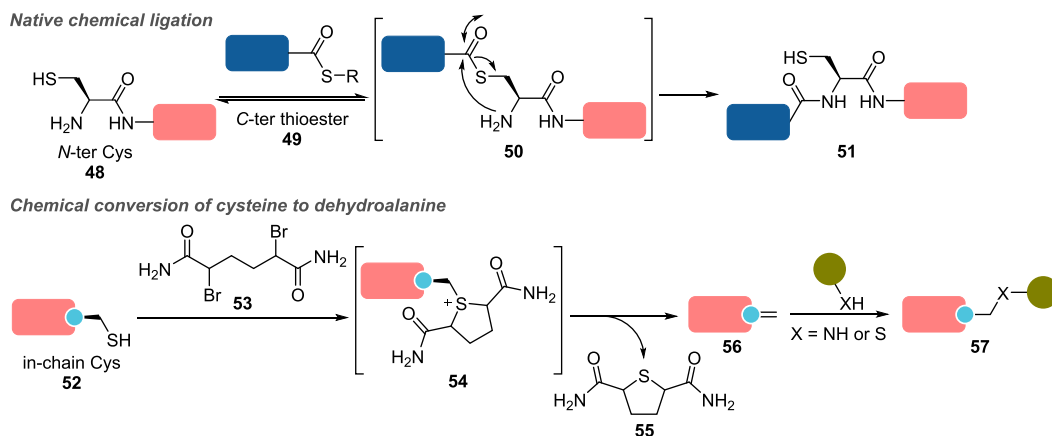
Figure 5. Selected examples of site-selective conjugation strategies using artificial proteins with engineered cysteine residues (R, alkyl or aryl).

reactivity is orthogonal to that of natural residues. Using tailored reactions, the selective labelling of these exogenous entities is greatly favoured and yields homogenous protein conjugates.

3.1. Site-selective labelling of artificial proteins

As previously discussed, three main strategies have been investigated for the site-selective conjugation of artificial proteins, depending on the nature of the exogenous group incorporated: a natural amino acid, a non-canonical amino acid and peptide or protein tags.

3.1.1. Incorporating natural α-amino acids

One of the first strategies explored was the introduction on the POI of poorly abundant yet natural amino acids, on positions where further functionalization would be possible without altering the structure and function of the biomolecule, such as N- or C-terminal positions, or on a loop [95].

Because of its high nucleophilicity and the poor abundance of free thiols in proteins, engineering of (seleno)cysteines by site-directed mutagenesis has been extensively employed [96,97]. Indeed, as cysteines are usually found as disulfide bonds in proteins, incorporation of an unpaired residue allowed the modification of the engineered protein via the numerous previously reported chemoselective methods (vide supra), but also triggered the development of new strategies.

N-terminal cysteines **48** and their 1,2-amino thiol motif have been selectively modified with thioesters for instance, as best exemplified by native chemical ligation (figure 5) [98,99], but also with tailored reagents bearing a reactive unsaturated carbon centre, such as 2-cyanobenzothiazole and diversely substituted 2-formylphenylboronic acids [95], as we will detail later in §3.2.3 and figure 16.

The development of a chemoselective di-alkylation/elimination cascade also permitted the chemical conversion of in-chain (seleno)cysteines **52** to dehydroalanines (Dha) **56** using 1,4-dihaloalkyl reagents such as **53** [100,101]. The resulting alkene moiety of Dha could then be selectively modified under mild conditions by appropriate nucleophiles, essentially via aza- and thio-Michael additions, to yield conjugates **57**. This strategy has notably been applied to the generation of chemically defined ADC, using THIOMAB antibodies, a class of artificial proteins developed by Genentech in which key native residues (in this case, the two valine residues V205 of the light chain) have been replaced with cysteines [102].

Besides (seleno)cysteines, other low-abundant amino acids have been engineered on proteins, such as tryptophan, tyrosine, methionine and histidine, but the limited number of chemoselective reactions to conjugate them coupled with competition with naturally occurring residues has limited their widespread utility [96].

Thus, of all natural amino acids, only engineered (seleno)cysteines found valuable applications in site-selective conjugation and the production of homogeneous conjugates. Albeit chemoselective, the conjugation strategies employed to target these residues still had to be finely tuned to limit side reactivity and reach good yields and conversion rates, some critical aspects that can be addressed by employing UAA and bioorthogonal conjugation reactions instead.
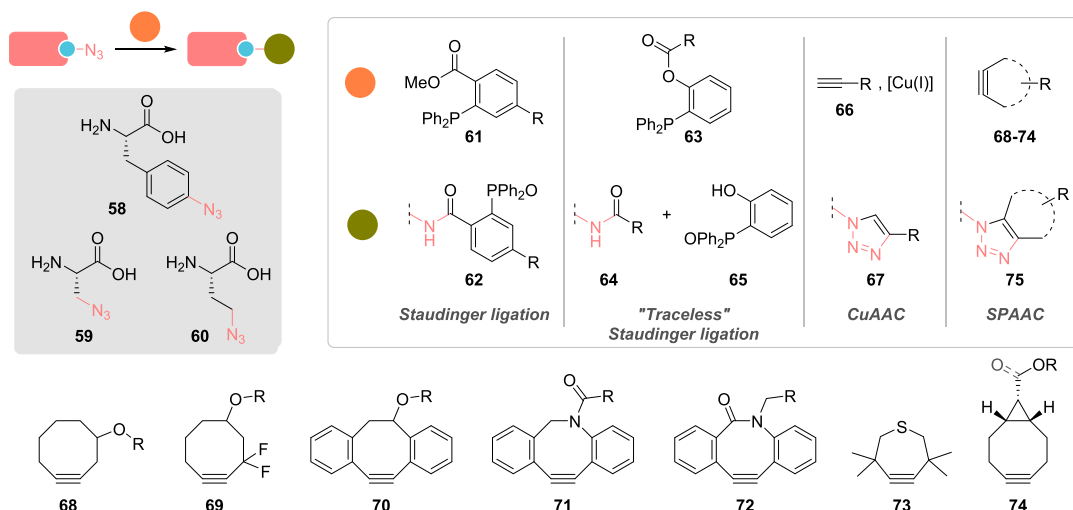
Figure 6. Incorporation of azido amino acids via genetic manipulation in proteins (R, alkyl). The introduced azido group can then be modified by Staudinger ligations, copper-catalysed alkyne-azide cycloaddition (CuAAC) and strain-promoted alkyne-azide cycloaddition (SPAAC).

3.1.2. Incorporating unnatural α-amino acids

The incorporation of a variety of UAA provided unique chemical handles for the site-selective modification of proteins [95,96,103–105]. Azide, alkyne, alkene, tetrazine and carbonyl-containing amino acids were designed and incorporated into proteins via genetic code expansion [106], offering the possibility to use bespoke reagents to modify these exogenous, unnatural moieties selectively without interfering with any natural functional group; in a *bioorthogonal* manner.

The Staudinger ligation between an azide and a triarylphosphine reported by Saxon & Bertozzi for the modification of azido-glycoproteins in living cells was the pioneering work demonstrating the vast potential of bioorthogonal reactions for the conjugation of unnatural proteins (figure 6) [107]. This technique and its variations, such as the traceless Staudinger [108], were then extended to the direct modification of mutant proteins that had incorporated *p*-azidophenylalanine **58** [109,110], azidoalanine **59** [111] or azidohomoalanine **60** amino acid residues [111]. However, because of slow kinetics (second-order rate constants classically found in the $10^{-3}\,M^{-1}\,s^{-1}$ range), the retained triarylphosphine oxide appendage in ligation products **62**, the sensitivity of phosphines **61** and **63** toward oxidation and the possibly competing iminophosphorane hydrolysis, the Staudinger ligation progressively lost its interest to the benefit of the copper(I)-catalysed alkyne-azide cycloaddition (CuAAC) and the selective formation of 1,4-triazole conjugates **67** [112–115]. As this rapid reaction (i.e. k_{obs} comprised between 10 and $100\,M^{-1}\,s^{-1}$) was found to be bioorthogonal and highly efficient under physiological conditions, it found myriad applications for the modification of recombinant proteins containing UAA with either azido or alkynyl substituents. Despite the problematic generation of harmful reactive oxygen species by the copper/ascorbate/atmospheric oxygen triad, effective procedures for the CuAAC labelling of live cells have nevertheless been reported [116].

As an alternative to the use of copper catalysts, Bertozzi and co-workers [117] had the revolutionary idea to adapt strain-promoted alkyne-azide cycloaddition (SPAAC) to the field of bioorthogonal chemistry in 2004, showing that cyclooctyne **68** could react rapidly at room temperature with azido-glycoproteins without the need for any catalytic system. Initially reported by Blomquist & Liu [118] before being studied in more detail by Wittig & Krebs [119], this facile reaction between cycloalkynes and azides results from the high reactivity of the bent triple bond, forced out of the ideal 180° bond angle by the cyclic and rigid framework. Because SPAAC kinetics were initially found to be similar to that of the slow Staudinger ligation [112], a number of diversely substituted strained alkynes were developed, playing essentially on stereoelectronic parameters, in order to accelerate the rate of this cycloaddition. Cyclooctynes and cycloheptynes were determined to be the best compromise between stability and reactivity, explaining the prevalence of these skeletons in the various reagents developed over the years; e.g. difluorinated cyclooctyne **69** (DIFO) [120–122], dibenzocyclooctynol **70** (DIBO) [123], dibenzoazacyclooctyne **71** (DIBAC) [124], biarylazacyclooctynone **72** (BARAC) [125], tetramethylthiacycloheptyne **73** (TMTH) [126–128] and bicyclo[6.1.0]nonyne **74** (BCN) [129]. While
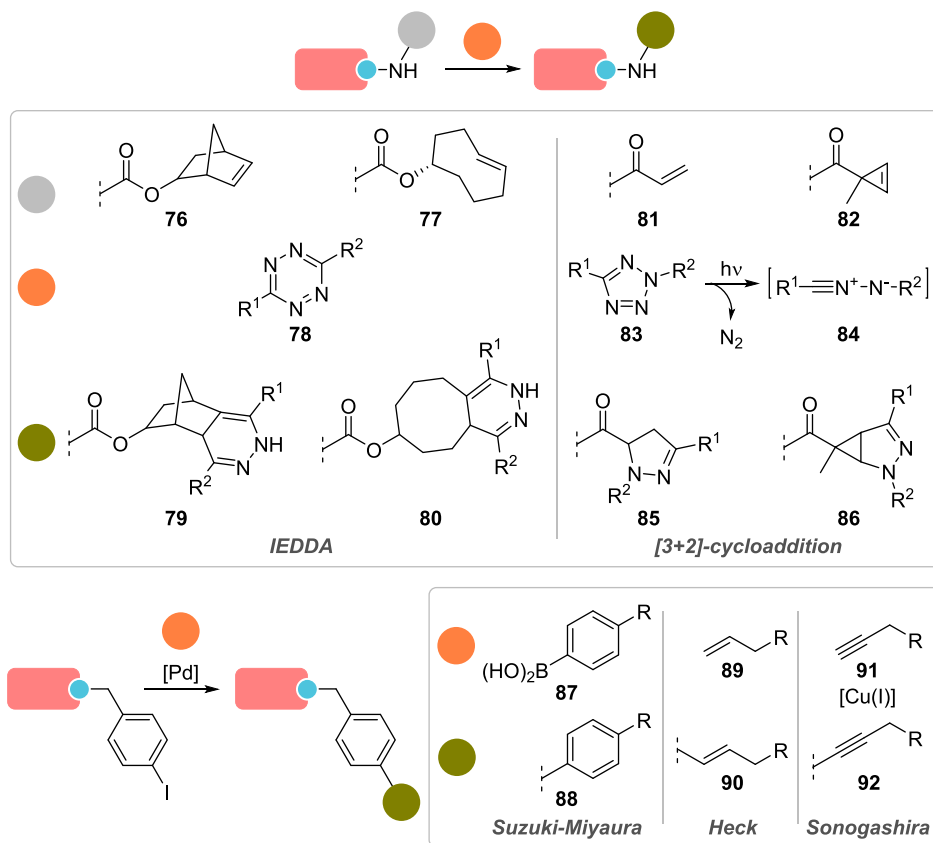
Figure 7. Site-selective modification of UAA on artificial proteins using bioorthogonal reactions. UAA 76, 77, 81 and 82 are lysine derivatives; UAA with *p*-iodobenzyl side-chain are either cysteine or phenylalanine derivatives 121 and 116, incorporated either chemically or by genetic code expansion, respectively.

cycloalkynes are essentially used as small molecular probes to functionalize azide-containing artificial proteins, the reverse approach has also been documented with cyclooctyne-containing UAA being genetically encoded into proteins [130,131].

Parallel to the development of SPAAC for the labelling of proteins, inverse-electron demand Diels–Alder reactions (IEDDA) also emerged as a new promising tool (figure 7). Reacting electron-rich dienes, such as *trans*-cyclooctenes or norbornenes, with tetrazine dienophiles led to exceedingly fast bioorthogonal reactions with second-order rate constants up to $10^5 \, \text{M}^{-1} \, \text{s}^{-1}$ [132–134]. New reactive UAA such as **76** and **77** were thus developed and applied to the conjugation of several artificial proteins, including cell-surface, cytosolic and nuclear proteins in live cells [131,135–138].

Another approach to fast bioorthogonal reaction for site-selective protein modification employs the [3 + 2]-cycloaddition between a transient nitrile-imine **84**, generated *in situ* by photo-irradiation of tetrazole **83** and a variety of alkenes, such as **81** and **82** (figure 7) [139]. This spurred the development of various alkenyl and tetrazole UAA (see **104–106**; figure 8), which were site-specifically incorporated and labelled in various proteins in live cells [140–143]. While second-order rate constants of this [3 + 2]-cycloaddition were found to be three to four orders of magnitude lower—*viz.* between 10 and $100 \, \text{M}^{-1} \, \text{s}^{-1}$—than that of IEDDA, it possesses the advantage of a possible spatial and temporal control, as UV irradiation is required to trigger the formation of the reactive nitrile-imine species.

Drifting away from cycloadditions, incorporation of carbonyl groups in recombinant proteins has also been investigated due to their limited occurrence in native proteins and their selective modification with α-nucleophiles, such as alkoxyamines and hydrazines, to form oximes and hydrazones, respectively. Despite not being bioorthogonal *per se*—amines condensation to form imines is also possible—this strategy has still found several applications in chemical biology over the years [144]. While site-selective generation of terminal aldehydes could be easily achieved on proteins possessing *N*-terminal serine or threonine residues, via periodate-mediated oxidative cleavage [145], aldehyde and ketone-containing UAA have also been designed [146,147], such as *p*- and *m*-acetyl-
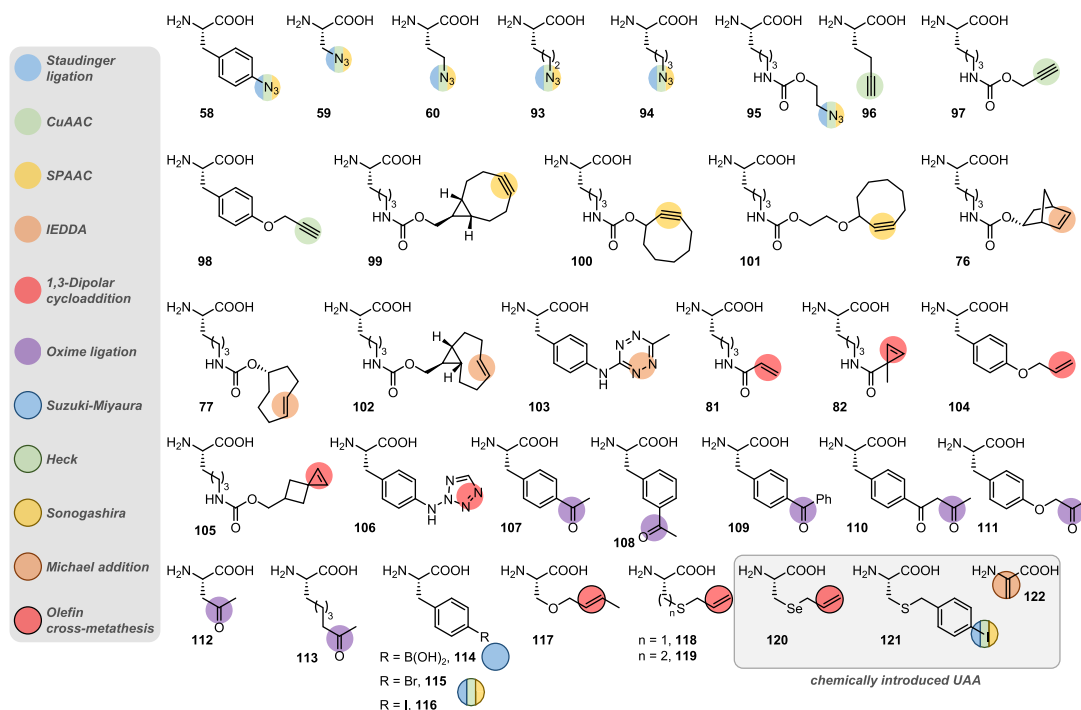
Figure 8. Examples of unnatural amino acids that have been incorporated into artificial proteins alongside the bioorthogonal technique employed to functionalize them.

L-phenylalanine **107** and **108**, respectively [148–154], p-benzoyl-L-phenylalanine **109** [150,152] and 2-amino-8-oxononanoic acid (KetoK) **113** (figure 8) [155]. However, the slow kinetics associated with carbonyl condensation—second-order rate constants ranging from 10^{-4} to $10^{-3}\,M^{-1}\,s^{-1}$, making this transformation the slowest of all bioorthogonal reactions—coupled with the need for acidic pH to achieve acceptable conversion have limited the use of this labelling technique.

Besides CuAAC, numerous other transition-metal catalysed strategies for the modification of UAA have surfaced in recent years (figure 7) [103,156]. Associated with excellent chemoselectivity, bioorthogonality and high yields, palladium-catalysed cross-couplings between aryl/alkenyl halides and a variety of reactive partners such as boronic acids **87** (Suzuki–Miyaura) [157–165], alkenes **89** (Heck) [166–168] and terminal alkynes **91** (Sonogashira) have been described for site-selective protein conjugation [166,168,169]. Early use of these reactions was hampered by technical difficulties as they needed to proceed in aqueous media and under low protein concentration, necessitating elevated temperatures in certain cases and high catalyst loading that would lead to protein degradation and side reactions, explaining the low yields that often accompanied these cross-couplings [157,158,167,168]. Improvement in this field has essentially been realized thanks to the development of more effective catalysts. The Davis group reported a phosphine-free palladium complex that promoted Suzuki–Miyaura cross-couplings between various boronic acids and the p-iodobenzyl cysteine residue (chemical incorporation by cysteine alkylation; see **121**, figure 8) of a mutant protein with excellent conversion (greater than 95%) and under mild conditions (pH 8.0, 37 °C). In a subsequent study, they also addressed the issue of non-specific palladium coordination to proteins in Suzuki–Miyaura cross-couplings by using 3-mercaptopropionic acid as a scavenger in order to monitor the conjugation of a p-iodophenylalanine-modified maltose binding protein by mass spectrometry [159,160].

In addition to copper and palladium, ruthenium completes the podium of most-employed transition metals for the conjugation of recombinant proteins, essentially in the form of the Hoveyda–Grubbs II catalyst to promote olefin cross-metathesis (CM) of allylic groups-containing UAA **117–120** [103,170]. (On a side note, several mentions of ring-opening metathesis polymerization (ROMP) can also be found in the literature regarding protein-polymer conjugation [171]; however, the metathesis step was essentially used to generate polymers and not to conjugate them to proteins, thus falling out of the scope of this review.) Benefitting from the strong allylic chalcogen effect observed in olefin CM, S-allylcysteine (Sac) **118** and its selenocysteine analogue (Seac) **120** were identified as excellent substrates for site-selective conjugation of artificial proteins. In general, these UAA were chemically introduced in place of a cysteine residue via direct allylation or transient Dha formation as previously discussed (see **56**,
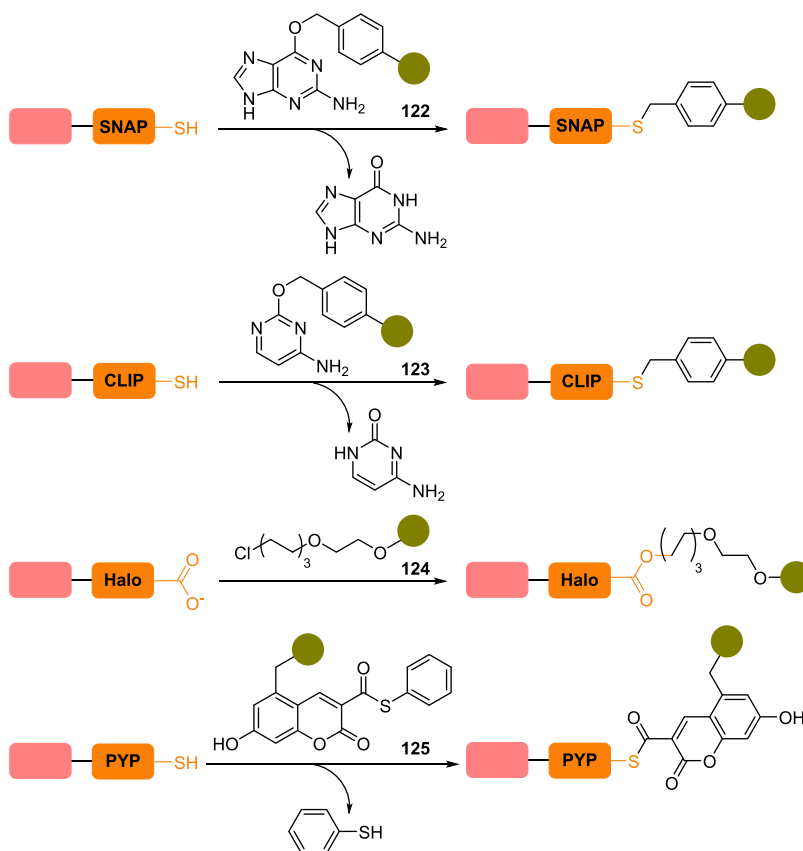
Figure 9. Use of protein tags for the site-selective conjugation of chimeric proteins.

figure 5 and **122**, figure 8), but genetic incorporation of diverse UAA with alkenyl side-chains has also been reported [172,173]. Experimentally, the resulting artificial proteins were conjugated with a wide panel of allylic reagents in rates similar to that of SPAAC for Seac substrates (i.e. approx. $0.3 \, M^{-1} \, s^{-1}$) [174]. Interestingly, addition of magnesium(II) salts to the buffer was found to be critical for successful on-protein olefin CM, in order to prevent unproductive binding of the ruthenium catalyst to the protein surface, as did the use of *t*-BuOH as a co-solvent to facilitate catalyst dissolution [175].

To date, dozens of UAA have been successfully incorporated into proteins—a selection of which is presented in figure 8, listed according to the bioorthogonal technique employed for their functionalization—almost exclusively by genetic code expansion. Using bespoke reagents, bioorthogonal and site-selective conjugation of these artificial proteins has been made possible in live cells, highlighting the strong advantages of these techniques. While chemical strategies for the modification of UAA are plentiful, peptide and protein tags offer a complementary and interesting alternative as they open the door to enzyme-mediated, site-selective conjugation of artificial proteins.

3.1.3. Incorporating enzymatic recognition sequences

Spurred by the powerful applications of bioorthogonal conjugation of UAA-containing artificial proteins, researchers also investigated the genetic engineering of fusion proteins, i.e. chimeric proteins resulting from the assembly of a POI with a peptide or a protein tag that can be covalently and selectively labelled with a small molecule. This selectivity usually comes from the repurposing of a known biological activity of the protein tags, which are essentially derived from enzymes in the context of bioconjugation. The usual method for producing these fusion proteins is via the construction of a corresponding recombinant plasmid, then expressed in different types of cells, thus offering access to live cell labelling applications.

As a prime example, the group led by Kai Johnsson reported in 2003 the use of a mutant enzyme—[W160]hAGT—derived from the human O^6-alkylguanine-DNA alkyltransferase (hAGT), a DNA repair protein that acts as a scavenger of O^6-alkylated guanine nucleobase by transferring the alkyl group to one of its reactive cysteine residues (figure 9) [176]. Taking advantage of this transformation, O^6-benzylguanine derivatives **122** were designed to label rapidly (rate constants up to $10^4 \, M^{-1} \, s^{-1}$) and

selectively engineered hAGT and fusion proteins thereof, leading to a plethora of applications, including *in vivo*, under the trade name SNAP-tag [177]. In the same vein, the CLIP-tag was also developed, which could be functionalized with *O*-benzylcytosine derivatives **123** on a reactive cysteine residue [178]. Interestingly, despite similar structural features, SNAP-tag and CLIP-tag were shown to be orthogonal to one another, allowing the simultaneous and selective labelling of two different fusion proteins.

Switching to another family of fusion proteins, the Halo tag was developed based on *Rhodococcus* haloalkane dehalogenase (DhaA), a bacterial enzyme possessing a reactive aspartate residue (D106) that could be selectively labelled with chloroalkane substrates **124** [179]. Carefully designed mutants were produced, which helped to improve the recombinant enzyme's binding rate (up to a whopping $10^6 \, M^{-1} \, s^{-1}$) and led to hydrolytically stable conjugates. Bacterial enzymes are also a source of inspiration for the development of other fusion proteins: the photoactive yellow protein (PYP) tag [180,181], which can be selectively functionalized on its reactive cysteine residue C69 by *trans*-thioesterification of coumarin derivatives **125**, or non-catalytic β-lactamases that can be site-selectively acylated on the serine residue S70 [182].

Although the use of fusion protein proved to be a very powerful strategy (e.g. fast labelling kinetics, site-selective conjugation, easy access to structurally diverse payloads, real-time analysis of fast biological processes in living cells), the molecular weight of these chimeric assemblies appeared to be too high in certain cases and altered some key properties of the POI (e.g. structure, function, dynamics, localization) [104].

To bypass this issue, shorter peptide tags (0.6–6 kDa) were also designed in parallel and installed on the *N*- or *C*-terminus of the POI (figure 10). It was thought that their smaller size would limit their detrimental impact on the protein's structure and function [104]. In a seminal work, Tsien and co-workers [183,184] reported a tetracysteine tag **126** (CCXXCC; X = any proteinogenic amino acid) able to bind bis-arsenic fluorophores **127** with high selectivity and affinity. Based on a similar design, the tetraserine tag SSPGSS **128** was developed to react selectively with bis-boronic dyes **129** to form macrocyclic boronates [185]. Unlike tetracysteine tag **126**, whose motif is absent in native proteins, the SSPGSS sequence can, however, be found in more than 100 native human proteins, thus affecting the protein selectivity level of this tag in live cells. In 2007, Hamachi and co-workers produced the undecapeptide tag $CysAla_6Asp_4$ (CA_6D_4) **130** in which the tetra-aspartate motif can selectively coordinate a zinc(II)-containing electrophilic probe of general structure **131**, helping to direct its reactive α-chloroacetyl motif (highlighted in blue) close to the tag's cysteine residue, resulting in a site-selective alkylation of the peptide tag. Similarly, polyhistidine tags have been employed to complex a tosylate-equipped nickel(II) probe to direct the selective alkylation of one of the histidine tag imidazole, but metal cation-free methods employing similar tags have also been reported [186,187]. More recently, Pentelute and co-workers [188] reported the shortest reactive tag sequence to date, the *C*-terminal FCPF tetrapeptide **132** functioning as a π-clamp and allowing the selective labelling of the tag's cysteine with perfluoroaromatic reagents **133**. It is worth stressing that the vast majority of these conjugation reactions occur under mild conditions—a prerequisite when working with live cells—and in minutes to hours.

Albeit falling slightly out of the scope of this review, one can also say a few words about the modification of peptide tags by formylglycine-generating enzyme (FGE). In eukaryotes, FGE in combination with a copper cofactor catalyses the conversion of a cysteine residue to an aldehyde-containing formylglycine (fGly) residue in the minimal consensus sequence CXPXR (X = any proteinogenic amino acid but proline) of sulfatases [189,190]. Taking advantage of the promiscuity of FGE, the incorporation of the aldehyde tag CXPXR—usually in the form of the hexapeptide LCTPSR **134**, the wild-type sequence found in human type I sulfatases—in various fusion proteins, including antibodies, allowed the smooth FGE-catalysed installation of fGly residues (see **135**, figure 10), which were then selectively functionalized via oxime or hydrazone ligation [191–193]. Unlike enzyme-mediated bioconjugation techniques (e.g. using sortase A, transglutaminases, protein farnesyltransferase), which will not be discussed in this review, the enzyme in this particular case only *assists* in the conjugation of the POI by converting it to a transient, reactive electrophile but does not *participate* in its subsequent derivatization, making conjugation a well and truly chemical step.

Thus, site-selective chemical conjugation of artificial proteins offered prominent results thanks to the incorporation of natural or UAA, either as single residues or as part of peptide and protein tags. However, generation of such proteins can be a costly and time-consuming process. Coupled with a restricted access to only a certain degree of conjugation (classically from 1 to 2) and potential immunogenicity issues due to the artificial incorporation of amino acids or peptide sequences in the POI, bioconjugates of artificial proteins present some limitations that prevent their widespread use,
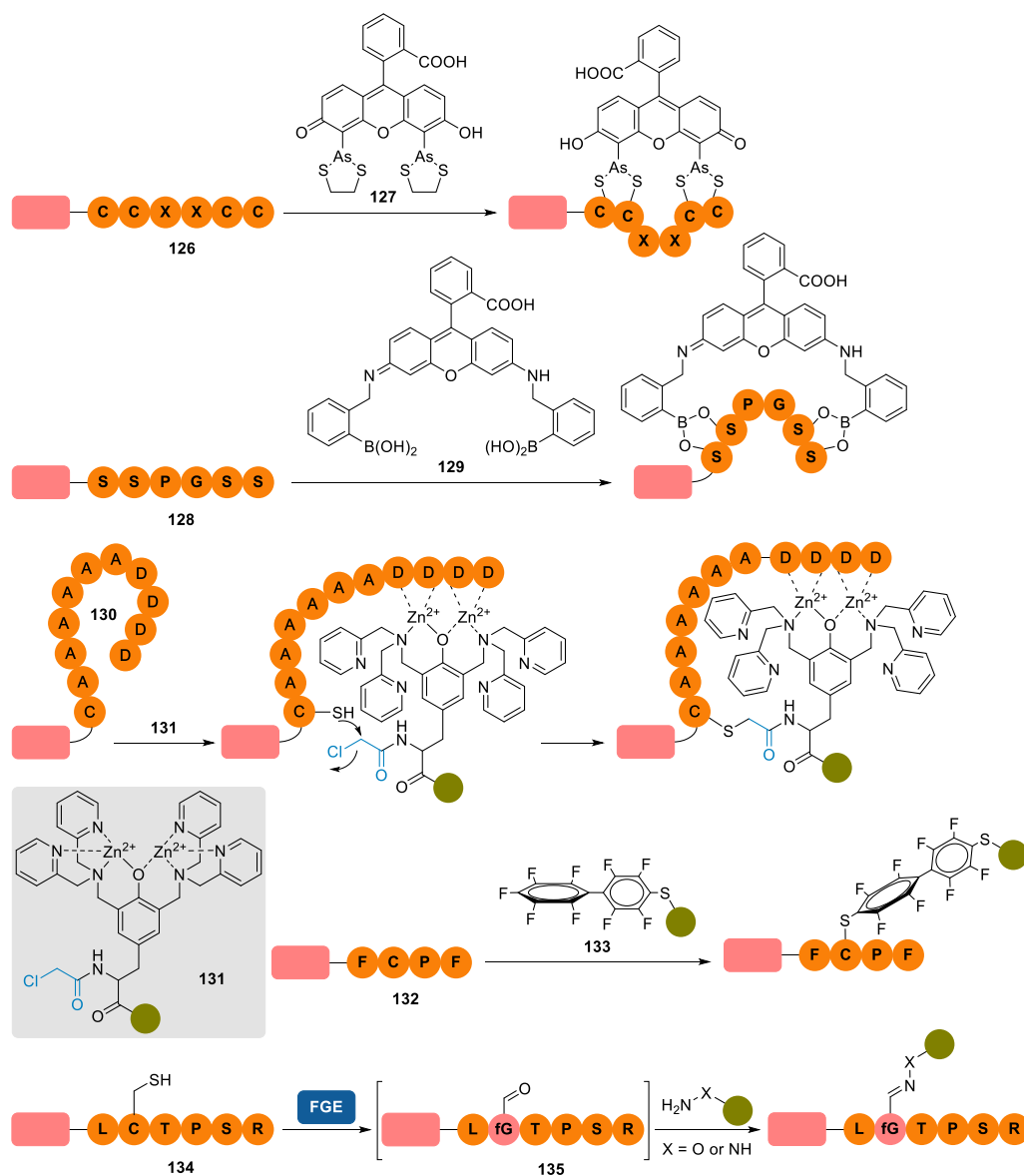
Figure 10. Installation of small peptide tags on the *C*-terminus of a protein of interest and their selective labelling with bespoke probes.

undermining their panacea-like image. While far more challenging chemically speaking, site-selective methods for the conjugation of native proteins can be seen as a way to overcome these limitations.

3.2. Site-selective labelling of native proteins

As an alternative to artificial proteins, the development of site-selective strategies for the conjugation of native proteins has emerged over the past few years. Compelling results have been obtained *in vitro*, in complex biological environment, notably thanks to ligand-directed proximity-driven conjugation [194]. However, these would not be discussed in the following sections, as the observed site-selectivity originates from a preliminary molecular recognition between a protein and its ligand and not from the reactive chemical component itself. Thus, the chemical strategies about to be discussed aim to be applicable to a broad range of native proteins, and have been designed to target various types of amino acids. Unsurprisingly, the most abundant lysine has been one the most studied residues.

3.2.1. Lysine residues

As previously mentioned, lysines are both abundant and surface-exposed in the vast majority of proteins, thus complicating their site-selective conjugation. However, some approaches have been proposed to

tackle this challenge on native proteins, relying on kinetic control, bespoke electrophilic reagents, multicomponent reactions (MCR) or template-directed approaches (figure 11).

One of the most common strategies for the chemoselective modification of lysines is based on the use of activated esters (see §2.2.1.). In 2012, Weil and co-workers [195] showed that a site-selective variation of this approach was possible, with the successful modification of RNase A and lysozyme C via a kinetically controlled labelling. Portionwise addition of sub-stoichiometric amounts of NHS-biotin derivative **136**—i.e. 0.005 equivalent per minute, over 100 min—led to the modification of a single lysine residue on the POI, K1 in both cases, which was explained by stereoelectronic effects: best solvent accessibility in the case of RNase A; hydrogen bond between K1 side-chain's ammonium and a neighbouring carboxylate group, increasing K1 nucleophilic character in lysozyme C. As expected from the conjugation conditions, site-selectivity occurred to the detriment of partial conversion and limited yield, even though this could be overcome by affinity purification on avidin column of the biotin-containing monoconjugates.

Moving on to bigger proteins, Cellitti and co-workers [30] studied the single labelling of K188 on the Fab fragment of the monoclonal antibody trastuzumab (50 kDa). Building on a previously published patent reporting the use of pentafluorophenyl esters **138** to achieve this single residue conjugation (up to 70% labelling) [196], they showed that lowering the temperature to 4 °C or using flow chemistry to perform the conjugation helped to increase the yield and reach nearly quantitative labelling of this residue. Interestingly, site-directed mutagenesis studies highlighted the extreme importance of two neighbouring residues, H189 and D151, whose mutation to alanine residues led to little to no conjugation of the resulting mutants. While it was hypothesized that the single labelling of K188 was explained by the formation of a triad charge-relay system, no definitive mechanism was proposed by the authors.

The fact that the local microenvironment of a residue can influence its reactivity is at the basis of how enzymes exert their biological role. In the case of lysine, it has been found that ε-amino groups located at internal positions in proteins could have pK_a values as low as 5.3, more than 5 units lower than the classical 10.4 value observed in water [197]. Being less prone to being protonated at physiological pH, such residues represent an attractive target in the context of site-selective conjugation, leading to the development of tailored reagents.

Barbas and co-workers [198] showed that α,β-unsaturated sulfonamides **140**—and their CuAAC-functionalized derivatives—could selectively modify the K64 residue of human serum albumin (HSA) in human plasma, leading to exceedingly stable adducts *in vitro* ($t_{1/2} > 28$ d in human plasma at 37 °C) and *in vivo* ($t_{1/2} \sim 40$ h in BALB/c mice). The influence of HSA's tertiary structure on the reactivity of K64 was highlighted by an absence of reaction between **140** and the simple lysine amino acid, even though no further theory was proposed to explain this observation.

Influence of local microenvironments on α,β-unsaturated sulfonamides reactivity and lysine conjugation was further supported by Bernardes and co-workers [199] a few years later. In a seminal work published in 2018, they reported the use of sulfonyl acrylate reagent **142**, reminiscent of Barbas' sulfonamide **140**, for the site-selective modification of the most reactive lysine residue in various proteins. It was anticipated that lysines' ε-amino groups with the lowest pK_a—due to their location in a rather hydrophobic microenvironment—would be the most nucleophilic and could react selectively with **142** via a chair-like H-bonded transition state. Comparison between theoretical predictions of pK_a and peptide mapping analyses run on a handful of monoconjugated proteins seemed to corroborate this postulate: K573 (theoretical pK_a 10.39) was identified as the sole conjugation site in HSA, K33 (th. pK_a 9.48) in lysozyme, K100 (th. pK_a 10.13) in the C2A domain of Synaptotagmin-I (C2Am) and K300 (th. pK_a 10.35) in Annexin V. Interestingly, these findings contrast with that of Barbas (*vide supra*) and Rai (*vide infra*), who identified K64 in HSA and K116 in lysozyme as the sole conjugated lysine residue, respectively, showing that the reagent structure has a strong influence in the site-selectivity of the conjugation, and not only the residue's accessibility and pK_a value.

Taking advantage of the *in situ* generation of formate ester **144** from the oxygen-mediated oxidation of benzaldehyde **145**, Rai and co-workers [200] reported the chemoselective and site-selective formylation of various proteins: RNase A (K37), melittin (K21), ubiquitin (K38), lysozyme C (K116), myoglobin (K145), cytochrome c (K5, albeit N-terminus acetylation was necessary in this case to achieve site-selective conjugation) and histone (K37), with moderate to good conversion (50–82% after 24–48 h of reaction). Interestingly, all formylated lysines were found to be located in ordered secondary structures (i.e. α-helix, loop, β-sheet) and not necessarily the most solvent-accessible ones.

Drifting away from the classical use of reactive electrophiles, the same group also opted to investigate MCR for the site-selective conjugation of lysines. Using either formaldehyde **147** or benzaldehyde
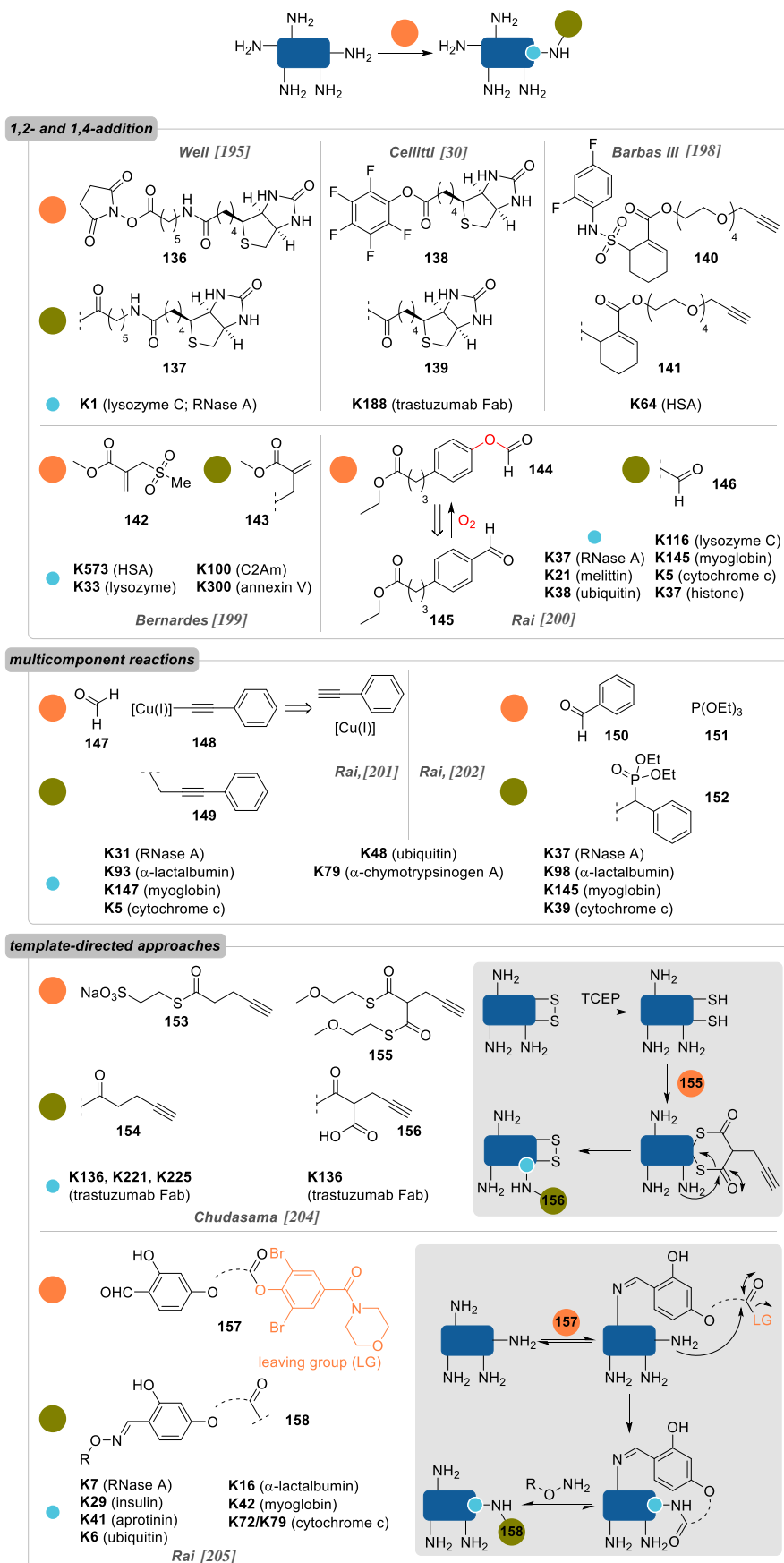
Figure 11. Overview of site-selective chemical conjugation of native proteins on lysine residues (R, alkyl chain with payloads of interest; TCEP, *tris*(2-carboxyethyl)phosphine).

derivatives **150**, ε-amino groups were first converted to latent imines, which reacted in a subsequent step with carefully selected nucleophiles—either *in situ* formed copper(I) phenylacetylide **148** (A^3 coupling) [201] or triethylphopshite **151** (phospha-Mannich) [202], respectively—with excellent chemo- and site-selectivity. Of the various proteins evaluated in both studies, some were shown to be labelled on the same lysine residue, irrespective of the MCR employed under identical reaction conditions: ubiquitin (on K48) and α-chymotrypsinogen A (K79). The majority, however, showed dissimilar conjugation sites depending on the strategy employed—RNase A (phospha-Mannich: K37; A^3 coupling: K31), α-lactalbumin (phospha-Mannich: K98; A^3 coupling: K93), cytochrome C (phospha-Mannich: K39; A^3 coupling: K5), myoglobin (phospha-Mannich: K145; A^3 coupling: K147)—stressing even more the prevalence of the reagent structure on the conjugation site.

As proposed by Gothelf and co-workers in 2014, another strategy for the site-selective conjugation of native proteins employs template-directed approaches. They took advantage of histidine-rich regions in different proteins to form transient metal complexes with divalent cations, helping to direct an NHS ester equipped with a *tris*-nitrilotriacetic acid chelating handle to label only a handful of neighbouring lysines [203]. This idea of using another amino acid as a temporary anchor point to direct—via electrostatic interactions or covalent bond formation—the modification of a neighbouring lysine residue can also be found in subsequent studies. Chudasama and co-workers [204] showed that cysteines could act as valuable relays for the site-selective modification of lysines on the Fab fragment of trastuzumab with thioesters **153** and **155**, in a strategy reminiscent of native chemical ligation (figure 11). After reduction of the disulfide bond of the Fab fragment, a first *trans*-thioesterification between the cysteine thiol and a thioester was followed with a 'cysteine-to-lysine transfer' (i.e. *S,N*-acyl migration) on a proximal lysine. While this first approach could not be coined site-selectively *per se*, with three labelled lysines in the vicinity of the disulfide bond being identified (K136, K221 and K225), the use of bis-thioester **155** for the formation of an intermediate bridged conjugate led to the sole conjugation of K136 (80% conversion), presumably thanks to conformational effects.

More recently, Rai and co-workers [205] pursued their effort toward site-selective conjugation of native proteins and described a lysine-directed lysine modification method. They designed a library of reagents of general structure **157** bearing two reactive groups, an aldehyde and a phenol ester, separated by a spacer of variable length. The aldehyde first led to the rapid but reversible formation of imines with multiple solvent-accessible lysines, helping to direct the intramolecular acylation of a neighbouring lysine by the phenol ester in a slow but irreversible second step. Addition of an *O*-alkoxyamine solution in a final step permits the removal of all remaining imines while functionalizing the POI in the form of conjugates **158**—a procedure that also gave the possibility of trapping and isolating pure mono-labelled conjugates on hydrazide-activated resins before a trans-tagging release with *O*-alkoxyamine probes. Site-selectivity was achieved on different proteins by varying the chemical structure of the spacer: RNase A was selectively modified on K7 (34% conversion), insulin on K29 (48%), aprotinin on K41 (86%), ubiquitin on K6 (35%), α-lactalbumin on K16 (37%) and myoglobin on K42 (26%). Interestingly, varying the spacer also led to different labelling sites on cytochrome C, either K72 or K79 (33% versus 28% conversion, respectively), offering a valuable modularity to this strategy.

Compelling results have thus been obtained on a panel of native proteins of diverse molecular weights—from as low as 2.8 kDa to greater than 66 kDa—via different approaches. As is apparent from these studies, the microenvironment of the lysine and the chemical structure of the electrophile are both equally important, and the slightest variation in one or the other parameters can drastically alter the site-selectivity of the strategy. While labelled lysine residues were systematically identified by peptide mapping studies in almost all studies, difficulties in identifying the precise conjugation site seem to have emerged with full antibodies, possibly because of their high molecular weight and their numerous lysines. Focusing instead on cysteine conjugation for full antibodies helped to overcome this limitation.

3.2.2. Cysteine and cystine residues

As previously mentioned, cysteines with free side-chain thiols are relatively scarce in nature, forming usually homo- or hetero-dimers via oxidized disulfide bonds, such as the cystine residues bridging heavy and light chains in antibodies. While reduction of the inter-chain disulfide bonds of the most abundant IgG1 can give access to eight free thiols, their chemoselective labelling can have a detrimental impact on the structure, shape, stability and activity of the protein, while still leading to heterogeneous mixtures of conjugates due to incomplete conversion [206]. Thus, chemical methods to conjugate covalently free cysteines 'in pairs', effectively rebridging the disulfide bond and cross-linking the protein, has been proposed as a way to address this issue [207].

Brocchini and co-workers [208,209] were the first ones to reconstruct covalent bonds between two reduced cysteines in proteins by forming a three-carbon bond between the two thiols with monosulfone cross-linking reagents **159B** deriving from bissulfone precursor **159A** by simple incubation at pH 7.8 (figure 12). After disulfide reduction, a successive thio-Michael addition/sulfinic acid elimination/ second thio-Michael addition (see mechanism in sidebar) allowed the facile PEGylation of interferon α-2b, as well as L-asparaginase and the Fab fragment of the monoclonal antibody Q4120/ADP318, without any apparent disulfide scrambling. Building on this work, monosulfone reagent **160B** was successfully applied as a rebridging reagent for the production of homogeneous trastuzumab Fab fragment conjugates [210]. The presence of a tetrazine handle allowed facile functionalization of the resulting conjugates with *trans*-cyclooctene-equipped payloads (i.e. fluorophores and short PEG).

Applied to the intact antibody, MMAE-sulfone reagent **161B** led to the favoured formation of a DAR4 conjugate (78%), showing both *in vitro* and *in vivo* targeted cytotoxicity [211].

Next-generation maleimides (NGM), such as dibromomaleimides **162** and **163**, and dithiophenolmaleimides **164** and **165**, were also employed as rebridging reagents of disulfide bonds [212–215]. The incorporation of two nucleofuges across the maleimide double bond allowed the consecutive addition of two thiols via a similar addition–elimination–addition sequence as previously described for sulfones **159B–161B**. These reagents helped construct numerous ADCs, some with improved properties over conventional maleimide-conjugated species in terms of efficacy, toxicity and pharmacokinetics. Excellent yields and conversion have been reported, with DAR 4 being the sole immunoconjugates detected in certain cases, albeit concurrent generation of half-antibody isoforms via intra-chain bridging and disulfide scrambling could systematically be observed.

In 2011, Caddick and co-workers [55] reported the conjugation of proteins and the bridging of cyclic peptides with dibromopyridazinedione **166** (PD). While being structurally related to NGM, PD possess the advantage of presenting four points of attachment: the two sp^2 carbons of the alkene moiety, reacting chemoselectively with free thiols through two consecutive addition–elimination steps, and the two pyridazine nitrogens, grafted with different bioorthogonal tags allowing the dual labelling of the POI. A fewyears later, Caddick and co-workers [216] reported the use of PD **167** for the reconstruction of inter-chain bonds of the full antibody trastuzumab. Since the reagent was designed to bear two orthogonal handles, a strained and a terminal alkyne, the rebridged antibody could be further functionalized via successive SPAAC and CuAAC, yielding immunoconjugates with two different payloads per reconstructed disulfide bond.

Based on these results, new PD reagents were also designed to afford homogeneous antibody conjugates with lower DAR. DAR4 conjugates were obtained by only adding a single functionalization site on PD, like the strained alkyne in PD **168** or the drug MMAE in PD **169**, resulting in the generation of a panel of PD with very effective rebridging abilities—less than 5% of half-antibody was generally detected, and homogeneous DAR4 ADC could be produced in excellent yield [217,218]. Connecting two PD together with a PEG linker—whose length was critical for the successful outcome of the strategy—allowed the simultaneous rebridging of two disulfides on trastuzumab, while the incorporation of one (**171**) or two (**170**) terminal alkynes per PD dimer offered a selective access to DAR2 or DAR4 conjugates, respectively [219]. Further optimization of pyridazinediones led to valuable applications, such as the possibility of conducting disulfide reduction and rebridging in a single step with 'two-in-one' reagents [220], or the generation of homogeneous bispecific Fab fragments conjugates [221].

Another rebridging reagent, the dibromo oxetane derivative **172**, was proposed in 2017 by Bernardes and co-workers [222]. Albeit being deprived of a functional handle for further derivatization of the resulting protein conjugates, dibromooxetanes proved to be effective for the stapling of cysteine residues in different proteins, including antibody fragments. Intriguingly, replacing the oxetane motif with an alkene altered the reactivity of the electrophile, leading only to negligible stapling [223].

Thiol-yne reaction was also proposed as a disulfide-rebridging photochemical strategy by Griebenow *et al.* [224] (figure 12). After reduction of the Fab fragment of the antibody M14-G07, the rebridging was carried out under UV light at 365 nm in the presence of 6-heptynoic acid **174** and lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) as radical initiator. While the presence of covalently disulfide-stapled Fab was demonstrated, conversion was poor (40%) and the reaction necessitated elevated antibody concentrations to be effective (46.1 mg ml^{-1}).

Structurally related to this strategy, addition to electron-deficient alkenes—such as divinyl triazine **176**—and alkynes was also explored by several groups, including ours, developing various rebridging reagents that were tested on native proteins—mostly full antibodies [225–228]. While potent ADCs could be produced in this way, disulfide scrambling was systematically observed, as evidenced by the formation of half-antibodies in all cases.
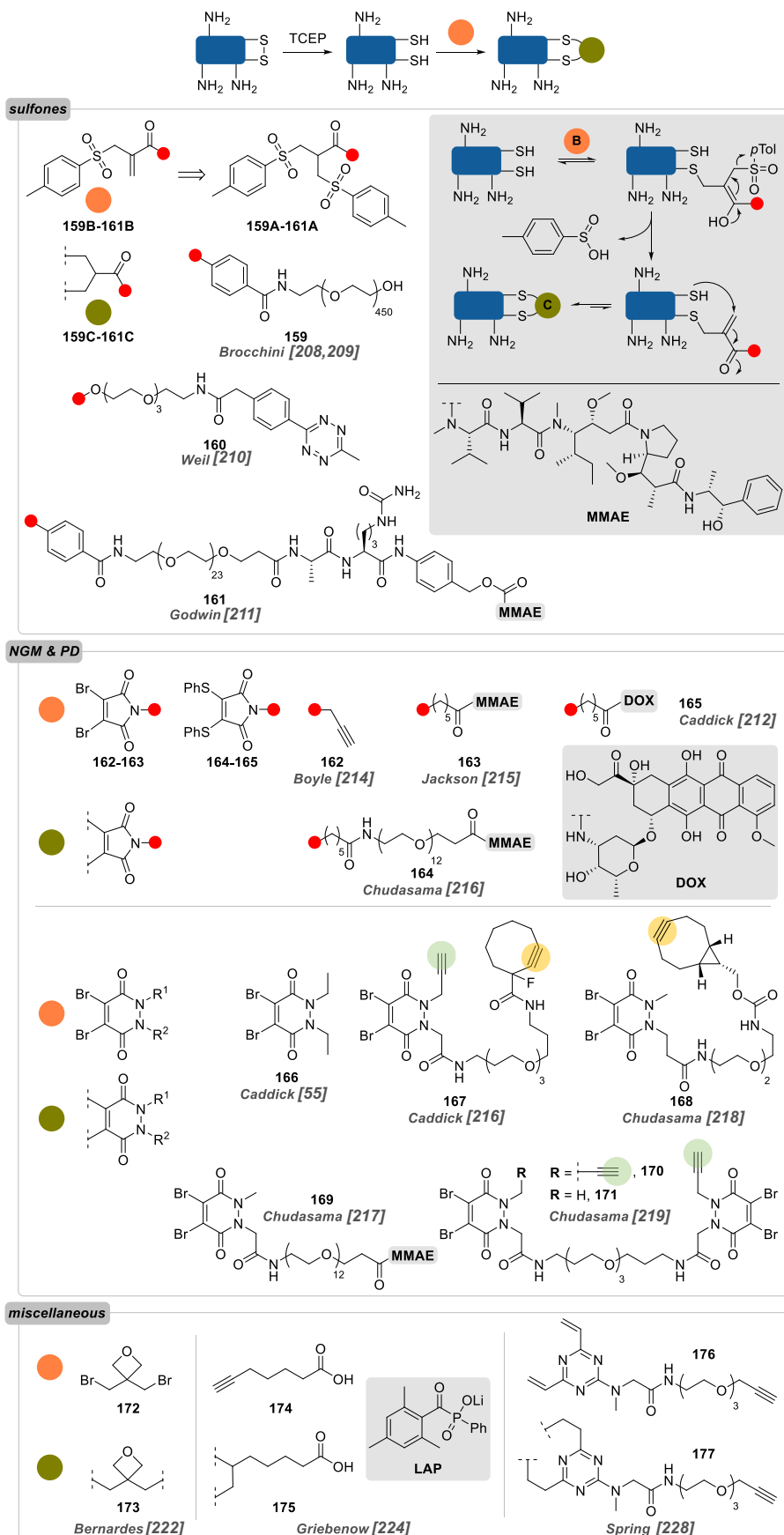
Figure 12. Overview of site-selective chemical conjugation of native proteins on cystine (TCEP, *tris*(2-carboxyethyl) phosphine; coloured circles on compounds **167**, **168**, **170** and **171** indicate functionalization sites as detailed in figure 7).
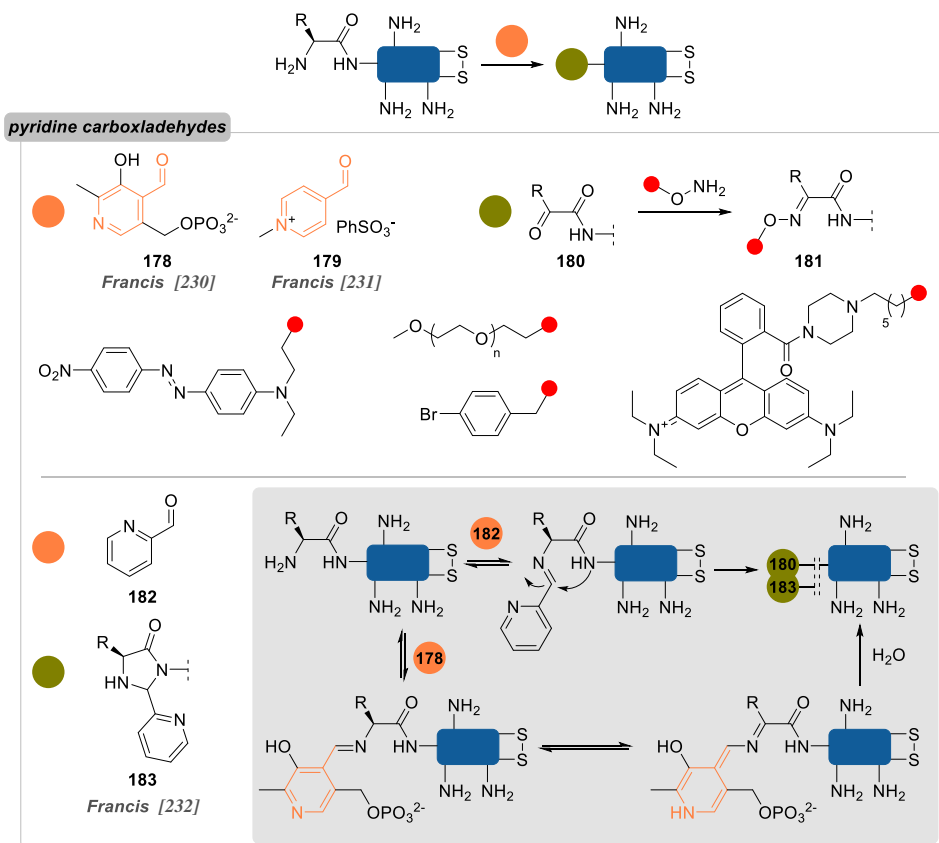
Figure 13. N-terminus selective reagents based on pyridine carboxaldehyde scaffolds (R, side-chain of proteinogenic α-amino acids).

Cysteines rebridging reagents have thus offered access to site-selectively conjugated native proteins in excellent yields. NGM and PD have been the most studied reagents in this regard, and were shown to be the most effective at minimizing disulfide scrambling. More interestingly, varying their structure permitted a controlled access to DAR8, 4 and 2 ADCs.

While lysine and cysteine take most of the light, other residues have also been investigated as valuable targets for accessing site-selective conjugation of native proteins. With a lower pK_a than the ε-amino side-chain of lysine, α-amino groups of N-terminus possess a peculiar reactivity of particular interest in this context, explaining why it has gained increasing attention in recent years.

3.2.3. N-terminal residues

Numerous reports about general strategies for the chemo- and site-selective modification of α-amino groups of native proteins can already be found, including in the recent literature [229]. Most of these methods employ carbonyls, and aromatic aldehydes in particular, but sulfonamides, phthalimides or enzymes were also demonstrated to be valuable candidates for this job. Pushing the constraints even further, focusing on the conjugation of particular N-terminal amino acids—notably N-terminal cysteine, glycine and proline—offers an extra level of selectivity.

In 2006, Francis and co-workers [230] were the first to report the use of pyridine carboxaldehydes for the selective oxidation of the N-terminus (figure 13). Using pyridoxal-5-phosphate (PLP) **178**, N-terminal α-amino acid residues were converted to pyruvamides **180** and further derivatized to oximes **181** equipped with various payloads. N-terminal valine, glycine, lysine, methionine and aspartate residues of several native proteins were successfully modified, with conversions up to 80%, but it was anticipated that other residues would not be suited for this approach: serine, threonine or cysteine due to possible oxazolidine/thiazolidine ring formation; tryptophan, because of competing Pictet–Spengler cyclization; and proline, due to a lack of reactivity. Alternatively, in 2013, they identified N-methylpyridinium-4-carboxaldehyde **179** (Rapoport's salt) as a new and more effective reagent, proceeding via the same mechanism, and applied it to the selective labelling of the heavy-chains

N-terminal glutamates of the antibody trastuzumab (67% conversion), while the light chains *N*-terminal aspartate were left unlabelled [231]. As the scope of *N*-terminal α-amino acids was found to be limited with these reagents, 2-pyridinecarboxyaldehyde (2-PCA) **182** was later reported as a more broadly applicable strategy for the site-selective conjugation of peptides and proteins, irrespective of the nature of the *N*-terminus [232]. Proceeding via another mechanism, this reagent led to the formation of imidazolidinones **183**, with the imine formed in a first step undergoing intramolecular addition of the *n* + 1 α-amino group in a second step—explaining why proteins with proline as the *n* + 1 residue failed to be conjugated via this strategy. While this conjugation proved to be partially reversible at 37 °C, it benefitted from a large protein scope, with good conversions being observed in most cases. Even though *N*-terminal serine-containing proteins were listed as potential limitations to the PLP strategy (*vide supra*), thioredoxin was still successfully labelled with 2-PCA (63% conversion) and the potentially competing oxazolidine formation not discussed by the authors.

Further evidence for an absence of oxazolidine ring formation when *N*-terminal serine is reacted with aromatic aldehydes has been given by Chou and co-workers [233], who conducted the reductive amination of six native proteins and different nonapeptides varying only by their *N*-terminal residues (figure 14). With benzaldehyde **184** in the presence of sodium cyanoborohydride, nearly quantitative conversions to the *N*-terminal alkylated peptides were observed in all cases with the exception of cysteine, where thiazolidine ring formation rivalled that of reductive amination. Interestingly, with 2-PCA, only alkylation was observed and not imidazolidinone formation as previously described by Francis, which could be explained by slightly different reaction conditions between the two studies— different substrates, different pH—but also by kinetic/thermodynamic considerations, imidazolidinone ring formation being reversible.

Taking inspiration from the previous strategies developed by Francis and co-workers, Wong and co-workers [234] recently reported 2-ethynylbenzaldehyde **186** (2-EBA) as a new *N*-terminus conjugation reagent. A cascade starting with imine formation followed by a 6-*endo-dig* cyclization delivered an isoquinolinium derivative as the final product **187** and was used to label peptides and proteins. In the latter case, conversions were found to be modest for all three proteins evaluated—lysozyme, RNAse A and BCarg mutant—and competition between *N*-terminal α-amines and lysine side-chain ε-amine could systematically be observed for conversions over 10%. Side reactions competing with isoquinolinium cyclization after imine formation—e.g. oxazolidine, thiazolidine, imidazolidinone— were not evidenced, suggesting again the profound influence of the aldehyde stereoelectronic effect on the reactivity of its α-imine derivatives.

Another strategy making use of aldehyde reactivity and so-called aldehyde capture ligation (ACL) was described in 2015 by Arora and co-workers [235]. In a mechanism reminiscent of that of native chemical ligation, the ACL employs benzeneselenoesters **188** as acylating reagents, whose reactivity is enhanced by a neighbouring *o*-aldehyde group—up to 100-fold compared to the unsubstituted benzeneselenoester analogue—helping to promote hemiaminal/imine formation and direct the subsequent *Se*-to-*N* intramolecular acyl shift (see mechanism, figure 14). Site-selective conjugation of ubiquitin could be obtained with just five equivalents of **188**, albeit 40 h were necessary to attain good conversion.

This idea that a fast and reversible first reaction, such as imine formation, can accelerate the kinetics of a second irreversible reaction and thus confer a certain degree of site-selectivity to the overall transformation can be found in the work of Rai and co-workers [236], who showed in 2018 that *N*-terminus selective modification of RNase A and aprotinin could be achieved with epoxides and sulfonate esters *only* if they were linked to a benzaldehyde handle in a bifunctional reagent such as **190** and **192**, respectively. By contrast to their subsequent work in 2020 in which they used a similar approach to selectively label a single lysine or histidine residue (see **157** and **158**, figure 11; and **240**–**242**, figure 18), in this case it is the same nitrogen—the α-amino group—that acts as both the directing group and the labelling site, with the transient imine formed in the first step being selectively and irreversibly alkylated in the second. Such a difference in reactivity pattern is remarkable given the great similarity between these two approaches (both in terms of the structure of the reagents and the reaction conditions employed) and further emphasizes the sensitivity of site-selective chemical conjugation techniques to the slightest variations.

While aromatic aldehydes have been shown to be privileged reagents for the *N*-terminus selective conjugation of native proteins, several other electrophiles have also been reported.

Ketenes **194** were found to selectively modify the *N*-terminus of insulin, lysozyme, RNaseA and BCArg, even though low conversions were generally observed, as well as traces of di- and tri-modified conjugates (figure 15) [237].
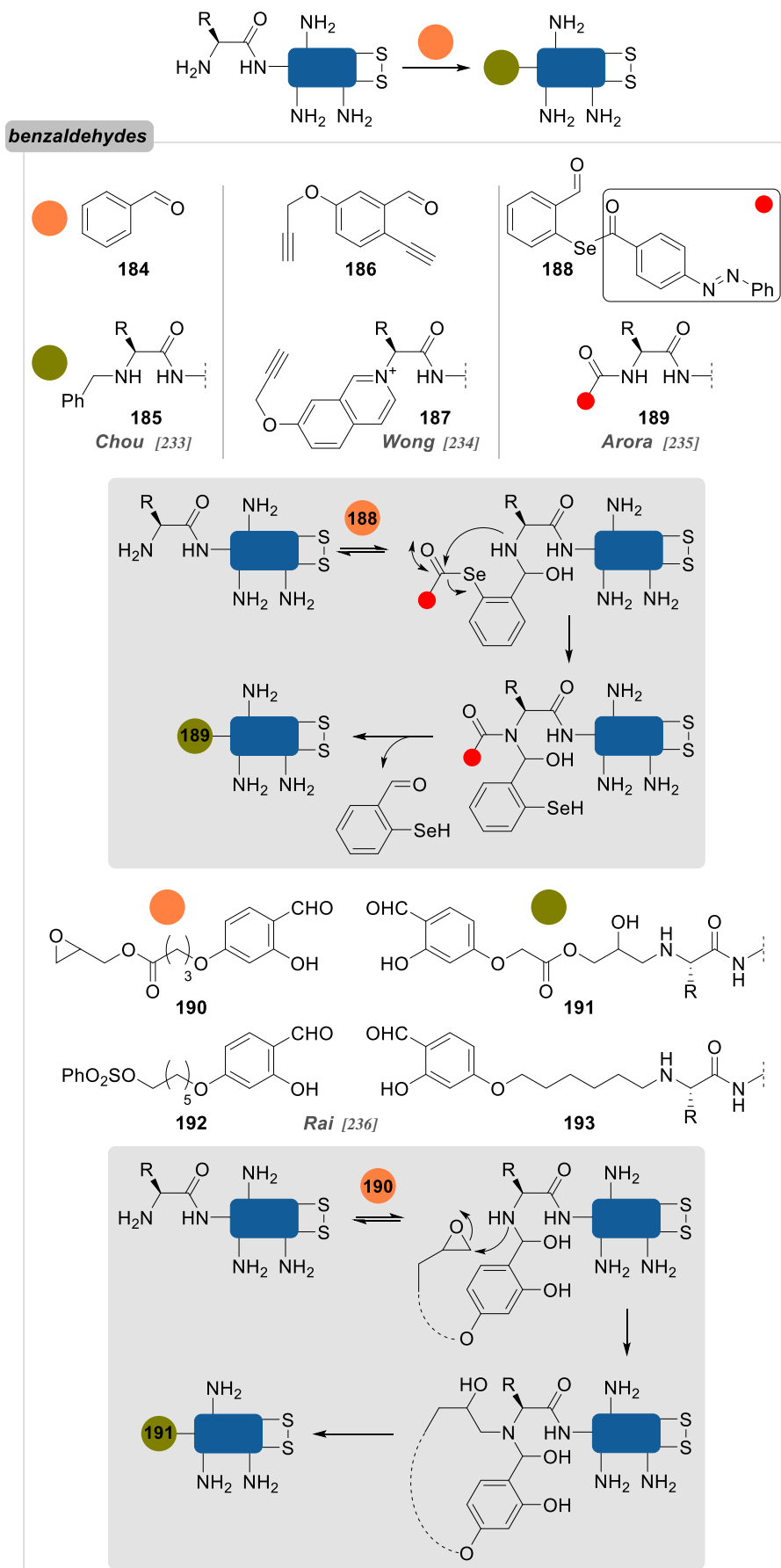
Figure 14. *N*-terminus selective reagents based on benzaldehyde derivatives (R, side-chain of proteinogenic α-amino acids).
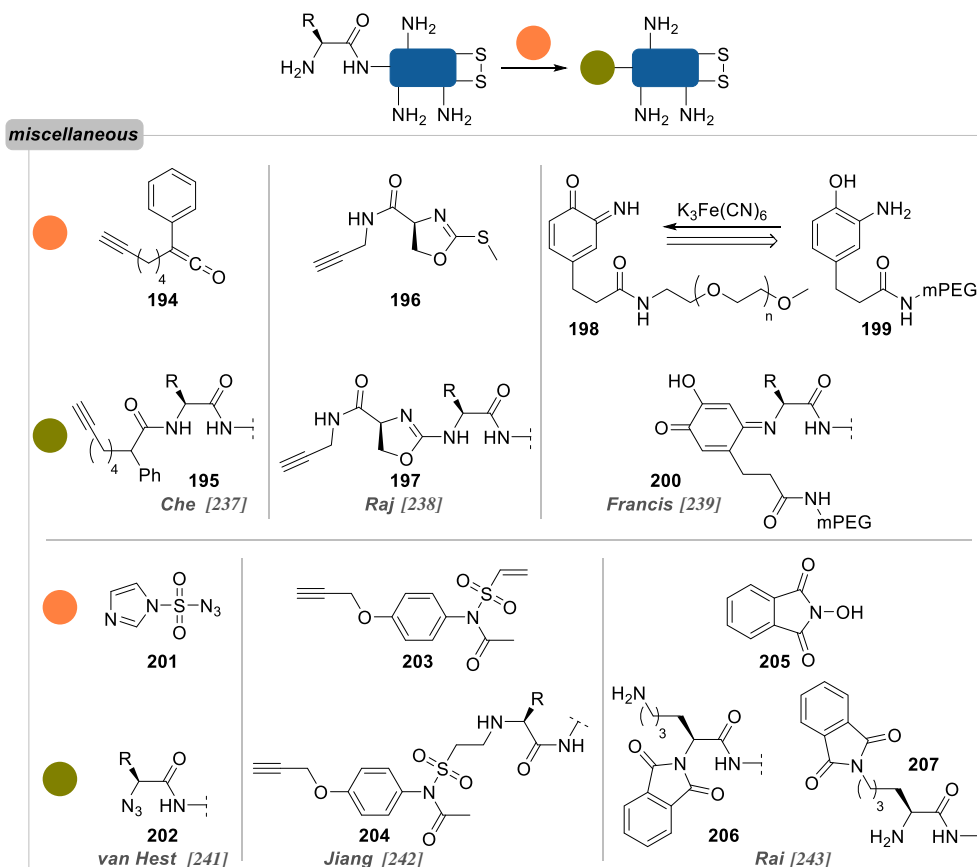
Figure 15. *N*-terminus selective reagents based on miscellaneous derivatives (R, side-chain of proteinogenic α-amino acids).

2-Methylthio oxazolines **196** have also been applied to the site-selective conjugation of native proteins with almost complete conversion in several cases [238]. In addition, the oxazoline group helped to increase ionization of peptide fragments by mass spectrometry and enhance the sensitivity of the method, making this conjugation strategy valuable for proteomics studies even though it was only limited to relatively low molecular-weight proteins (i.e. less than or equal to 17 kDa).

Francis and co-workers [239] proposed an oxidative strategy for the efficient labelling of α-amino groups. Relying on the *in situ* chemical oxidation (using potassium ferricyanide) of *o*-aminophenols **199** to reactive *o*-iminoquinone intermediates **198**, formation of *o*-quinone adducts **200** occurred selectively on *N*-terminus. Moderate conversions were obtained on various native proteins, presumably because of the very nature of their *N*-terminal residue or their solvent accessibility: MS2 (*N*-ter Ala; 0%), lysozyme (*N*-ter Ala; 38%), RNase A (*N*-ter Lys; 30%), myoglobin (*N*-ter Gly; 34%) and chymotrypsinogen (*N*-ter Cys; 42%). Interestingly, *N*-terminal proline residues seemed to be favoured substrates in this transformation, as exemplified by a mutant *N*-ter-proline GFP leading to better conversion (61%) than its wild-type, *N*-ter-Met analogue (24%) under the same conjugation conditions.

In an effort to improve the key oxidation step of phenol derivatives, the same group later reported the use of tyrosinase to convert benchstable 4-substituted phenol reagents to their corresponding *o*-quinone analogues, a reaction that potassium ferricyanide failed to promote [240]. The resulting quinones gave excellent *N*-ter-proline selectivity and high yields, even though small amounts of double modification were consistently observed on several proteins and under various reaction conditions. Importantly, quinone payloads were shown to be reactive toward alkoxyamines and free thiols, presumably because of oxime formation and thio-Michael addition, a potential advantage for the introduction of different payloads post-conjugation that has not been exploited further by the authors.

A handful of reports also detail the use of sulfonyl-based reagents for *N*-terminus selective conjugation. Imidazole-1-sulfonyl azide **201** was used by van Hest and co-workers [241] as a convenient way to convert amines to azides on proteins via diazo transfer. *N*-terminus selectivity was demonstrated on two elastin-like polypeptides and on CalB, even though the presence of minute

amounts of ε-azidation side reaction could consistently be detected. Intriguingly, lysozyme led only to lysine conjugation, presumably because of the poor solvent accessibility of N-terminus.

While vinyl sulfones were shown to be cysteine-selective reagents (see §2.2.2., compound **36**), vinyl sulfonamide **203** was successfully employed for the site-selective conjugation of α-amino groups of somatostatin, lysozyme and RNase [242]. Mildly acidic conditions (i.e. pH 6.0) offered excellent site-selectivity, with only trace amounts of double conjugation observed, even though conversions were rather low, with unconjugated proteins being the most abundant species in the cases of lysozyme and RNase A.

Finally, phthalimidation of primary amines was also found to be effective for the residue-selective modification of RNase A [243]. Using a single equivalent of N-hydroxyphthalimide **205** led to the sole conjugation of K1 (50% conversion), even though clear evidence that α-phthalimidation (**206**) occurred preferentially over ε-phthalimidation (**207**) was not given.

Increasing the level of selectivity further, some strategies recently surfaced focusing on the conjugation of some specific N-terminal residue, taking advantage of a peculiar reactivity caused by structural aspects (e.g. the C_α-unsubstituted glycine, the secondary amine of proline) or by the presence of a complementary reactive group on the side-chain.

Taking inspiration from the synthesis of D-luciferin, Rao and co-workers [244] reported in 2009 a water-compatible condensation reaction for the selective labelling of 1,2- and 1,3-aminothiols in general, and of N-terminal cysteines in particular (figure 16), using 2-cyanobenzothiazoles **208** (CBT). While reversible, this ligation was shown to possess elevated kinetics (greater than $9.0\,M^{-1}\,s^{-1}$) and was applied to the rapid conjugation of several proteins in live cells in just 30 minutes, thanks to the development of various CBT-containing probes equipped with fluorophores or biotin.

Cyclization of the 1,2-aminothiol motif of N-terminal cysteines to 4,5-dihydrothiazole was also employed by Tsai and co-workers in 2020 with 2-((alkylthio)(aryl)methylene)-malononitrile (TAMM) reagents **210** [245]. The reaction proceeded smoothly under similar conditions to those employed by Rao with CBT **208**, and with comparable kinetics, allowing the bioorthogonal conjugation of various peptides and artificial proteins, including in live cells.

Focusing instead on the formation of thioazolidine analogues, Gois and Gao independently described 2-formylphenyl boronic acid **212** (2-FPBA) for the modification of N-terminal cysteines at neutral pH of a library of peptides and of the model protein villin headpiece subdomain [246,247]. While thioazolidine formation typically requires harsh conditions and prolonged reaction times, the presence of a boronic acid helps accelerate its formation—as demonstrated by second order rate constants comprised between 10^2 and $10^3\,M^{-1}\,s^{-1}$ and reactions necessitating only a few minutes to reach completion—and stabilize the resulting heterocycle via the formation of an N–B dative bond. However, this stabilization is only partial, as thiazolidine adducts were shown to dissociate in the presence of free cysteine, hydroxylamine derivatives or under mildly acidic pH.

To address this limitation, Gao and co-workers [248] hypothesized that the acylation of the thiazolidine nitrogen would prevent imine reformation and ring-opening, and thus lead to more stable conjugates. Developing boronic acid reagents **214** based on an acetylated salicaldehyde motif, they managed to observe formation of the expected adducts **216**, with the O-to-N-acetyl shift from boron adduct **215** being the rate limiting step. Complete conversions were typically observed within 2 h at pH 6.0 with various peptides and proteins, and the strategy was also applied to the chemical modification of a phage library. As hypothesized, the intramolecular acetyl transfer led to a substantial gain in stability for the thiazolidine adducts, which were shown to be insensitive to the presence of free cysteine, and stable for days at pH 5.0.

Switching to another specific family of N-terminal residues, Rai and co-workers [249] showed in 2019 that N-terminal glycines could form 1,2-amino alcohols **218** when reacted with ortho-substituted benzaldehydes **217** (figure 16). Notwithstanding the excessive amount of reagent employed (500 equivalents), the reaction proceeded under smooth conditions—pH 7.8, 24 h, 25 °C—and proved to be selective for N-terminal glycines as proteins exhibiting C_α-substituted N-terminal residues were not labelled in competition experiments.

More recently, selective modification of N-terminal prolines was also proposed. Francis and co-workers [239,240] used quinone derivatives as conjugation reagents, generated in situ via chemical or enzymatic methods (vide supra), while Raj and co-workers [250] employed the Petasis reaction for the preferential modification of peptides with an N-terminal proline (figure 16). Applied to proteins, a mixture of 2-PCA **182** and p-methoxyphenyl boronic acid **219** did not lead to any detectable conjugates when myoglobin (N-ter Gly) and α-lactalbumin (N-ter Ala) were evaluated, but successfully modified creatine kinase and aldolase (both N-ter Pro), even though conversions were not
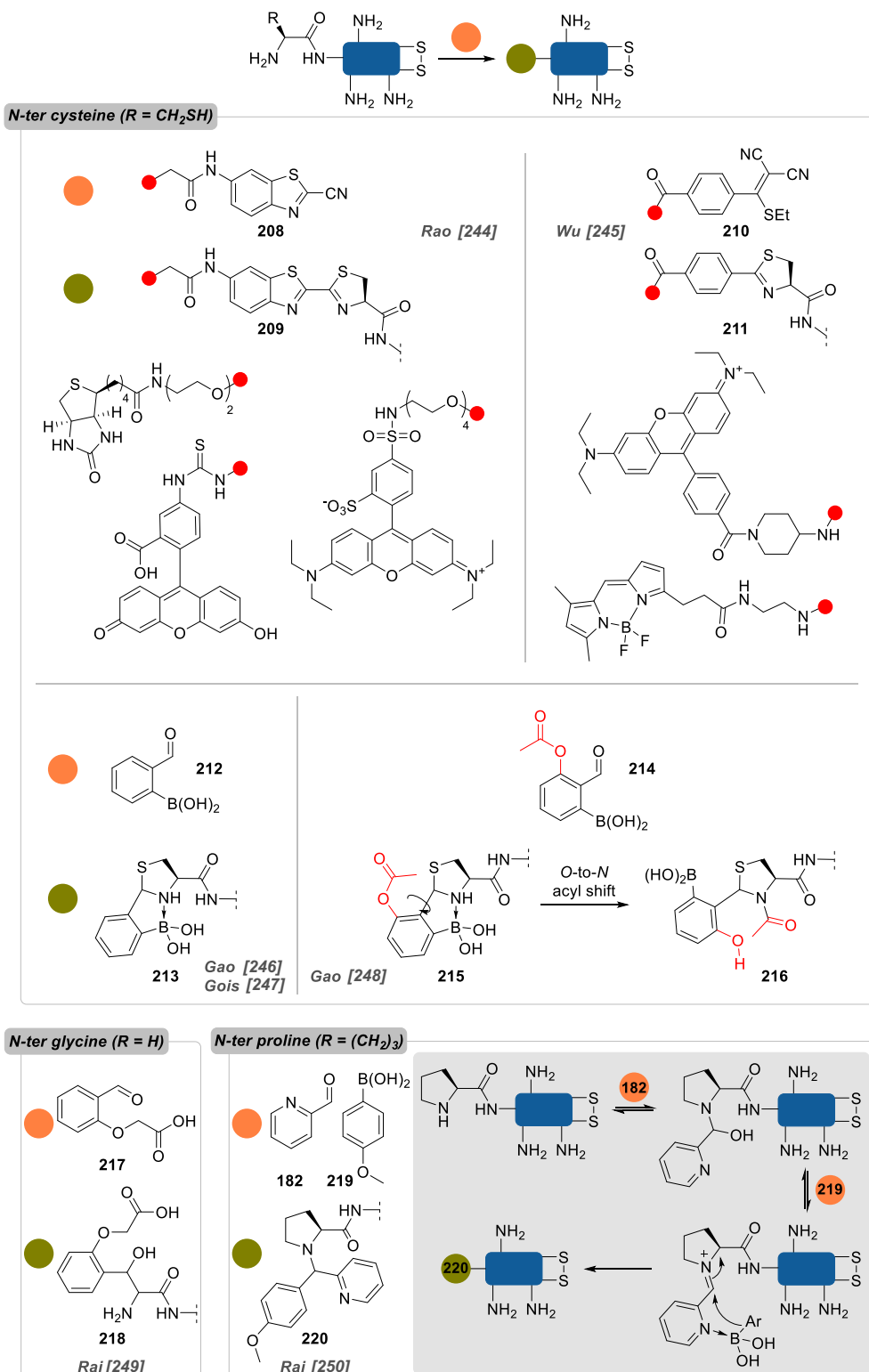
Figure 16. N-terminal cysteine, glycine and proline selective strategies based on the formation of 4,5-dihydrothiazole/thioazolidine, 1,2-aminoalcohol and N-alkylated conjugates, respectively.

given and no further analyses were conducted to confirm the site-selectivity of the conjugation. Interestingly, formation of imidazolidinone as reported by Francis *et al.* (figure 13) with 3-hydroxy-2-pyridinecarbaldehyde could be suppressed in this case to favour Petasis reaction by finely tuning the reaction conditions and amount of boronic acid nucleophile.

Strategies focusing on N-terminus selective modification have thus played on various parameters to produce successful results: the usually observed solvent accessibility of α-amines coupled with their

lower pK_a value compared with ε-amines, but also the presence on the α-carbon of both the protein's amide backbone and relatively acidic protons, both of which could confer a particular reactivity to transient imines formed on these positions. Similarly, the nature of the N-terminal residue itself can confer an additional degree of selectivity and a handful of strategies targeting only specific residues have thus been successfully devised, essentially centred on N-terminal cysteine, glycine and proline.

In the following paragraph, we will talk about the site-selective conjugation of residues which have not benefited—yet—from extensive research compared with their predecessors; these concern tyrosine, tryptophan, histidine, methionine, arginine, serine, aspartate, glutamate and C-terminal residues.

3.2.4. Miscellaneous

Tyrosines have become an interesting target for facile site-selective modification of proteins due to their moderate abundancy at the surface of proteins and hence limited solvent accessibility. Extensive study of tyrosine modification by the group led by Francis in the early 2000s led to the development of valuable chemoselective strategies and reagents—e.g. diazonium reagents, three-component Mannich-type reactions, palladium-catalysed O-alkylation—which were shown to lead to site-selective modifications of different proteins, from bacteriophage MS2 to chymotrypsinogen A [73,251,252].

More recently, Hulme and co-workers [253] developed an original catch-and-release strategy for the site-selective tagging of proteins (figure 17). Using agarose beads derivatized with diazonium groups **221**, they managed to trap proteins via electrophilic aromatic substitution of solvent-accessible tyrosine residues. Interestingly, they demonstrated that site-selectivity was possible, at least for RNase A, with only one (Y115) out of the three most exposed tyrosines being conjugated, which could be possibly explained by minimized electrostatic repulsion between the resin and the region around this residue. While the yield of protein capture was rather low—33% at best—addition of sodium dithionate led to the cleavage of the azobenzene group of **222** and the release of the modified protein, bearing an o-aminophenol group **223** that could be selectively oxidized and derivatized following Francis protocol (*vide supra*).

Besides 'two-electron chemistry' approaches relying on nucleophile/electrophile duets, another prominent strategy for tyrosine modification has taken advantage of the SOMOphilic character of the phenol ring—its propensity to react selectively with radical species. While single-electron oxidation (SEO) of tyrosines' phenol ring to C-centred radical species has been used to produce cross-linked proteins and protein oligomers (see box, figure 17) [254], a similar approach has been employed by different groups to generate reactive N-centred radical reagents—either enzymatically, using a combination of horseradish peroxidase (HRP) and hydrogen peroxide, or via electrochemical and photochemical techniques—[255] in order to label selectively only surface-exposed, hydrogen-bonded residues. Nakamura and co-workers [256] employed N-methylated luminol radical derivatives **224** and observed excellent chemo- and site-selectivity on different proteins, including antibodies such as trastuzumab in which Y57 was identified as the unique conjugation site [255]. However, site-selectivity was lost when multiple surface-exposed residues were present, as in the case of rituximab where four tyrosines were found to be responsible for the antibody's conjugation (Y32, Y52, Y101 and Y102). Very recently, the MacMillan group detailed the photoredox-catalysed coupling between tyrosines and phenoxazines **227**. When exposed to blue light (34 W, 440 nm) in the presence of lumiflavin as the photocatalyst, the phenoxazine ring was converted to the electrophilic N-centred radical **228** that coupled selectively with the electron-rich phenol groups of tyrosines. In addition, the presence of two formyl groups on the phenoxazine ring allowed for further functionalization of the conjugates by classical hydrazine ligation [257]. Most notably, this strategy led to exquisite site-selectivity on nine different native proteins with molecular weights ranging from 5.8 to 77.0 kDa. The observed selectivity was rationalized by the preferential reactivity of phenol side-chains engaged in hydrogen bonding with their local microenvironment and exposed to the surface of proteins. Conversions were higher than 70% on average and site-selectivity levels essentially over 90%, with a single labelled tyrosine identified in five cases—in the most prominent example, Y136 was the sole conjugation site identified on serotransferrin (77.0 kDa), despite the presence of 25 other tyrosines. Curiously, while following a very similar approach based on the same rationale (i.e. surface-exposed tyrosines react best), Nakamura and MacMillan identified two different conjugation sites on horse myoglobin (which only possesses two tyrosine residues): Y146 for MacMillan, Y103 for Nakamura. The absence of reactivity of Y146 in the latter case was suggested to be due to its shielding from **225** by the heme-coordinated H93 residue—we could imagine that the same heme unit could direct the
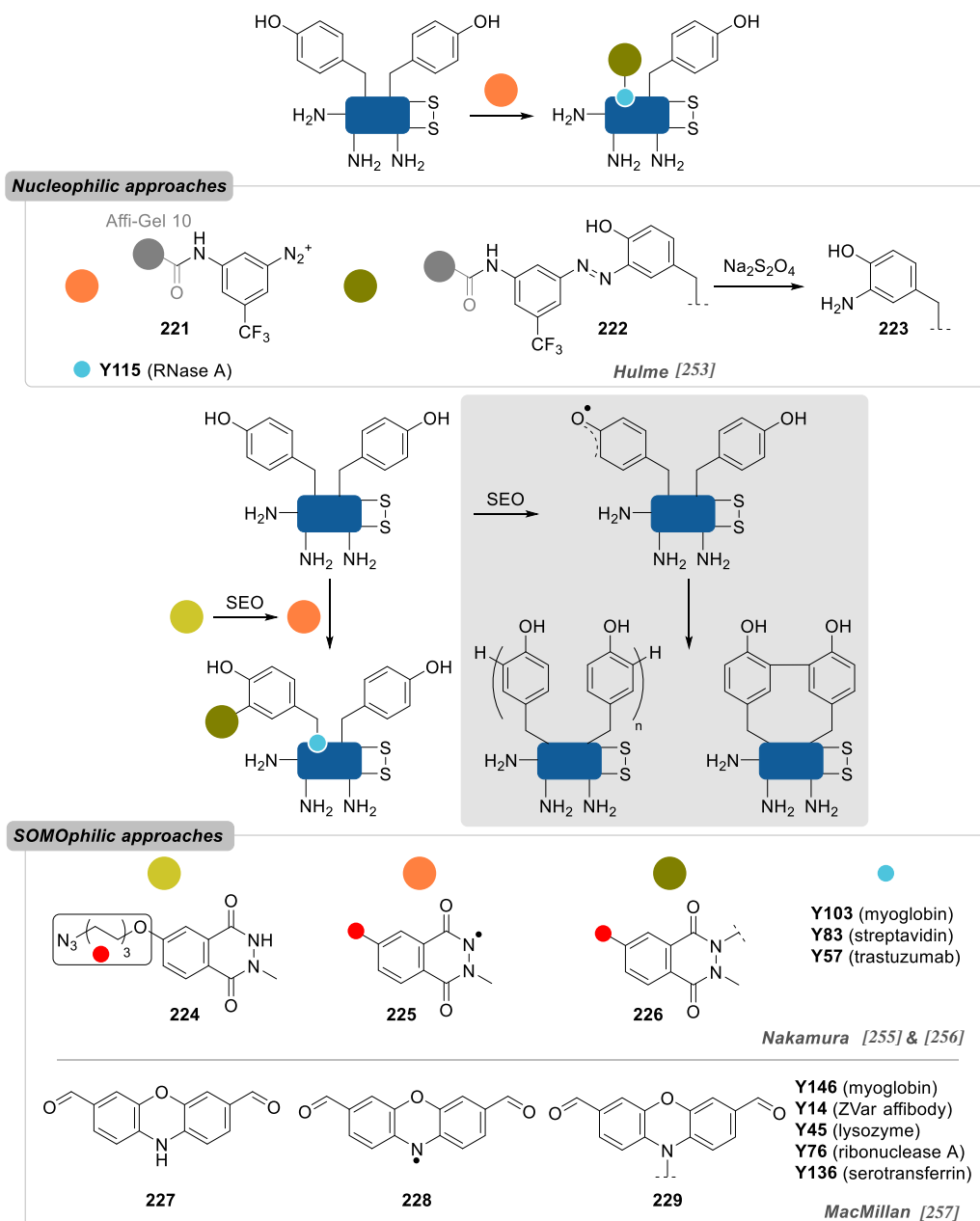
Figure 17. Overview of site-selective chemical conjugation of native proteins on tyrosine (SEO, single-electron oxidation).

preferential approach of phenoxazine radical **228** towards Y146 via electrostatic interactions and explain MacMillan's study results.

Turning to another amino acid residue equipped with an electron-rich side-chain, tryptophan has also been envisioned as an attractive target for site-selective modification. Being the least abundant proteinogenic amino acid in proteins—to the exception of the exotic pyrrolysine and selenocysteine—chemoselectivity and site-selectivity could, therefore, coincide in the case where only one tryptophan residue is present in the POI. Over the years, different research groups have tried to modify this amino acid selectively, albeit with limited success on proteins due to the harsh conditions being often required. Two main families of reagents have been investigated for this purpose: metallocarbenes and radical species.

In 2004, Antos & Francis [258] proposed using metallocarbenes—generated from rhodium(II) acetate and diazo precursor **230** in the presence of hydroxylamine as ligand for the intermediate complex—for the chemoselective modification of the indole rings of tryptophans—conjugation products **231** obtained as mixture of N- and C2-alkylated products—on two proteins: myoglobin (two W residues) and subtilisin Carlsberg (one W residue) (figure 18). Due to the residues' poor solvent accessibility in these proteins,

conjugation had to be performed under exceedingly harsh conditions—pH 3.5 for myoglobin, 1.5 for subtilisin Carlsberg—in order to denature the proteins and expose the indole rings, leading quite predictably to an absence of site-selectivity in the case of myoglobin. Replacing hydroxylamine with N-(tert-butyl)hydroxylamine did not alter the chemoselectivity of the method but allowed us to conjugate peptides such as melittin under milder conditions (i.e. pH 6.0) [259]. However, thermally induced denaturation was still needed to conjugate effectively native proteins with unexposed tryptophan residues.

Single modification of tryptophan on various native proteins was achieved by Kanai and co-workers [260] under smoother conditions by mixing five equivalents of the 9-azabicyclo[3.3.1]-nonane N-oxyl (ABNO) derivative **232** with three equivalents of sodium nitrite, presumably generating reactive oxoammonium species **233**. This resulted in the sole formation of mono-labelled conjugates with conversions oscillating between 11% and 64% on five different proteins—lysozyme (64%), myoglobin (53%), concanavalin A (11%), BSA (25%) and β_2-microglobulin (15%). Site-selectivity was demonstrated by X-ray diffraction analysis of a lysozyme conjugate, highlighting W62 as the only conjugated tryptophan residue over the six present in this protein, presumably because of a less hindered environment.

Radical trifluoromethylation of tryptophans using sodium trifluoromethanesulfinate (NaTFMS or Langlois reagent) and tert-butyl hydroperoxide (tBuOOH) was also described by Davis and co-workers [261]. Good chemoselectivity was observed, with tryptophan being identified as primary modification sites—other aromatics-containing amino acids were also found to be partially labelled in some cases. Logically, this translated to a good level of site-selectivity on pantothenate synthetase—possessing a single tryptophan residue—which was lost with more complex targets—i.e. horse-heart myoglobin (two tryptophans, W7 and W14; both labelled) and lysozyme (six tryptophans, W28, W62, W63, W108, W111 and W123; all labelled except W111).

Taylor and co-workers [262] proposed photochemical reactions as a valuable strategy for the modification of these residues due to tryptophan's propensity to participate in photoinduced electron transfer events. Employing diversely functionalized N-substituted pyridinium salts **236** and UV-B irradiation in the presence of glutathione, chemoselective C2-carbamylation of tryptophans' indole ring was observed via a mechanism that remains elusive. Fine tuning of the reaction conditions allowed excellent conversions to be obtained on few peptides and led also—like Kanai's ABNO strategy—to the selective coupling of W62 on lysozyme (up to 94% conversion) even though double modification could also be detected, accounting for up to 25% of conjugated species. This second labelling event could, however, be located on the same W residue, as evidenced during the optimization study conducted on the peptide octreotide.

After tyrosine and tryptophan, the last member of the electron-rich-aromatics family of α-amino acids is histidine and its imidazole ring. Chemoselective modification of this residue is challenging due to the presence of more nucleophilic groups on proteins that can outcompete the imidazole ring for substitutions and to the absence of a particular reactive feature of this aromatic ring that could be used to label it selectively—such as phenols' facile generation of C-centred radicals, or indoles' high molar absorptivity and propensity to undergo photoionization, for example. Consequently, general, broadly applicable strategies for histidine modifications have regularly showed chemoselectivity issues [263], circumscribing site-selective reactions to only a few examples.

In an attempt to expand the portfolio of histidine-selective reagents, Joshi & Rai [264] screened several electrophiles on the model protein ubiquitin, which contains only a single histidine residue, and identified 2-cyclohexenone **238** as the best candidate (46% conversion; figure 18). The presence of a ketone in the final conjugate **239** allowed further functionalization of the protein with alkoxyamine probes via oxime formation, which also helped suppress the reversibility of the conjugation reaction under physiological conditions (PB, 0.1 M, pH 7.0, 25 °C). This strategy was applied to the site-selective conjugation of several native proteins, from insulin (H10, 52%) to myoglobin (H113, 25%) and His-tagged proteins in cell lysates.

Based on their previously reported linchpin-directed modification (figures 11 and 14), Rai and co-workers [265] developed a variation adapted to the modification of histidines (figure 18). In this latter case, reagents with two reactive handles separated by linkers of different sizes and structures were designed (**240–242**). On one extremity, an aromatic aldehyde will form Schiff bases with lysines (fast, reversible); this will help to direct the approach of a monosubstituted epoxide—located on the other extremity—towards a proximal histidine to form N-alkylated adducts (slow, irreversible). Site-selective, monolabelling of various proteins was demonstrated, albeit on different histidines and in lower conversion compared with their cyclohexenone approach (vide supra) for the same targets. This
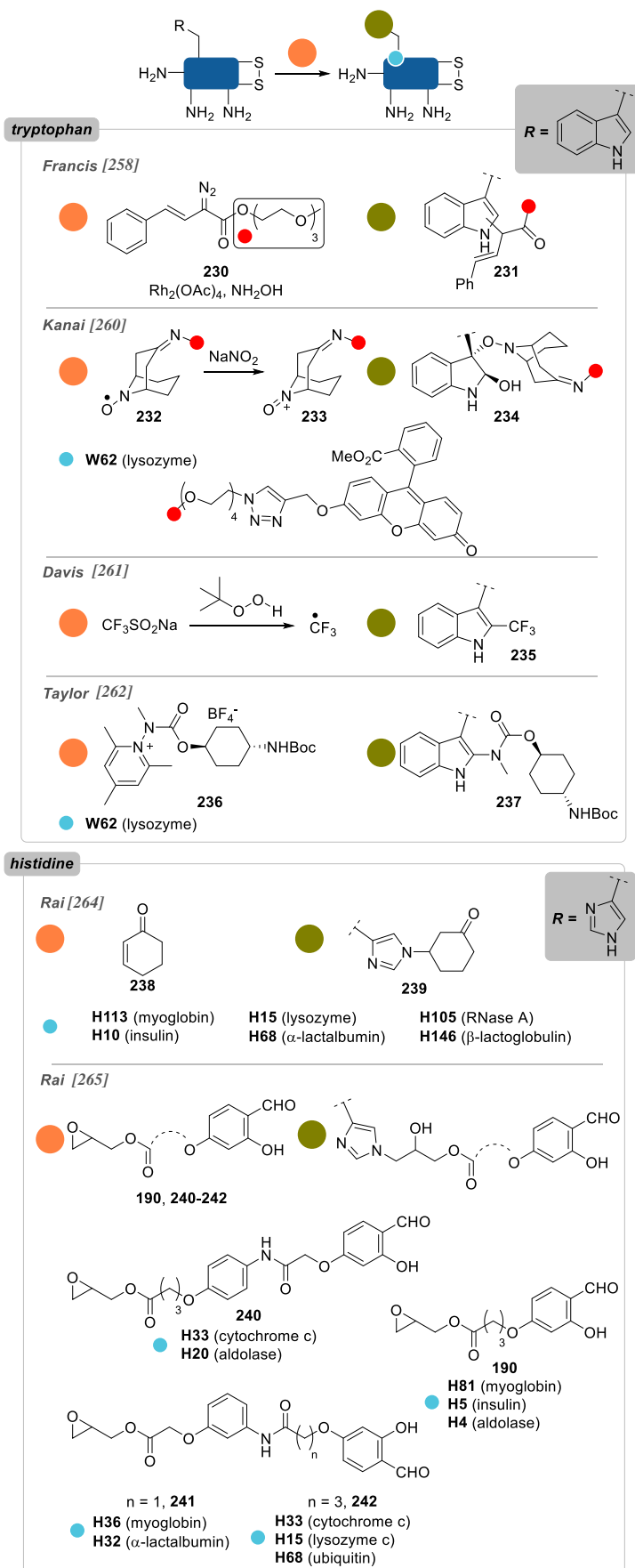
Figure 18. Overview of site-selective chemical conjugation of native proteins on tryptophan and histidine.
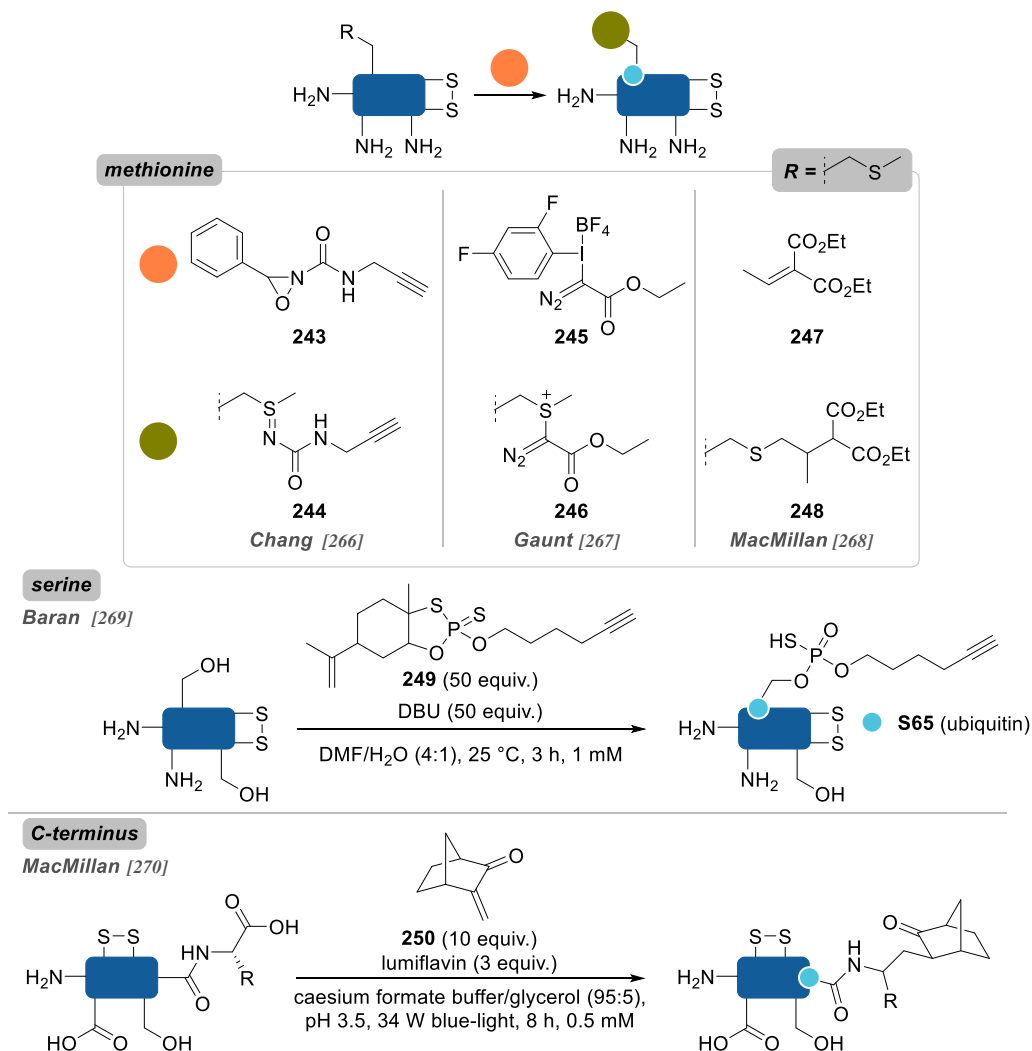
Figure 19. Site-selective strategies for the chemical conjugation of methionine, serine and N-terminal residues of native proteins (R, side-chain of proteinogenic α-amino acids).

degree of selectivity seems to depend on both the linker structure (no N-terminal labelling was identified with **190**, as opposed to what had been observed on RNase A and aprotinin; figure 14) and the protein size: while the Fab fragment of trastuzumab could be site-selectively conjugated (H189; light chain) with **190**, this site-selectivity eroded when the full antibody was used instead. Interestingly, single labelling of a unique protein (cytochrome C) could be achieved when reacted in a mixture with six other proteins.

Similar to tryptophan, methionine residues are also both scarce—the second least abundant amino acids in proteins—and poorly solvent accessible, tending to be buried deep into proteins' tertiary structures. Killing two birds with one stone, developing a methionine-selective conjugation strategy would thus likely lead to a site-selective approach. While limited, a few methods can be found in the early literature, essentially relying on S-alkylation at low pH, methionines' thioether groups being the sole side-chain residue not protonated under harsh acidic conditions. While such stringent conditions helped compensate for the low intrinsic nucleophilicity of the thioether group, the potential instability of the resulting sulfonium coupled with the absence of biocompatible conditions explained why this approach had been disregarded. Offering a new look at methionine conjugation, two seminal papers from the late 2010s reported smooth and chemoselective strategies for the labelling of this residue.

In 2017, Toste, Chang and co-workers [266] detailed the use of oxaziridine reagents such as **243** for oxidizing thioether to sulfimides **244** within minutes at room temperature (PBS, pH 7.4; figure 19). Only surface accessible methionines were found to react, which resulted in the nearly quantitative conversion of all nine native methionine residues on native calmodulin to their sulfimide derivatives. The strategy proved to be highly versatile and was used to produce ADC—employing artificial trastuzumab though, due to the lack of solvent-accessible methionine in the wild-type antibody—and to identify

proteins with reactive methionines in cell lysates in a chemoproteomic approach. One year later, in 2018, the group led by Gaunt reported an alternative strategy, this time employing hypervalent iodine(III) reagents such as λ^3-iodane **245** [267]. The reaction proceeded in the presence of TEMPO—to minimize the formation of oxidative by-products—and thiourea—whose role was not fully understood—under acidic conditions, and led to the formation of sulfoniums **246** equipped with a diazo motif that could be further functionalized by photocatalysis. Here as well, the method proved to favour solvent-accessible methionines, and while excellent levels of monofunctionalization could be observed on various proteins, identification of the modified methionines was not discussed.

More recently, the MacMillan group offered a third original and effective strategy for methionine conjugation, based on a photoredox catalytic protocol similar to that of tyrosine conjugation with phenoxazines derivatives discussed previously (*vide supra*) [268]. In the presence of flavin as photocatalyst, methionines' thioether functions could be converted to α-thio radicals under blue-light irradiation, and trapped with various Michael acceptors, such as diethyl ethylidenemalonate **247**. Under smooth conditions, a handful of proteins were shown to be chemoselectively conjugated on methionine residues, with acceptable site-selectivity levels in cases where more than one methionine residue is present—up to approximately 70% for myoglobin (M55 versus M131, 5 : 2) and human growth hormone (M14 versus M125 and M170, 11 : 5).

Finally, it is worth mentioning some isolated examples of novel strategies aiming at the conjugation of previously unreported residues.

Baran and co-workers [269] paved the way toward serine-selective conjugation with phosphorus(V)-based electrophile **249**. In the presence of DBU, modification of serine could be detected, both as free amino acid and residue in peptides, in high yield and great selectivity. Despite lower conversion (less than or equal to 40%), transposition of this method to ubiquitin delivered a mono-labelled conjugate with S65 as the sole modification site. In the authors' own words, these reagents still need to be optimized to make them more reactive, more stable in aqueous conditions (reaction is done in 80% of DMF) and compatible with cysteine residues.

Finally, based on the difference in oxidation potentials between the side-chain carboxylates of aspartate and glutamate and the C-terminal α-amino carboxylate, Macmillan and co-workers [270] envisioned that the latter could be selectively modified via single-electron transfer. Employing lumiflavin as a water-compatible photocatalyst, the C-terminus of three different native proteins was efficiently modified with a variety of α,β-unsaturated carbonyls, including 3-methylene-2-norbornanone **250**, via photoredox chemistry under highly acidic conditions (pH 3.5). While conversions were rather limited, this strategy proved to be able to differentiate between the A- and B-chain C-termini of insulin, the former being selectively alkylated. This could be due to a variation in oxidation potentials between the two groups or to a directing effect from a lipophilic region nearby the A-chain C-terminus, adsorbing the photocatalyst.

4. Conclusion

The chemical conjugation of proteins has thus gone a long way since its first applications in the tanning industry more than 150 years ago thanks to the progressive understanding of proteins' structure and composition which fostered the development of chemical reagents and techniques with improved selectivity. Essentially centred on the most nucleophilic and solvent-accessible lysine and cysteine residues in the late 1990s, chemoselective strategies have benefitted from the marked interest of the community of synthetic chemists, which led to the development of various techniques for the precise labelling of functional groups that were previously inaccessible—e.g. methionines' thioethers, tryptophans' indoles, serines' primary alcohols, N-ter and C-ter residues. In parallel, synthetic chemistry *and* synthetic biology both powered heavily the development of site-selective protein conjugation with the expansion of UAA mutagenesis and the rise of bioorthogonal reactions, resulting in an exponential growth of reports and applications in the past decade. While not as effective yet, chemical strategies for the site-selective conjugation of native proteins have also been developed, some of which even showed promising levels of protein selectivity. The challenge in the case of native proteins lies in the understanding of the slight variations in reactivity of the various surface-exposed reactive entities, governed by the three-dimensional conformation of the protein and the microenvironment surrounding the residues and affecting their nucleophilicity, basicity and access. Working with these constraints, innovation in the site-selective conjugation of native proteins has been driven by two main approaches: the development of new reagents (e.g. bifunctional probes,

phosphorus(V)-based electrophiles, hypervalent iodine(III) compounds) and that of novel techniques (e.g. photocatalysis, linchpin-directed modifications), with a clear overlap between these two areas in several cases. While predicting what the future holds in store for the field is nearly impossible, as it might depend heavily on the applications of all reported strategies to date, one can assume that the involvement of new techniques and new fields of research can lead to improved conjugation strategies. The 2010s have seen a renewed interested in electrochemistry from the community of synthetic chemists, which could one day permeate the field of bioconjugation. As another underexplored technique, mechanochemistry could offer interesting perspectives. Having transitioned from an exotic oddity to an industrially applicable method, one can imagine it could lead to new reactivity profiles by modifying local microenvironments at the surface of the protein. As shown by a recent work from the Bernardes group [199], computational chemistry can help design tailored reagents for site-selective applications. Coupled with X-ray crystallography to decipher proteins' three-dimensional structure, it could facilitate our understanding of the key factors influencing reactivity and lead to new synthetic strategies taking advantage of the spatial proximity between two or more reactive residues to narrow down the potential number of conjugation sites. While several strategies have already been reported, one can finally expect an increased participation from transition-metal catalysis in protein bioconjugation, even though the nitrogen- and oxygen-rich structure of these biomolecules—and the presence of several conjugated π-systems—can complicate this approach. Put together, this could foster spectacular breakthroughs, leading one day to the conjugation of presumably unreactive substrates—the family of alkyl-substituted amino acids—in the same manner as Csp^3-H activation has reshaped the field of C–C bond formation. Paradoxically, the uncertainty about the future direction of the field is also the source of one of the only certainties to date: the field of protein bioconjugation still requires continuous development of new and novel chemical techniques and reagents—even for already well-studied residues such as lysines and cysteines—for the very reason that we never know what beneficial applications they might one day have.

References

1. Bird RE, Lemmel SA, Yu X, Zhou QA. 2021 Bioorthogonal chemistry and its applications. *Bioconjugate Chem.* 32, 2457–2479 (doi:10.1021/acs.bioconjchem.1c00461)

2. Brittain WDG, Coxon CR. 2021 Perfluoroaryl and perfluoroheteroaryl reagents as emerging new tools for peptide synthesis, modification and bioconjugation. *Chem. Eur. J.* 27, 1–15. (doi:10.1002/chem.202103305)

3. Haque M, Forte N, Baker JR. 2021 Site-selective lysine conjugation methods and applications towards antibody–drug conjugates. *Chem. Commun.* 57, 10 689–10 702. (doi:10.1039/D1CC03976H)

4. Allouche EMD, Grinhagena E, Waser J. 2021 Hypervalent iodine-mediated late-stage peptide and protein functionalization. *Angew. Chem. Int. Ed.* 60, 2–17. (doi:10.1002/anie.202112287)

5. Kang MS, Kong TWS, Khoo JYX, Loh T-P. 2021 Recent developments in chemical conjugation strategies targeting native amino acids in proteins and their applications in antibody–drug conjugates. *Chem. Sci.* 12, 13 613–13 647. (doi:10.1039/D1SC02973H)

6. Means GE, Feeney RE. 1990 Chemical modifications of proteins: history and applications. *Bioconjugate Chem.* 1, 2–12. (doi:10.1021/bc00001a001)

7. Yadav DK, Yadav N, Khurana SMP. 2014. Chapter 26 - vaccines: present status and applications. In *Animal biotechnology* (eds AS Verma, A Singh), pp. 491–508. San Diego, CA: Academic Press.

8. van Slyke DD. 1910 Eine Methode zur quantitativen Bestimmung der aliphatischen Aminogruppen; einige Anwendungen derselben in der Chemie der Proteine, des Harns und der Enzyme. *Ber. Dtsch. Chem. Ges.* 43, 3170–3181.

9. Hintermann L. 2010 *Comprehensive organic name reactions and reagents*. Hoboken, NJ: John Wiley & Sons, Inc.

10. Fields R. 1971 The measurement of amino groups in proteins and peptides. *Biochem. J.* 124, 581–590. (doi:10.1042/bj1240581)

11. Ellman GL. 1959 Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82, 70–77. (doi:10.1016/0003-9861(59)90090-6)

12. Hoare DG, Koshland DE. 1967 A method for the quantitative modification and estimation of carboxylic acid groups in proteins. *J. Biol. Chem.* 242, 2447–2453. (doi:10.1016/S0021-9258(18)95981-8)

13. Barman TE, Koshland DE. 1967 A colorimetric procedure for the quantitative determination of tryptophan residues in proteins. *J. Biol. Chem.* 242, 5771–5776. (doi:10.1016/S0021-9258(18)99367-1)

14. Olcott HS, Fraenkel-Conrat H. 1947 Specific group reagents for proteins. *Chem. Rev.* 41, 151–197. (doi:10.1021/cr60128a004)

15. Herriott RM. 1947 Reactions of native proteins with chemical reagents. In *Advances in protein chemistry*, vol. 3 (eds ML Anson, JT Edsall), pp. 169–225. Cambridge, MA: Academic Press.

16. Balls AK, Jansen EF. 1952 Stoichiometric inhibition of chymotrypsin. In *Advances in enzymology and related areas of molecular biology* (ed. D Purich), pp. 321–343. Hoboken, NJ: John Wiley & Sons, Ltd.

17. Hermanson G. 2008 *Bioconjugate Techniques - 2nd Edition*. Cambridge MA: Academic Press.

18. Kratzer U, Sommersdorf C, Maier S, Wagner TR, Templin M, Joos TO, Rothbauer U, Zeck A, Poetz O. 2021 Tris(hydroxymethyl)aminomethane compatibility with *N*-hydroxysuccinimide ester chemistry: biotinylation of peptides and proteins in TRIS buffer. *Bioconjugate Chem.* **32**, 1960–1965. (doi:10.1021/acs.bioconjchem.1c00283)

19. Jentoft N, Dearborn DG. 1979 Labeling of proteins by reductive methylation using sodium cyanoborohydride. *J. Biol. Chem.* **254**, 4359–4365. (doi:10.1016/S0021-9258(17)30016-9)

20. Baslé E, Joubert N, Pucheault M. 2010 Protein chemical modification on endogenous amino acids. *Chem. Biol.* **17**, 213–227. (doi:10.1016/j. chembiol.2010.02.008)

21. Algar WR. 2017 A brief introduction to traditional bioconjugate chemistry In *Chemoselective and bioorthogonal ligation reactions* (eds WR Algar, PE Dawson, IL Medintz), pp. 1–36. Hoboken, NJ: John Wiley & Sons, Ltd.

22. Boutureira O, Bernardes GJL. 2015 Advances in chemical protein modification. *Chem. Rev* **115**, 2174–2195. (doi:10.1021/cr500399p)

23. Koniev O, Wagner A. 2015 Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **44**, 5495–5551. (doi:10.1039/C5CS00048C)

24. Gildersleeve JC, Oyelaran O, Simpson JT, Allred B. 2008 Improved procedure for direct coupling of carbohydrates to proteins via reductive amination. *Bioconjugate Chem.* **19**, 1485–1490. (doi:10.1021/bc800153t)

25. Smith GP. 2006 Kinetics of amine modification of proteins. *Bioconjugate Chem.* **17**, 501–506. (doi:10.1021/bc0503061)

26. Dovgan I, Ursuegui S, Erb S, Michel C, Kolodych S, Cianférani S, Wagner A. 2017 Acyl fluorides: fast, efficient, and versatile lysine-based protein conjugation via plug-and-play strategy. *Bioconjugate Chem.* **28**, 1452–1457. (doi:10. 1021/acs.bioconjchem.7b00141)

27. Berndt DC, Faburada AL. 1982 Reaction of acyl azide and amines. Kinetics and mechanism. *J. Org. Chem.* **47**, 4167–4169. (doi:10.1021/ jo00142a033)

28. Ślósarczyk AT, Ramapanicker R, Norberg T, Baltzer L. 2012 Mixed pentafluorophenyl and o-fluorophenyl esters of aliphatic dicarboxylic acids: efficient tools for peptide and protein conjugation. *RSC Adv.* **2**, 908–914. (doi:10. 1039/C1RA00530H)

29. Anderson GW, Zimmerman JE, Callahan FM. 1964 The use of esters of N-hydroxysuccinimide in peptide synthesis. *J. Am. Chem. Soc.* **86**, 1839–1842. (doi:10.1021/ja01063a037)

30. Pham GH *et al.* 2018 Tuning a protein-labeling reaction to achieve highly site selective lysine conjugation. *Chembiochem* **19**, 799–804. (doi:10.1002/cbic.201700611)

31. Radzicka A, Wolfenden R. 1996 Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *J. Am. Chem. Soc.* **118**, 6105–6109. (doi:10.1021/ja954077c)

32. Lee YC, Stowell CP, Krantz MJ. 1976 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry* **15**, 3956–3963. (doi:10.1021/ bi00663a008)

33. Bozler H, Jany KD, Pfleiderer G. 1983 Synthesis and application of a fluorescent imido ester for specific labelling of amino groups in proteins. *Biochim. Biophys. Acta* **749**, 238–243. (doi:10. 1016/0167-4838(83)90230-3)

34. Hermanson GT. 2013 In *Bioconjugate techniques (third edition)* (ed. GT Hermanson), pp. 229–258. Boston, MA: Academic Press.

35. Riggs JL, Seiwald RJ, Burckhalter JH, Downs CM, Metcalf TG. 1958 Isothiocyanate compounds as fluorescent labeling agents for immune serum. *Am. J. Pathol.* **34**, 1081–1097.

36. Jorbágy A, Király K. 1966 Chemical characterization of fluorescein isothiocyanate-protein conjugates. *Biochim. Biophys. Acta* **124**, 166–175. (doi:10.1016/0304-4165(66)90325-4)

37. Tuls J, Geren L, Millett F. 1989 Fluorescein isothiocyanate specifically modifies lysine 338 of cytochrome P-450scc and inhibits adrenodoxin binding. *J. Biol. Chem.* **264**, 16 421–16 425. (doi:10.1016/S0021-9258(19)84723-3)

38. Poole LB. 2015 The basics of thiols and cysteines in redox biology and chemistry. *Free Radical Biol. Med.* **80**, 148–157. (doi:10.1016/j. freeradbiomed.2014.11.013)

39. Cal PMSD, Bernardes GJL, Gois PMP. 2014 Cysteine-selective reactions for antibody conjugation. *Angew. Chem. Int. Ed.* **53**, 10 585–10 587. (doi:10.1002/anie.201405702)

40. Gurd FRN. 1972 [34a] Carboxymethylation. *Methods Enzymol.* **25**, 424–438. (doi:10.1186/ s12885-016-2135-2)

41. Friedman M, Krull LH, Cavins JF. 1970 The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl)cysteine. *J. Biol. Chem.* **245**, 3868–3871. (doi:10.1016/S0021-9258(18)62930-8)

42. Sechi S, Chait BT. 1998 Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **70**, 5150–5158. (doi:10.1021/ac9806005)

43. Chin CCQ, Wold F. 1975 Some chemical properties of carboxymethyl derivatives of amino acids. *Arch. Biochem. Biophys.* **167**, 448–451. (doi:10.1016/0003-9861(75)90486-5)

44. Zhang C, Vinogradova EV, Spokoyny AM, Buchwald SL, Pentelute BL. 2019 Arylation chemistry for bioconjugation. *Angew. Chem. Int. Ed.* **58**, 4810–4839. (doi:10.1002/anie. 201806009)

45. Baldwin AD, Kiick KL. 2011 Tunable degradation of maleimide–thiol adducts in reducing environments. *Bioconjugate Chem.* **22**, 1946–1953. (doi:10.1021/bc2001auv)

46. Ravasco JMJM, Faustino H, Trindade A, Gois PMP. 2019 Bioconjugation with maleimides: a useful tool for chemical biology. *Chem. Eur. J.* **25**, 43–59. (doi:10.1002/chem.201803174)

47. Partis MD, Griffiths DG, Roberts GC, Beechey RB. 1983 Cross-linking of protein by ω-maleimido alkanoylN-hydroxysuccinimido esters. *J. Protein Chem.* **2**, 263–277. (doi:10.1007/BF01025358)

48. Gorin G, Martic PA, Doughty G. 1966 Kinetics of the reaction of N-ethylmaleimide with cysteine and some congeners. *Arch. Biochem. Biophys.* **115**, 593–597. (doi:10.1016/0003-9861(66)90079-8)

49. Brewer CF, Riehm JP. 1967 Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. *Anal. Biochem.* **18**, 248–255. (doi:10.1016/0003-2697(67)90007-3)

50. Gennari A, Wedgwood J, Lallana E, Francini N, Tirelli N. 2020 Thiol-based michael-type addition. A systematic evaluation of its controlling factors. *Tetrahedron* **76**, 131637. (doi:10.1016/j.tet.2020.131637)

51. Saito F, Noda H, Bode JW. 2015 Critical evaluation and rate constants of chemoselective ligation reactions for stoichiometric conjugations in water. *ACS Chem. Biol.* **10**, 1026–1033. (doi:10.1021/cb5006728)

52. Philippe S, Boeckler C, Frisch B, Schuber F. 2000 Differential reactivity of maleimide and bromoacetyl functions with thiols: application to the preparation of liposomal diepitope constructs. *Bioconjugate Chem.* **11**, 118–123. (doi:10.1021/bc990122k)

53. Smith MEB, Schumacher FF, Ryan CP, Tedaldi LM, Papaioannou D, Waksman G, Caddick S, Baker JR. 2010 Protein modification, bioconjugation, and disulfide bridging using bromomaleimides. *J. Am. Chem. Soc.* **132**, 1960–1965. (doi:10.1021/ja908610s)

54. Schumacher FF, Nobles M, Ryan CP, Smith MEB, Tinker A, Caddick S, Baker JR. 2011 *In situ* maleimide bridging of disulfides and a new approach to protein PEGylation. *Bioconjugate Chem.* **22**, 132–136. (doi:10.1021/bc1004685)

55. Chudasama V, Smith MEB, Schumacher FF, Papaioannou D, Waksman G, Baker JR, Caddick S. 2011 Bromopyridazinedione-mediated protein and peptide bioconjugation. *Chem. Commun.* **47**, 8781–8783. (doi:10.1039/c1cc12807h)

56. Fontaine SD, Reid R, Robinson L, Ashley GW, Santi DV. 2015 Long-term stabilization of maleimide–thiol conjugates. *Bioconjugate Chem.* **26**, 145–152. (doi:10.1021/bc5005262)

57. Kalia D, Pawar SP, Thopate JS. 2017 Stable and rapid thiol bioconjugation by light-triggered thiomaleimide ring hydrolysis. *Angew. Chem. Int. Ed.* **56**, 1885–1889. (doi:10.1002/anie. 201609733)

58. Kalia D, Malekar PV, Parthasarathy M. 2016 Exocyclic olefinic maleimides: synthesis and application for stable and thiol-selective bioconjugation. *Angew. Chem. Int. Ed.* **55**, 1432–1435. (doi:10.1002/anie.201508118)

59. De Geyter E, Antonatou E, Kalaitzakis D, Smolen S, Iyer A, Tack L, Ongenae E, Vassilikogiannakis G, Madder A. 2021 5-Hydroxy-pyrrolone based building blocks as maleimide alternatives for protein bioconjugation and single-site multi-functionalization. *Chem. Sci.* **12**, 5246–5252. (doi:10.1039/D0SC05881E)

60. Bernardim B *et al.* 2016 Stoichiometric and irreversible cysteine-selective protein modification using carbonylacrylic reagents. *Nat.*

Commun. **7**, 13128. (doi:10.1038/ncomms13128)

61. Ariyasu S, Hayashi H, Xing B, Chiba S. 2017 Site-specific dual functionalization of cysteine residue in peptides and proteins with 2-azidoacrylates. *Bioconjugate Chem.* **28**, 897–902. (doi:10.1021/acs.bioconjchem.7b00024)

62. Bernardim B, Dunsmore L, Li H, Hocking B, Nuñez-Franco R, Navo CD, Jiménez-Osés G, Burtoloso ACB, Bernardes GJL. 2020 Precise installation of diazo-tagged side-chains on proteins to enable *in vitro* and in-cell site-specific labeling. *Bioconjugate Chem.* **31**, 1604–1610. (doi:10.1021/acs.bioconjchem.0c00232)

63. Masri MS, Friedman M. 1988 Protein reactions with methyl and ethyl vinyl sulfones. *J. Protein Chem.* **7**, 49–54. (doi:10.1007/BF01025413)

64. Morpurgo M, Veronese FM, Kachensky D, Harris JM. 1996 Preparation and characterization of poly(ethylene glycol) vinyl sulfone. *Bioconjugate Chem.* **7**, 363–368. (doi:10.1021/bc9600224)

65. Morales-Sanfrutos J, Lopez-Jaramillo J, Ortega-Muñoz M, Megia-Fernandez A, Perez-Balderas F, Hernandez-Mateo F, Santoyo-Gonzalez F. 2010 Vinyl sulfone: a versatile function for simple bioconjugation and immobilization. *Org. Biomol. Chem.* **8**, 667–675. (doi:10.1039/B920576D)

66. Gil de Montes E, Jiménez-Moreno E, Oliveira BL, Navo CD, Cal PMSD, Jiménez-Osés G, Robina I, Moreno-Vargas AJ, Bernardes GJL. 2019 Azabicyclic vinyl sulfones for residue-specific dual protein labelling. *Chem. Sci.* **10**, 4515–4522. (doi:10.1039/C9SC00125E)

67. Kasper M-A *et al.* 2019 Cysteine-selective phosphonamidate electrophiles for modular protein bioconjugations. *Angew. Chem. Int. Ed.* **58**, 11 625–11 630. (doi:10.1002/anie.201814715)

68. Baumann AL, Schwagerus S, Broi K, Kemnitz-Hassanin K, Stieger CE, Trieloff N, Schmieder P, Hackenberger CPR. 2020 Chemically induced vinylphosphonothiolate electrophiles for thiol–thiol bioconjugations. *J. Am. Chem. Soc.* **142**, 9544–9552. (doi:10.1021/jacs.0c03426)

69. Stieger CE, Franz L, Körlin F, Hackenberger CPR. 2021 Diethynyl phosphinates for cysteine-selective protein labeling and disulfide rebridging. *Angew. Chem. Int. Ed.* **60**, 15 359–15 364. (doi:10.1002/anie.202100683)

70. Abegg D, Frei R, Cerato L, Prasad Hari D, Wang C, Waser J, Adibekian A. 2015 Proteome-wide profiling of targets of cysteine reactive small molecules by using ethynyl benziodoxolone reagents. *Angew. Chem. Int. Ed.* **54**, 10 852–10 857. (doi:10.1002/anie.201505641)

71. Tessier R, Ceballos J, Guidotti N, Simonet-Davin R, Fierz B, Waser J. 2019 'Doubly orthogonal' labeling of peptides and proteins. *Chem* **5**, 2243–2263. (doi:10.1016/j.chempr.2019.06.022)

72. Tessier R *et al.* 2020 Ethynylation of cysteine residues: from peptides to proteins *in vitro* and in living cells. *Angew. Chem. Int. Ed.* **59**, 10 961–10 970. (doi:10.1002/anie.202002626)

73. Joshi NS, Whitaker LR, Francis MB. 2004 A three-component mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **126**, 15 942–15 943. (doi:10.1021/ja0439017)

74. Pauly H. 1904 Über die Konstitution des Histidins. I. Mitteilung. *Hoppe-Seyleŕs Zeitschrift für physiologische Chemie* **42**, 508–518. (doi:10.1515/bchm2.1904.42.5-6.508)

75. Leier S, Richter S, Bergmann R, Wuest M, Wuest F. 2019 Radiometal-containing aryl diazonium salts for chemoselective bioconjugation of tyrosine residues. *ACS Omega* **4**, 22 101–22 107. (doi:10.1021/acsomega.9b03248)

76. Szijj PA, Kostadinova KA, Spears RJ, Chudasama V. 2020 Tyrosine bioconjugation – an emergent alternative. *Org. Biomol. Chem.* **18**, 9018–9028. (doi:10.1039/D0OB01912G)

77. Dorta DA, Deniaud D, Mével M, Gouin SG. 2020 Tyrosine conjugation methods for protein labelling. *Chem. Eur. J.* **26**, 14 257–14 269. (doi:10.1002/chem.202001992)

78. Takahashi K. 1968 The reaction of phenylglyoxal with arginine residues in proteins. *J. Biol. Chem.* **243**, 6171–6179. (doi:10.1016/S0021-9258(18)94475-3)

79. Schwarzenbolz U, Henle T, Haeßner R, Klostermeyer H. 1997 On the reaction of glyoxal with proteins. *Z. Lebensm. Unters. Forsch.* **205**, 121–124. (doi:10.1007/s002170050137)

80. Kato H, van Chuyen N, Utsunomiya N, Okitani A. 1986 Changes of amino acids composition and relative digestibility of lysozyme in the reaction with .ALPHA.-dicarbonyl compounds in aqueous system. *J. Nutr. Sci. Vitaminol. (Tokyo)* **32**, 55–65. (doi:10.3177/jnsv.32.55)

81. Dovgan I, Erb S, Hessmann S, Ursuegui S, Michel C, Muller C, Chaubet G, Cianférani S, Wagner A. 2018 Arginine-selective bioconjugation with 4-azidophenyl glyoxal: application to the single and dual functionalisation of native antibodies. *Org. Biomol. Chem.* **16**, 1305–1311. (doi:10.1039/C7OB02844J)

82. Tamura T, Hamachi I. 2014 Recent progress in design of protein-based fluorescent biosensors and their cellular applications. *ACS Chem. Biol.* **9**, 2708–2717. (doi:10.1021/cb500661v)

83. Nischan N, Hackenberger CPR. 2014 Site-specific PEGylation of proteins: recent developments. *J. Org. Chem.* **79**, 10 727–10 733. (doi:10.1021/jo502136n)

84. Turecek PL, Bossard MJ, Schoetens F, Ivens IA. 2016 PEGylation of biopharmaceuticals: a review of chemistry and nonclinical safety information of approved drugs. *J. Pharm. Sci.* **105**, 460–475.

85. Lambert JM, Berkenblit A. 2018 Antibody–drug conjugates for cancer treatment. *Annu. Rev. Med.* **69**, 191–207. (doi:10.1146/annurev-med-061516-121357)

86. Beck A, Goetsch L, Dumontet C, Corvaïa N. 2017 Strategies and challenges for the next generation of antibody–drug conjugates. *Nat. Rev. Drug Discovery* **16**, 315–337. (doi:10.1038/nrd.2016.268)

87. Schumacher D, Hackenberger CPR, Leonhardt H, Helma J. 2016 Current status: site-specific antibody drug conjugates. *J. Clin. Immunol.* **36**, 100–107. (doi:10.1007/s10875-016-0265-6)

88. Gautier V, Boumeester AJ, Lössl P, Heck AJR. 2015 Lysine conjugation properties in human IgGs studied by integrating high-resolution native mass spectrometry and bottom-up proteomics. *Proteomics* **15**, 2756–2765. (doi:10.1002/pmic.201400462)

89. Junutula JR *et al.* 2008 Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index. *Nat. Biotechnol.* **26**, 925–932. (doi:10.1038/nbt.1480)

90. Rawale DG, Thakur K, Sreekumar P, Sajeev TK, Ramesh A, Adusumalli R, Mishra RK, Rai V. 2021 Linchpins empower promiscuous electrophiles to enable site-selective modification of histidine and aspartic acid in proteins. *Chem. Sci.* **12**, 6732–6736. (doi:10.1039/D1SC00335F)

91. Yamada K, Ito Y. 2019 Recent chemical approaches for site-specific conjugation of native antibodies: technologies toward next-generation antibody–drug conjugates. *Chembiochem* **20**, 2729–2737. (doi:10.1002/cbic.201900178)

92. Bai C *et al.* 2020 Site-specific conjugation of the indolinobenzodiazepine DGN549 to antibodies affords antibody–drug conjugates with an improved therapeutic index as compared with lysine conjugation. *Bioconjugate Chem.* **31**, 93–103. (doi:10.1021/acs.bioconjchem.9b00777)

93. Sadiki A *et al.* 2020 Site-specific conjugation of native antibody. *Antib. Ther.* **3**, 271–284. (doi:10.1093/abt/tbaa027)

94. Procter G. 1998 *Stereoselectivity in organic synthesis*. Oxford, New York: Oxford University Press.

95. Rawale DG, Thakur K, Adusumalli SR, Rai V. 2019 Chemical methods for selective labeling of proteins. *Eur. J. Org. Chem.* **2019**, 6749–6763. (doi:10.1002/ejoc.201900801)

96. Hoyt EA, Cal PMSD, Oliveira BL, Bernardes GJL. 2019 Contemporary approaches to site-selective protein modification. *Nat. Rev. Chem.* **3**, 147–171. (doi:10.1038/s41570-019-0079-1)

97. Chalker JM, Bernardes GJL, Lin YA, Davis BG. 2009 Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J.* **4**, 630–640. (doi:10.1002/asia.200800427)

98. Dawson PE, Muir TW, Clark-Lewis I, Kent SB. 1994 Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779. (doi:10.1126/science.7973629)

99. Agouridas V, El Mahdi O, Diemer V, Cargoët M, Monbaliu J-CM, Melnyk O. 2019 Native chemical ligation and extended methods: mechanisms, catalysis, scope, and limitations. *Chem. Rev.* **119**, 7328–7443. (doi:10.1021/acs.chemrev.8b00712)

100. Dadová J, Galan SR, Davis BG. 2018 Synthesis of modified proteins via functionalization of dehydroalanine. *Curr. Opin Chem. Biol.* **46**, 71–81. (doi:10.1016/j.cbpa.2018.05.022)

101. Chalker JM *et al.* 2011 Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2**, 1666–1676. (doi:10.1039/c1sc00185j)

102. Freedy AM *et al.* 2017 Chemoselective installation of amine bonds on proteins through Aza-Michael ligation. *J. Am. Chem. Soc.* **139**, 18 365–18 375. (doi:10.1021/jacs.7b10702)

103. Isenegger PG, Davis BG. 2019 Concepts of catalysis in site-selective protein modifications. *J. Am. Chem. Soc.* **141**, 8005–8013. (doi:10.1021/jacs.8b13187)

104. Sakamoto S, Hamachi I. 2019 Recent progress in chemical modification of proteins. *Anal. Sci.* **35**, 5–27. (doi:10.2116/analsci.18R003)

105. Spicer CD, Davis BG. 2014 Selective chemical protein modification. *Nat. Commun.* **5**, 4740. (doi:10.1038/ncomms5740)

106. Chin JW. 2017 Expanding and reprogramming the genetic code. *Nature* **550**, 53–60. (doi:10.1038/nature24031)

107. Saxon E, Bertozzi CR. 2000 Cell surface engineering by a modified Staudinger reaction. *Science* **287**, 2007–2010. (doi:10.1126/science.287.5460.2007)

108. Bednarek C, Wehl I, Jung N, Schepers U, Bräse S. 2020 The Staudinger ligation. *Chem. Rev.* **120**, 4301–4354. (doi:10.1021/acs.chemrev.9b00665)

109. Tsao M-L, Tian F, Schultz PG. 2005 Selective Staudinger modification of proteins containing p-Azidophenylalanine. *Chembiochem* **6**, 2147–2149. (doi:10.1002/cbic.200500314)

110. Serwa R, Wilkening I, DelSignore G, Mühlberg M, Claußnitzer I, Weise C, Gerrits M, Hackenberger CPR. 2009 Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. *Angew. Chem. Int. Ed.* **48**, 8234–8239. (doi:10.1002/anie.200902118)

111. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. 2002 Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl Acad. Sci. USA* **99**, 19–24. (doi:10.1073/pnas.012583299)

112. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. 2006 A comparative study of bioorthogonal reactions with azides. *ACS Chem. Biol.* **1**, 644–648. (doi:10.1021/cb6003228)

113. Bernardes GJL, Linderoth L, Doores KJ, Boutureira O, Davis BG. 2011 Site-selective traceless staudinger ligation for glycoprotein synthesis reveals scope and limitations. *Chembiochem* **12**, 1383–1386. (doi:10.1002/cbic.201100125)

114. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. 2002 A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **41**, 2596–2599. (doi:10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4)

115. Tornøe CW, Christensen C, Meldal M. 2002 Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064. (doi:10.1021/jo011148j)

116. Hong V, Steinmetz NF, Manchester M, Finn MG. 2010 Labeling live cells by copper-catalyzed alkyne—azide click chemistry. *Bioconjugate Chem.* **21**, 1912–1916. (doi:10.1021/bc100272z)

117. Agard NJ, Prescher JA, Bertozzi CR. 2004 A strain-promoted [3+2] azide—alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15 046–15 047. (doi:10.1021/ja044996f)

118. Blomquist AT, Liu LH. 1953 Many-membered carbon rings. VII. Cycloöctyne. *J. Am. Chem. Soc.* **75**, 2153–2154. (doi:10.1021/ja01105a039)

119. Wittig G, Krebs A. 1961 Zur Existenz niedergliedriger Cycloalkine, I. *Chem. Ber.* **94**, 3260–3275. (doi:10.1002/cber.19610941213)

120. Codelli JA, Baskin JM, Agard NJ, Bertozzi CR. 2008 Second-generation difluorinated cyclooctynes for copper-free click chemistry. *J. Am. Chem. Soc.* **130**, 11 486–11 493. (doi:10.1021/ja803086r)

121. Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, Lo A, Codelli JA, Bertozzi CR. 2007 Copper-free click chemistry for dynamic *in vivo* imaging. *Proc. Natl Acad. Sci. USA* **104**, 16 793–16 797. (doi:10.1073/pnas.0707090104)

122. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. 2008 *In vivo* imaging of membrane-associated glycans in developing zebrafish. *Science* **320**, 664–667. (doi:10.1126/science.1155106)

123. Ning X, Guo J, Wolfert MA, Boons G-J. 2008 Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast Huisgen cycloadditions. *Angew. Chem. Int. Ed.* **47**, 2253–2255. (doi:10.1002/anie.200705456)

124. Debets MF, van Berkel SS, Schoffelen S, Rutjes FPJT, van Hest JCM, van Delft FL. 2010 Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. *Chem. Commun.* **46**, 97–99. (doi:10.1039/B917797C)

125. Jewett JC, Sletten EM, Bertozzi CR. 2010 Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *J. Am. Chem. Soc.* **132**, 3688–3690. (doi:10.1021/ja100014q)

126. Krebs A, Kimling H. 1970 3.3.6.6-tetramethyl-1-thiacyclohpeptin ein isolierbares siebenring-acetylen. *Tetrahedron Lett.* **11**, 761–764. (doi:10.1016/S0040-4039(01)97823-2)

127. King M, Baati R, Wagner A. 2012 New tetramethylthiepinium (TMTI) for copper-free click chemistry. *Chem. Commun.* **48**, 9308–9309. (doi:10.1039/c2cc35034c)

128. Weterings J, Rijcken CJF, Veldhuis H, Meulemans T, Hadavi D, Timmers M, Honing M, Ippel H, Liskamp RMJ. 2020 TMTHSI, a superior 7-membered ring alkyne containing reagent for strain-promoted azide–alkyne cycloaddition reactions. *Chem. Sci.* **11**, 9011–9016. (doi:10.1039/D0SC03477K)

129. Dommerholt J, Schmidt S, Temming R, Hendriks LJA, Rutjes FPJT, vanHest JCM, Lefeber DJ, Friedl P, vanDelft FL. 2010 Readily accessible bicyclononynes for bioorthogonal labeling and three-dimensional imaging of living cells. *Angew. Chem. Int. Ed.* **49**, 9422–9425. (doi:10.1002/anie.201003761)

130. Plass T, Milles S, Koehler C, Schultz C, Lemke EA. 2011 Genetically encoded copper-free click chemistry. *Angew. Chem. Int. Ed.* **50**, 3878–3881. (doi:10.1002/anie.201008178)

131. Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. 2012 Genetic encoding of bicyclononynes and *trans*-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels–Alder reactions. *J. Am. Chem. Soc.* **134**, 10 317–10 320. (doi:10.1021/ja302832g)

132. Blackman ML, Royzen M, Fox JM. 2008 Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels—Alder reactivity.

J. Am. Chem. Soc. **130**, 13 518–13 519. (doi:10.1021/ja8053805)

133. Devaraj NK, Weissleder R, Hilderbrand SA. 2008 Tetrazine-based cycloadditions: application to pretargeted live cell imaging. *Bioconjugate Chem.* **19**, 2297–2299. (doi:10.1021/bc8004446)

134. Taylor MT, Blackman ML, Dmitrenko O, Fox JM. 2011 Design and synthesis of highly reactive dienophiles for the tetrazine—*trans*-cyclooctene ligation. *J. Am. Chem. Soc.* **133**, 9646–9649. (doi:10.1021/ja201844c)

135. Plass T, Milles S, Koehler C, Szymański J, Mueller R, Wießler M, Schultz C, Lemke EA. 2012 Amino acids for Diels-Alder reactions in living cells. *Angew. Chem. Int. Ed.* **51**, 4166–4170. (doi:10.1002/anie.201108231)

136. Lang K, Davis L, Torres-Kolbus J, Chou C, Deiters A, Chin JW. 2012 Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem* **4**, 298–304. (doi:10.1038/nchem.1250)

137. Kaya E, Vrabel M, Deiml C, Prill S, Fluxa VS, Carell T. 2012 A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. *Angew. Chem. Int. Ed.* **51**, 4466–4469. (doi:10.1002/anie.201109252)

138. Kang K, Park J, Kim E. 2017 Tetrazine ligation for chemical proteomics. *Proteome Sci.* **15**, 15. (doi:10.1186/s12953-017-0121-5)

139. Song W, Wang Y, Qu J, Madden MM, Lin Q. 2008 A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angew. Chem. Int. Ed.* **47**, 2832–2835. (doi:10.1002/anie.200705805)

140. Wang J et al. 2010 A biosynthetic route to photoclick chemistry on proteins. *J. Am. Chem. Soc.* **132**, 14 812–14 818. (doi:10.1021/ja104350y)

141. Song W, Wang Y, Qu J, Lin Q. 2008 Selective functionalization of a genetically encoded alkene-containing protein via 'photoclick chemistry' in bacterial cells. *J. Am. Chem. Soc.* **130**, 9654–9655. (doi:10.1021/ja803598e)

142. Ramil CP, Lin Q. 2014 Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. *Curr. Opin Chem. Biol.* **21**, 89–95. (doi:10.1016/j.cbpa.2014.05.024)

143. Yu Z, Pan Y, Wang Z, Wang J, Lin Q. 2012 Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. *Angew. Chem. Int. Ed.* **51**, 10 600–10 604. (doi:10.1002/anie.201205352)

144. Dirksen A, Dawson PE. 2008 Rapid oxime and hydrazone ligations with aromatic aldehydes for biomolecular labeling. *Bioconjugate Chem.* **19**, 2543–2548. (doi:10.1021/bc800310p)

145. Geoghegan KF, Stroh JG. 1992 Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine. *Bioconjugate Chem.* **3**, 138–146. (doi:10.1021/bc00014a008)

146. Cornish VW, Hahn KM, Schultz PG. 1996 Site-specific protein modification using a ketone handle. *J. Am. Chem. Soc.* **118**, 8150–8151. (doi:10.1021/ja961216x)

147. Zeng H, Xie J, Schultz PG. 2006 Genetic introduction of a diketone-containing amino acid into proteins. *Bioorg. Med. Chem. Lett.* **16**, 5356–5359. (doi:10.1016/j.bmcl.2006.07.094)

148. Wang L, Zhang Z, Brock A, Schultz PG. 2003 Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **100**, 56–61. (doi:10.1073/pnas.0234824100)

149. Zhang Z, Smith BAC, Wang L, Brock A, Cho C, Schultz PG. 2003 A new strategy for the site-specific modification of proteins *in vivo*. *Biochemistry* **42**, 6735–6746. (doi:10.1021/bi0300231)

150. Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG. 2003 An expanded eukaryotic genetic code. *Science* **301**, 964–967.

151. Liu H, Wang L, Brock A, Wong C-H, Schultz PG. 2003 A method for the generation of glycoprotein mimetics. *J. Am. Chem. Soc.* **125**, 1702–1703. (doi:10.1021/ja029433n)

152. Ye S, Köhrer C, Huber T, Kazmi M, Sachdev P, Yan ECY, Bhagat A, RajBhandary UL, Sakmar TP. 2008 Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. *J. Biol. Chem.* **283**, 1525–1533. (doi:10.1074/jbc.M707355200)

153. Hutchins BM, Kazane SA, Staflin K, Forsyth JS, Felding-Habermann B, Schultz PG, Smider VV. 2011 Site-specific coupling and sterically controlled formation of multimeric antibody fab fragments with unnatural amino acids. *J. Mol. Biol.* **406**, 595–603. (doi:10.1016/j.jmb.2011.01.011)

154. Kim CH, Axup JY, Dubrovska A, Kazane SA, Hutchins BA, Wold ED, Smider VV, Schultz PG. 2012 Synthesis of bispecific antibodies using genetically encoded unnatural amino acids. *J. Am. Chem. Soc.* **134**, 9918–9921. (doi:10.1021/ja303904e)

155. Huang Y, Wan W, Russell WK, Pai P-J, Wang Z, Russell DH, Liu W. 2010 Genetic incorporation of an aliphatic keto-containing amino acid into proteins for their site-specific modifications. *Bioorg. Med. Chem. Lett.* **20**, 878–880. (doi:10.1016/j.bmcl.2009.12.077)

156. Antos JM, Francis MB. 2006 Transition metal catalyzed methods for site-selective protein modification. *Curr. Opin Chem. Biol.* **10**, 253–262. (doi:10.1016/j.cbpa.2006.04.009)

157. Ojida A, Tsutsumi H, Kasagi N, Hamachi I. 2005 Suzuki coupling for protein modification. *Tetrahedron Lett.* **46**, 3301–3305. (doi:10.1016/j.tetlet.2005.03.094)

158. Brustad E, Bushey ML, Lee JW, Groff D, Liu W, Schultz PG. 2008 A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed.* **47**, 8220–8223. (doi:10.1002/anie.200803240)

159. Chalker JM, Wood CSC, Davis BG. 2009 A convenient catalyst for aqueous and protein Suzuki–Miyaura cross-coupling. *J. Am. Chem. Soc.* **131**, 16 346–16 347. (doi:10.1021/ja907150m)

160. Spicer CD, Davis BG. 2011 Palladium-mediated site-selective Suzuki–Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun.* **47**, 1698–1700. (doi:10.1039/c0cc04970k)

161. Wang Y-S, Russell WK, Wang Z, Wan W, Dodd LE, Pai P-J, Russell DH, Liu WR. 2011 The de novo engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of l-phenylalanine and its derivatives. *Mol. Biosyst.* **7**, 714–717. (doi:10.1039/c0mb00217h)

162. Spicer CD, Triemer T, Davis BG. 2012 Palladium-mediated cell-surface labeling. *J. Am. Chem. Soc.* **134**, 800–803. (doi:10.1021/ja209352s)

163. Spicer CD, Davis BG. 2013 Rewriting the bacterial glycocalyx via Suzuki–Miyaura cross-coupling. *Chem. Commun.* **49**, 2747–2749. (doi:10.1039/c3cc38824g)

164. Dumas A, Spicer CD, Gao Z, Takehana T, Lin YA, Yasukohchi T, Davis BG. 2013 Self-liganded Suzuki-Miyaura coupling for site-selective protein PEGylation. *Angew. Chem. Int. Ed.* **52**, 3916–3921. (doi:10.1002/anie.201208626)

165. Gao Z, Gouverneur V, Davis BG. 2013 Enhanced aqueous Suzuki–Miyaura coupling allows site-specific polypeptide ^{18}F-labeling. *J. Am. Chem. Soc.* **135**, 13 612–13 615. (doi:10.1021/ja4049114)

166. Li J, Lin S, Wang J, Jia S, Yang M, Hao Z, Zhang X, Chen PR. 2013 Ligand-free palladium-mediated site-specific protein labeling inside gram-negative bacterial pathogens. *J. Am. Chem. Soc.* **135**, 7330–7338. (doi:10.1021/ja402424j)

167. Kodama K *et al.* 2006 Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *Chembiochem* **7**, 134–139. (doi:10.1002/cbic.200500290)

168. Kodama K *et al.* 2007 Site-specific functionalization of proteins by organopalladium reactions. *Chembiochem* **8**, 232–238. (doi:10.1002/cbic.200600432)

169. Li N, Lim RKV, Edwardraja S, Lin Q. 2011 Copper-free Sonogashira cross-coupling for functionalization of alkyne-encoded proteins in aqueous medium and in bacterial cells. *J. Am. Chem. Soc.* **133**, 15 316–15 319. (doi:10.1021/ja2066913)

170. Lin YA, Chalker JM, Davis BG. 2009 Olefin metathesis for site-selective protein modification. *Chembiochem* **10**, 959–969. (doi:10.1002/cbic.200900002)

171. Messina MS, Maynard HD. 2020 Modification of proteins using olefin metathesis. *Mater. Chem. Front.* **4**, 1040–1051. (doi:10.1039/C9QM00494G)

172. Bhushan B *et al.* 2018 Genetic incorporation of olefin cross-metathesis reaction tags for protein modification. *J. Am. Chem. Soc.* **140**, 14 599–14 603. (doi:10.1021/jacs.8b09433)

173. Ai H, Shen W, Brustad E, Schultz PG. 2010 Genetically encoded alkenes in yeast. *Angew. Chem. Int. Ed.* **49**, 935–937. (doi:10.1002/anie.200905590)

174. Lin YA, Boutureira O, Lercher L, Bhushan B, Paton RS, Davis BG. 2013 Rapid cross-metathesis for reversible protein modifications via chemical access to Se-allyl-selenocysteine in proteins. *J. Am. Chem. Soc.* **135**, 12 156–12 159. (doi:10.1021/ja403191g)

175. Lin YA, Chalker JM, Floyd N, Bernardes GJL, Davis BG. 2008 Allyl sulfides are privileged substrates in aqueous cross-metathesis: application to site-selective protein modification. *J. Am. Chem. Soc.* **130**, 9642–9643. (doi:10.1021/ja8026168)

176. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. 2003 A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89. (doi:10.1038/nbt765)

177. Yang G *et al.* 2015 Genetic targeting of chemical indicators in vivo. *Nat. Methods* **12**, 137–139. (doi:10.1038/nmeth.3207)

178. Gautier A, Juillerat A, Heinis C, Corrêa IR, Kindermann M, Beaufils F, Johnsson K. 2008 An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128–136. (doi:10.1016/j.chembiol.2008.01.007)

179. Los GV *et al.* 2008 HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382. (doi:10.1021/cb800025k)

180. Hori Y, Ueno H, Mizukami S, Kikuchi K. 2009 Photoactive yellow protein-based protein labeling system with turn-on fluorescence intensity. *J. Am. Chem. Soc.* **131**, 16 610–16 611. (doi:10.1021/ja904800k)

181. Hori Y, Norinobu T, Sato M, Arita K, Shirakawa M, Kikuchi K. 2013 Development of fluorogenic probes for quick no-wash live-cell imaging of intracellular proteins. *J. Am. Chem. Soc.* **135**, 12 360–12 365. (doi:10.1021/ja405745v)

182. Mizukami S, Watanabe S, Hori Y, Kikuchi K. 2009 Covalent protein labeling based on noncatalytic β-lactamase and a designed FRET substrate. *J. Am. Chem. Soc.* **131**, 5016–5017. (doi:10.1021/ja8082285)

183. Griffin BA, Adams SR, Tsien RY. 1998 Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269–272. (doi:10.1126/science.281.5374.269)

184. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, Llopis J, Tsien RY. 2002 New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J. Am. Chem. Soc.* **124**, 6063–6076. (doi:10.1021/ja017687n)

185. Halo TL, Appelbaum J, Hobert EM, Balkin DM, Schepartz A. 2009 Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid. *J. Am. Chem. Soc.* **131**, 438–439. (doi:10.1021/ja807872s)

186. Uchinomiya S, Nonaka H, Fujishima S, Tsukiji S, Ojida A, Hamachi I. 2009 Site-specific covalent labeling of His-tag fused proteins with a reactive Ni(ii)–NTA probe. *Chem. Commun.* **39**, 5880–5882. (doi:10.1039/b912025d)

187. Martos-Maldonado MC *et al.* 2018 Selective N-terminal acylation of peptides and proteins with a Gly-His tag sequence. *Nat. Commun.* **9**, 3307. (doi:10.1038/s41467-018-05695-3)

188. Zhang C, Welborn M, Zhu T, Yang NJ, Santos MS, Van Voorhis T, Pentelute BL. 2016 π-Clamp-mediated cysteine conjugation. *Nat. Chem.* **8**, 120–128. (doi:10.1038/nchem.2413)

189. Appel MJ *et al.* 2019 Formylglycine-generating enzyme binds substrate directly at a mononuclear Cu(I) center to initiate O_2 activation. *Proc. Natl Acad. Sci. USA* **116**, 5370–5375. (doi:10.1073/pnas.1818274116)

190. Krüger T, Dierks T, Sewald N. 2019 Formylglycine-generating enzymes for site-specific bioconjugation. *Biol. Chem.* **400**, 289–297. (doi:10.1515/hsz-2018-0358)

191. Rabuka D, Rush JS, deHart GW, Wu P, Bertozzi CR. 2012 Site-specific chemical protein

conjugation using genetically encoded aldehyde tags. *Nat. Protoc.* **7**, 1052–1067. (doi:10.1038/nprot.2012.045)

192. Holder PG, Jones LC, Drake PM, Barfield RM, Bañas S, de Hart GW, Baker J, Rabuka D. 2015 Reconstitution of formylglycine-generating enzyme with copper(II) for aldehyde tag conversion. *J. Biol. Chem.* **290**, 15 730–15 745. (doi:10.1074/jbc.M115.652669)

193. York D, Baker J, Holder PG, Jones LC, Drake PM, Barfield RM, Bleck GT, Rabuka D. 2016 Generating aldehyde-tagged antibodies with high titers and high formylglycine yields by supplementing culture media with copper(II). *BMC Biotechnol.* **16**, 23. (doi:10.1186/s12896-016-0254-0)

194. Tamura T, Hamachi I. 2019 Chemistry for covalent modification of endogenous/native proteins: from test tubes to complex biological systems. *J. Am. Chem. Soc.* **141**, 2782–2799. (doi:10.1021/jacs.8b11747)

195. Chen X, Muthoosamy K, Pfisterer A, Neumann B, Weil T. 2012 Site-selective lysine modification of native proteins and peptides via kinetically controlled labeling. *Bioconjugate Chem.* **23**, 500–508. (doi:10.1021/bc200556n)

196. United States, US8741291B2, 2014.

197. Isom DG, Castañeda CA, Cannon BR. 2011 Large shifts in pKa values of lysine residues buried inside a protein. *Proc. Natl Acad. Sci. USA* **108**, 5260–5265. (doi:10.1073/pnas.1010750108)

198. Asano S, Patterson JT, Gaj T, Barbas CF. 2014 Site-selective labeling of a lysine residue in human serum albumin. *Angew. Chem. Int. Ed.* **53**, 11 783–11 786. (doi:10.1002/anie.201405924)

199. Matos MJ et al. 2018 Chemo- and regioselective lysine modification on native proteins. *J. Am. Chem. Soc.* **140**, 4004–4017. (doi:10.1021/jacs.7b12874)

200. Purushottam L, Adusumalli SR, Chilamari M, Rai V. 2017 Chemoselective and site-selective peptide and native protein modification enabled by aldehyde auto-oxidation. *Chem. Commun.* **53**, 959–962. (doi:10.1039/C6CC09555K)

201. Chilamari M, Purushottam L, Rai V. 2017 Site-selective labeling of native proteins by a multicomponent approach. *Chem. A Eur. J.* **23**, 3819–3823. (doi:10.1002/chem.201605938)

202. Chilamari M, Kalra N, Shukla S, Rai V. 2018 Single-site labeling of lysine in proteins through a metal-free multicomponent approach. *Chem. Commun.* **54**, 7302–7305. (doi:10.1039/C8CC03311K)

203. Rosen CB et al. 2014 Template-directed covalent conjugation of DNA to native antibodies, transferrin and other metal-binding proteins. *Nat. Chem.* **6**, 804–809. (doi:10.1038/nchem.2003)

204. Forte N, Benni I, Karu K, Chudasama V, Baker JR. 2019 Cysteine-to-lysine transfer antibody fragment conjugation. *Chem. Sci.* **10**, 10 919–10 924. (doi:10.1039/C9SC03825F)

205. Adusumalli SR, Rawale DG, Thakur K, Purushottam L, Reddy NC, Kalra N, Shukla S, Rai V. 2020 Chemoselective and site-selective lysine-directed lysine modification enables single-site labeling of native proteins. *Angew. Chem. Int. Ed.* **59**, 10 332–10 336. (doi:10.1002/anie.202000062)

206. Akkapeddi P, Azizi S-A, Freedy AM, Cal PMSD, Gois PMP, Bernardes GJL. 2016 Construction of homogeneous antibody–drug conjugates using site-selective protein chemistry. *Chem. Sci.* **7**, 2954–2963. (doi:10.1039/C6SC00170J)

207. S. Mitra and R. G. Lawton, Reagents for the crosslinking of proteins by equilibrium transfer alkylation *J. Am. Chem. Soc.* 1979, **101**, 3097–3110. (doi:10.1021/ja00505a043)

208. Balan S, Choi J, Godwin A, Teo I, Laborde CM, Heidelberger S, Zloh M, Shaunak S, Brocchini S. 2007 Site-specific PEGylation of protein disulfide bonds using a three-carbon bridge. *Bioconjugate Chem.* **18**, 61–76. (doi:10.1021/bc0601471)

209. Shaunak S et al. 2006 Site-specific PEGylation of native disulfide bonds in therapeutic proteins. *Nat. Chem. Biol.* **2**, 312–313. (doi:10.1038/nchembio786)

210. Xu L, Raabe M, Zegota MM, Nogueira JCF, Chudasama V, Kuan SL, Weil T. 2020 Site-selective protein modification via disulfide rebridging for fast tetrazine/trans-cyclooctene bioconjugation. *Org. Biomol. Chem.* **18**, 1140–1147. (doi:10.1039/C9OB02687H)

211. Badescu G et al. 2014 Bridging disulfides for stable and defined antibody drug conjugates. *Bioconjugate Chem.* **25**, 1124–1136. (doi:10.1021/bc500148x)

212. Schumacher FF, Nunes JPM, Maruani A, Chudasama V, Smith MEB, Chester KA, Baker JR, Caddick S. 2014 Next generation maleimides enable the controlled assembly of antibody–drug conjugates via native disulfide bond bridging. *Org. Biomol. Chem.* **12**, 7261–7269. (doi:10.1039/C4OB01550A)

213. Nunes JPM et al. 2015 Functional native disulfide bridging enables delivery of a potent, stable and targeted antibody–drug conjugate (ADC). *Chem. Commun.* **51**, 10 624–10 627. (doi:10.1039/C5CC03557K)

214. Bryden F, Maruani A, Savoie H, Chudasama V, Smith MEB, Caddick S, Boyle RW. 2014 Regioselective and stoichiometrically controlled conjugation of photodynamic sensitizers to a HER2 targeting antibody fragment. *Bioconjugate Chem.* **25**, 611–617. (doi:10.1021/bc5000324)

215. Behrens CR et al. 2015 Antibody–drug conjugates (ADCs) derived from interchain cysteine cross-linking demonstrate improved homogeneity and other pharmacological properties over conventional heterogeneous ADCs. *Mol. Pharmaceutics* **12**, 3986–3998. (doi:10.1021/acs.molpharmaceut.5b00432)

216. Maruani A, Smith MEB, Miranda E, Chester KA, Chudasama V, Caddick S. 2015 A plug-and-play approach to antibody-based therapeutics via a chemoselective dual click strategy. *Nat. Commun.* **6**, 1–9. (doi:10.1038/ncomms7645)

217. Robinson E et al. 2017 Pyridazinediones deliver potent, stable, targeted and efficacious antibody–drug conjugates (ADCs) with a controlled loading of 4 drugs per antibody. *RSC Adv.* **7**, 9073–9077. (doi:10.1039/C7RA00788D)

218. Bahou C, Richards DA, Maruani A, Love EA, Javaid F, Caddick S, Baker JR, Chudasama V. 2018 Highly homogeneous antibody modification through optimisation of the synthesis and conjugation of functionalised dibromopyridinediones. *Org. Biomol. Chem.* **16**, 1359–1366. (doi:10.1039/C7OB03138F)

219. Lee MTW, Maruani A, Richards DA, Baker JR, Caddick S, Chudasama V. 2017 Enabling the controlled assembly of antibody conjugates with a loading of two modules without antibody engineering. *Chem. Sci.* **8**, 2056–2060. (doi:10.1039/C6SC03655D)

220. Lee MTW, Maruani A, Baker JR, Caddick S, Chudasama V. 2015 Next-generation disulfide stapling: reduction and functional re-bridging all in one. *Chem Sci* **7**, 799–802.

221. Maruani A, Szijj PA, Bahou C, Nogueira JCF, Caddick S, Baker JR, Chudasama V. 2020 A plug-and-play approach for the de novo generation of dually functionalized bispecifics. *Bioconjugate Chem.* **31**, 520–529. (doi:10.1021/acs.bioconjchem.0c00002)

222. Martínez-Sáez N et al. 2017 Oxetane grafts installed site-selectively on native disulfides to enhance protein stability and activity in vivo. *Angew. Chem. Int. Ed.* **56**, 14 963–14 967. (doi:10.1002/anie.201708847)

223. Lee B, Sun S, Jiménez-Moreno E, Neves AA, Bernardes GJL. 2018 Site-selective installation of an electrophilic handle on proteins for bioconjugation. *Bioorg. Med. Chem.* **26**, 3060–3064. (doi:10.1016/j.bmc.2018.02.028)

224. Griebenow N, Dilmaç AM, Greven S, Bräse S. 2016 Site-specific conjugation of peptides and proteins via rebridging of disulfide bonds using the thiol–yne coupling reaction. *Bioconjugate Chem.* **27**, 911–917. (doi:10.1021/acs.bioconjchem.5b00682)

225. Koniev O, Dovgan I, Renoux B, Ehkirch A, Eberova J, Cianférani S, Kolodych S, Papot S, Wagner A. 2018 Reduction–rebridging strategy for the preparation of ADPN-based antibody–drug conjugates. *Med. Chem. Commun.* **9**, 827–830. (doi:10.1039/C8MD00141C)

226. Walsh SJ, Omarjee S, Galloway WRJD, Kwan TT-L, Sore HF, Parker JS, Hyvönen M, Carroll JS, Spring DR. 2019 A general approach for the site-selective modification of native proteins, enabling the generation of stable and functional antibody–drug conjugates. *Chem. Sci.* **10**, 694–700. (doi:10.1039/C8SC04645J)

227. Walsh SJ, Iegre J, Seki H, Bargh JD, Sore HF, Parker JS, Carroll JS, Spring DR. 2020 General dual functionalisation of biomacromolecules via a cysteine bridging strategy. *Org. Biomol. Chem.* **18**, 4224–4230. (doi:10.1039/D0OB00907E)

228. Counsell AJ, Walsh SJ, Robertson NS, Sore HF, Spring DR. 2020 Efficient and selective antibody modification with functionalised divinyltriazines. *Org. Biomol. Chem.* **18**, 4739–4743. (doi:10.1039/D0OB01002B)

229. Wang S, Zhou Q, Chen X, Luo R-H, Li Y, Liu X, Yang L-M, Zheng Y-T, Wang P. 2021 Modification of N-terminal α-amine of proteins via biomimetic ortho-quinone-mediated oxidation. *Nat. Commun.* **12**, 1–9. (doi:10.1038/s41467-021-22654-7)

230. Gilmore JM, Scheck RA, Esser-Kahn AP, Joshi NS, Francis MB. 2006 N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed.* **45**, 5307–5311. (doi:10.1002/anie.200600368)

231. Witus LS, Netirojjanakul C, Palla KS, Muehl EM, Weng C-H, Iavarone AT, Francis MB. 2013 Site-specific protein transamination using *N*-methylpyridinium-4-carboxaldehyde. *J. Am. Chem. Soc.* **135**, 17 223–17 229. (doi:10.1021/ja408868a)

232. MacDonald JI, Munch HK, Moore T, Francis MB. 2015 One-step site-specific modification of native proteins with 2-pyridinecarboxaldehydes. *Nat. Chem. Biol.* **11**, 326–331. (doi:10.1038/nchembio.1792)

233. Chen D, Disotuar MM, Xiong X, Wang Y, Chou DH-C. 2017 Selective N-terminal functionalization of native peptides and proteins. *Chem. Sci.* **8**, 2717–2722. (doi:10.1039/C6SC04744K)

234. Deng J-R, Lai NC-H, Kung KK-Y, Yang B, Chung S-F, Leung AS-L, Choi M-C, Leung Y-C, Wong M-K. 2020 Conversion of triphenylphosphine oxide to organophosphorus via selective cleavage of C-P, O-P, and C-H bonds with sodium. *Commun. Chem.* **3**, 1–9. (doi:10.1038/s42004-019-0249-6)

235. Raj M, Wu H, Blosser SL, Vittoria MA, Arora PS. 2015 Aldehyde capture ligation for synthesis of native peptide bonds. *J. Am. Chem. Soc.* **137**, 6932–6940. (doi:10.1021/jacs.5b03538)

236. Adusumalli SR, Rawale DG, Rai V. 2018 Aldehydes can switch the chemoselectivity of electrophiles in protein labeling. *Org. Biomol. Chem.* **16**, 9377–9381. (doi:10.1039/C8OB02897D)

237. Chan AO-Y, Ho C-M, Chong H-C, Leung Y-C, Huang J-S, Wong M-K, Che C-M. 2012 Modification of N-terminal α-amino groups of peptides and proteins using ketenes. *J. Am. Chem. Soc.* **134**, 2589–2598. (doi:10.1021/ja208009r)

238. Tang KC, Raj M. 2021 One-step azolation strategy for site- and chemo-selective labeling of proteins with mass-sensitive probes. *Angew. Chem. Int. Ed.* **60**, 1797–1805. (doi:10.1002/anie.202007608)

239. Obermeyer AC, Jarman JB, Francis MB. 2014 N-terminal modification of proteins with *o*-aminophenols. *J. Am. Chem. Soc.* **136**, 9572–9579. (doi:10.1021/ja500728c)

240. Maza JC *et al.* 2019 Enzymatic modification of N-terminal proline residues using phenol derivatives. *J. Am. Chem. Soc.* **141**, 3885–3892. (doi:10.1021/jacs.8b10845)

241. Schoffelen S, van Eldijk MB, Rooijakkers B, Raijmakers R, Heck AJR, van Hest JCM. 2011 Metal-free and pH-controlled introduction of azides in proteins. *Chem. Sci.* **2**, 701–705. (doi:10.1039/c0sc00562b)

242. Huang R, Li Z, Ren P, Chen W, Kuang Y, Chen J, Zhan Y, Chen H, Jiang B. 2018 *N*-phenyl-*N*-aceto-vinylsulfonamides as efficient and chemoselective handles for N-terminal modification of peptides and proteins. *Eur. J. Org. Chem.* **2018**, 829–836. (doi:10.1002/ejoc.201701715)

243. Singudas R, Adusumalli SR, Joshi PN, Rai V. 2014 A phthalimidation protocol that follows protein defined parameters. *Chem. Commun.* **51**, 473–476. (doi:10.1039/C4CC08503E)

244. Ren H, Xiao F, Zhan K, Kim Y-P, Xie H, Xia Z, Rao J. 2009 A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins. *Angew. Chem. Int. Ed.* **48**, 9658–9662. (doi:10.1002/anie.200903627)

245. Zheng X, Li Z, Gao W, Meng X, Li X, Luk LYP, Zhao Y, Tsai Y-H, Wu C. 2020 Condensation of 2-((Alkylthio)(aryl)methylene)malononitrile with 1,2-aminothiol as a novel bioorthogonal reaction for site-specific protein modification and peptide cyclization. *J. Am. Chem. Soc.* **142**, 5097–5103. (doi:10.1021/jacs.9b11875)

246. Bandyopadhyay A, Cambray S, Gao J. 2016 Fast and selective labeling of N-terminal cysteines at neutral pH via thiazolidino boronate formation. *Chem. Sci.* **7**, 4589–4593. (doi:10.1039/C6SC00172F)

247. Faustino H, Silva MJSA, Veiros LF, Bernardes GJL, Gois PMP. 2016 Iminoboronates are efficient intermediates for selective, rapid and reversible N-terminal cysteine functionalisation. *Chem. Sci.* **7**, 5052–5058. (doi:10.1039/C6SC01520D)

248. Li K, Wang W, Gao J. 2020 Fast and stable N-terminal cysteine modification through thiazolidino boronate mediated acyl transfer. *Angew. Chem. Int. Ed.* **59**, 14 246–14 250. (doi:10.1002/anie.202000837)

249. Purushottam L, Adusumalli SR, Singh U, Unnikrishnan VB, Rawale DG, Gujrati M, Mishra RK, Rai V. 2019 Single-site glycine-specific labeling of proteins. *Nat. Commun.* **10**, 2539. (doi:10.1038/s41467-019-10503-7)

250. Sim YE, Nwajiobi O, Mahesh S, Cohen RD, Reibarkh MY, Raj M. 2019 Secondary amine selective Petasis (SASP) bioconjugation. *Chem. Sci.* **11**, 53–61. (doi:10.1039/C9SC04697F)

251. Hooker JM, Kovacs EW, Francis MB. 2004 Interior surface modification of bacteriophage MS2. *J. Am. Chem. Soc.* **126**, 3718–3719. (doi:10.1021/ja031790q)

252. Tilley SD, Francis MB. 2006 Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J. Am. Chem. Soc.* **128**, 1080–1081. (doi:10.1021/ja057106k)

253. Allan C, Kosar M, Burr CV, Mackay CL, Duncan RR, Hulme AN. 2018 A catch-and-release approach to selective modification of accessible tyrosine residues. *Chembiochem* **19**, 2443–2447. (doi:10.1002/cbic.201800532)

254. Minamihata K, Goto M, Kamiya N. 2011 Site-specific protein cross-linking by peroxidase-catalyzed activation of a tyrosine-containing peptide tag. *Bioconjugate Chem.* **22**, 74–81. (doi:10.1021/bc1003982)

255. Sato S, Matsumura M, Kadonosono T, Abe S, Ueno T, Ueda H, Nakamura H. 2020 Site-selective protein chemical modification of exposed tyrosine residues using tyrosine click reaction. *Bioconjugate Chem.* **31**, 1417–1424. (doi:10.1021/acs.bioconjchem.0c00120)

256. Sato S, Nakamura K, Nakamura H. 2017 Horseradish-peroxidase-catalyzed tyrosine click reaction. *Chembiochem* **18**, 475–478. (doi:10.1002/cbic.201600649)

257. Li BX, Kim DK, Bloom S, Huang RY-C, Qiao JX, Ewing WR, Oblinsky DG, Scholes GD, MacMillan DWC. 2021 Site-selective tyrosine bioconjugation via photoredox catalysis for native-to-bioorthogonal protein transformation. *Nat. Chem.* **13**, 902–908. (doi:10.1038/s41557-021-00733-y)

258. Antos JM, Francis MB. 2004 Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc.* **126**, 10 256–10 257. (doi:10.1021/ja047272c)

259. Antos JM, McFarland JM, Iavarone AT, Francis MB. 2009 Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. *J. Am. Chem. Soc.* **131**, 6301–6308. (doi:10.1021/ja900094h)

260. Seki Y, Ishiyama T, Sasaki D, Abe J, Sohma Y, Oisaki K, Kanai M. 2016 Transition metal-free tryptophan-selective bioconjugation of proteins. *J. Am. Chem. Soc.* **138**, 10 798–10 801. (doi:10.1021/jacs.6b06692)

261. Imiołek M, Karunanithy G, Ng W-L, Baldwin AJ, Gouverneur V, Davis BG. 2018 Selective radical trifluoromethylation of native residues in proteins. *J. Am. Chem. Soc.* **140**, 1568–1571. (doi:10.1021/jacs.7b10230)

262. Tower SJ, Hetcher WJ, Myers TE, Kuehl NJ, Taylor MT. 2020 Selective modification of tryptophan residues in peptides and proteins using a biomimetic electron transfer process. *J. Am. Chem. Soc.* **142**, 9112–9118. (doi:10.1021/jacs.0c03039)

263. Nakane K, Sato S, Niwa T, Tsushima M, Tomoshige S, Taguchi H, Ishikawa M, Nakamura H. 2021 Proximity histidine labeling by umpolung strategy using singlet oxygen. *J. Am. Chem. Soc.* **143**, 7726–7731. (doi:10.1021/jacs.1c01626)

264. Joshi PN, Rai V. 2019 Single-site labeling of histidine in proteins, on-demand reversibility, and traceless metal-free protein purification. *Chem. Commun.* **55**, 1100–1103. (doi:10.1039/C8CC08733D)

265. Adusumalli SR, Rawale DG, Singh U, Tripathi P, Paul R, Kalra N, Mishra RK, Shukla S, Rai V. 2018 Single-site labeling of native proteins enabled by a chemoselective and site-selective chemical technology. *J. Am. Chem. Soc.* **140**, 15 114–15 123. (doi:10.1021/jacs.8b10490)

266. Lin S *et al.* 2017 Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597–602. (doi:10.1126/science.aal3316)

267. Taylor MT, Nelson JE, Suero MG, Gaunt MJ. 2018 A protein functionalization platform based on selective reactions at methionine residues. *Nature* **562**, 563–568. (doi:10.1038/s41586-018-0608-y)

268. Kim J, Li BX, Huang RY-C, Qiao JX, Ewing WR, MacMillan DWC. 2020 Site-selective functionalization of methionine residues via photoredox catalysis. *J. Am. Chem. Soc.* **142**, 21 260–21 266. (doi:10.1021/jacs.0c09926)

269. Vantourout JC *et al.* 2020 Serine-selective bioconjugation. *J. Am. Chem. Soc.* **142**, 17 236–17 242. (doi:10.1021/jacs.0c05595)

270. Bloom S, Liu C, Kölmel DK, Qiao JX, Zhang Y, Poss MA, Ewing WR, MacMillan DWC. 2018 Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. *Nat. Chem.* **10**, 205–211. (doi:10.1038/nchem.2888)