

## Protein Functionalization

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 13757–13777

International Edition: doi.org/10.1002/anie.202012034

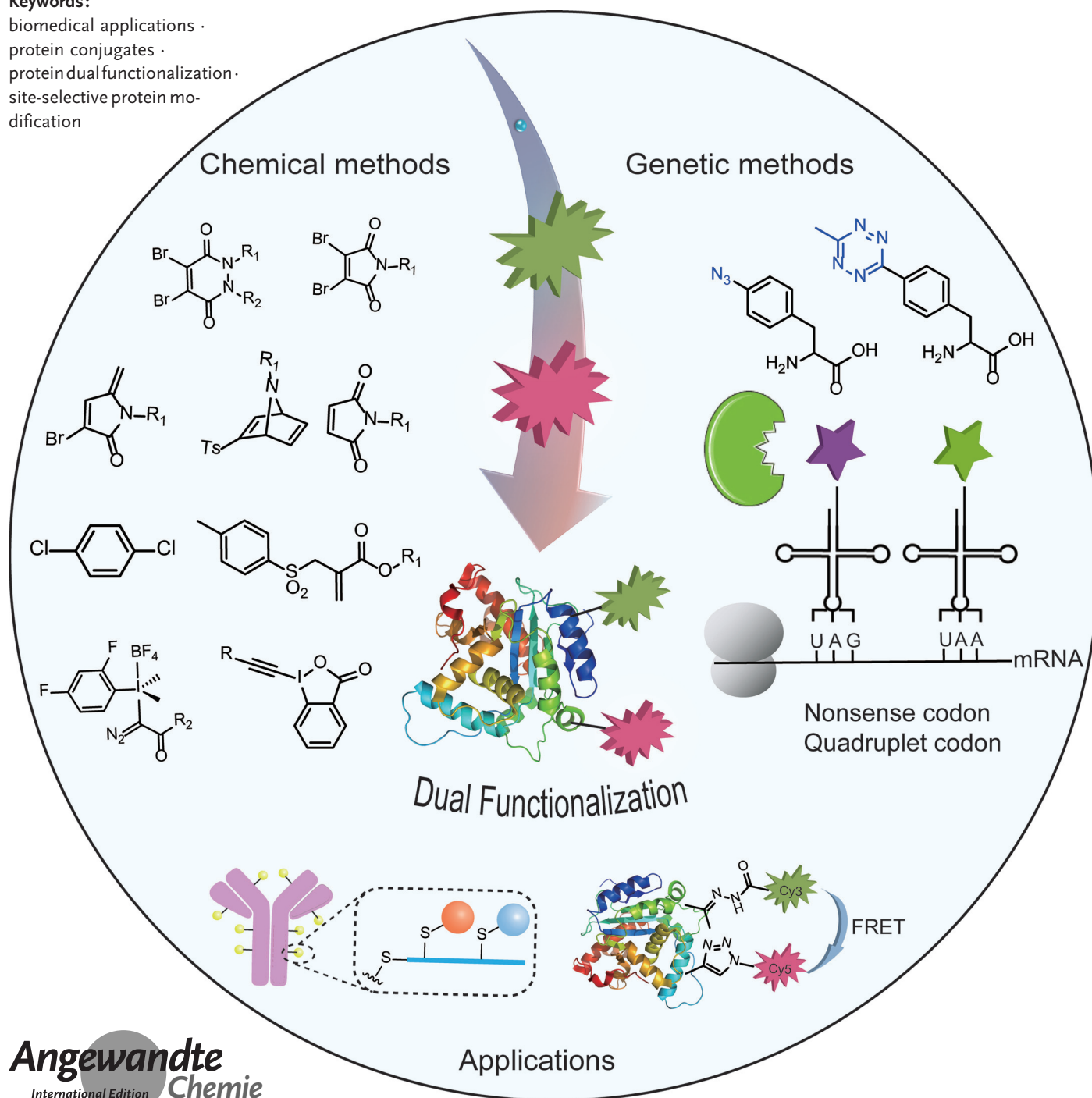
German Edition: doi.org/10.1002/ange.202012034

## Contemporary Approaches for Site-Selective Dual Functionalization of Proteins

Lujuan Xu, Seah Ling Kuan,\* and Tanja Weil\*

## Keywords:

biomedical applications ·  
protein conjugates ·  
protein dual functionalization ·  
site-selective protein modification



**S**ite-selective protein functionalization serves as an invaluable tool for investigating protein structures and functions in complicated cellular environments and accomplishing semi-synthetic protein conjugates such as traceable therapeutics with improved features. Dual functionalization of proteins allows the incorporation of two different types of functionalities at distinct location(s), which greatly expands the features of native proteins. The attachment and crosstalk of a fluorescence donor and an acceptor dye provides fundamental insights into the folding and structural changes of proteins upon ligand binding in their native cellular environments. Moreover, the combination of drug molecules with different modes of action, imaging agents or stabilizing polymers provides new avenues to design precision protein therapeutics in a reproducible and well-characterizable fashion. This review aims to give a timely overview of the recent advancements and a future perspective of this relatively new research area. First, the chemical toolbox for dual functionalization of proteins is discussed and compared. The strengths and limitations of each strategy are summarized in order to enable readers to select the most appropriate method for their envisaged applications. Thereafter, representative applications of these dual-modified protein bioconjugates benefiting from the synergistic/additive properties of the two synthetic moieties are highlighted.

## 1. Introduction

Proteins are ubiquitous in Nature, serving as the basic building blocks of life. They play many essential roles in a myriad of biological processes, such as molecular transport, energy conversion, inter- and intramolecular signaling.<sup>[1,2]</sup> Nature expands protein diversity by post-translational modifications (PTMs) after their biosynthesis in the ribosome, thus vastly enlarging their structural and functional repertoire by up to two orders of magnitude.<sup>[3,4]</sup> Inspired by Nature's elegance, scientists strived to modify proteins with diverse synthetic moieties, allowing for the creation of bioconjugates with high degree of structural perfection and new functional characteristics.<sup>[5]</sup>

In this context, the past decades revealed significant progress in the development of new methodologies for site-selective protein functionalization to install the desired functionalities at pre-defined sites.<sup>[3,6-10]</sup> These well-defined protein conjugates offer great prospects for a wide range of fields including biomedicine, bioimaging, biosensing and materials science.<sup>[5,11-17]</sup> Some promising examples include the conjugation of synthetic polymers to therapeutic proteins to improve their solubility and extend their plasma circulation half-life<sup>[13]</sup> or the attachment of anticancer drugs to antibodies forming antibody-drug conjugates (ADCs) for cell-targeted cancer therapy.<sup>[11,12]</sup>

Nevertheless, along with the increasing demand for multifunctional bioconjugates to perform more sophisticated biological studies in vitro as well as in vivo, appending only one type of functionality to proteins is often insufficient to

## From the Contents


1. Introduction	13758
2. Chemical toolbox for site-selective functionalization of proteins	13760
3. Synthetic strategies for dual functionalization of proteins	13762
4. Genetic engineering for dual functionalization of proteins	13764
5. Higher-level functionalization of proteins	13769
6. Applications	13769
7. Conclusion and Outlook	13774


customize proteins for the desired applications. For example, despite the clinical success of ADCs, classical ADCs equipped with a single type of anticancer drug could still suffer from low efficacy, drug resistance, unfavor-

able pharmacokinetics, immunogenicity, and the inherent hydrophobicity of the drug, which greatly hampers their further in vivo applications.<sup>[18,19]</sup> In addition, there is also an increasing demand for new therapeutic strategies which combine imaging agents and drug molecules within the protein, for example, antibody, for real-time monitoring during the treatment.<sup>[20]</sup> In another example, a protein was functionalized with three copies of cell-targeting somatostatin peptide and an enzyme. Remarkably, the resultant bioconjugate inhibited tumor growth already at 100-fold lower concentration than a clinically approved antibody acting via a similar mode of action. Furthermore, co-administration of this protein bioconjugate with an approved anticancer drug, doxorubicin, boosted its antitumor activity in a combination therapy approach.<sup>[21]</sup> However, these multifunctional protein

[\*] L. Xu, Dr. S. L. Kuan, Prof. Dr. T. Weil  
Max Planck Institute for Polymer Research  
Ackermannweg 10, 55128 Mainz (Germany)  
and

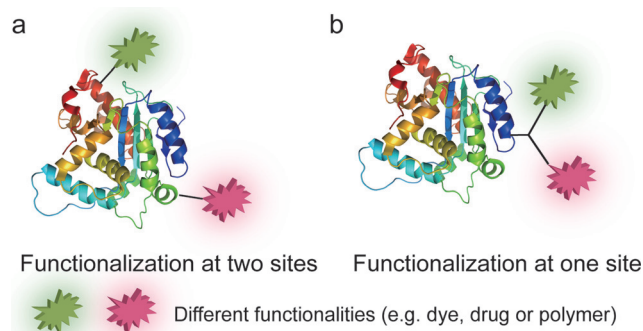
Institute of Inorganic Chemistry I, Ulm University  
Albert-Einstein-Allee 11, 89081 Ulm (Germany)  
E-mail: kuan@mpip-mainz.mpg.de  
weil@mpip-mainz.mpg.de

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/anie.202012034>.

 © 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

conjugates have been mainly achieved by statistical modification on the protein surface,<sup>[22,23]</sup> which results in heterogeneous mixtures with reduced protein activity as well as batch-to-batch variations. In view of biosafety, these limitations significantly hamper their further developments. Consequently, there is a pressing need to devise new strategies to generate protein bioconjugates that exhibit higher order of structural and functional complexity but retaining structural perfection. In this regard, site-selective dual functionalization of proteins has emerged as a promising strategy to customize proteins for the respective applications.

In this review, dual functionalization is defined as the incorporation of two different functionalities into proteins in a site-selective fashion, which is accomplished either at two different amino acids (AA) sites or at a single AA residue at the protein surface (Figure 1). In both cases, the reagents as well as the sequence of the bioconjugation reactions need to be carefully considered. Dual-modified protein bioconjugates can harness the function and properties imparted by the respective payloads, which complement the capabilities of biomolecules and significantly expand their functional arsenal.<sup>[24]</sup> For example, dual functionalization with two chromophores at distinct sites allows real-time monitoring of changes in protein conformations and dynamics upon ligand binding in native environments or in response to certain stimuli by Förster Resonance Energy Transfer (FRET) measurements.<sup>[25]</sup> This cannot be achieved by mono-functionalization of proteins. Moreover, a new generation of protein therapeutics/diagnostics can be derived by incorporating two different functionalities, as exemplified by integrating a drug and an imaging agent into one platform for simultaneous therapy and diagnostics applications (theranos-



**Figure 1.** Two different approaches for dual functionalization of proteins (a) at two different sites and (b) at one single site with a multi-functional linker.

tics) or inserting two different drugs into an antibody for combination therapy.<sup>[20,26,27]</sup>

Owing to the emerging interest for addressing previously “unanswered” fundamental scientific questions as well as offering practical solutions for biomedical applications, site-selective dual functionalization of proteins has seen rapid developments in the past ten years. In this review, we first summarize the synthetic strategies for protein dual functionalization to provide a timely overview of this burgeoning field, which serves as a guideline for the selection of the most appropriate method for the envisaged applications (Figure 2). In addition, the rationale and principles behind these synthetic strategies as well as the strength and inherent limitations are discussed. In the last section, some representative applications of dual-functionalized protein conjugates



*Lujuan Xu studied chemistry at Zhengzhou University, where she received her Bachelor's (2014) and Master's (2016) degrees. From 2014 to 2016, she conducted her Master's thesis on the synthesis of small organic molecules for optoelectronic devices in the group of Prof. Jian Pei at Peking University. In 2017 she received a CSC scholarship to undertake her PhD with Tanja Weil at the MPIP. Her research focuses on the development of new chemical methodologies for site-selective modification of proteins.*

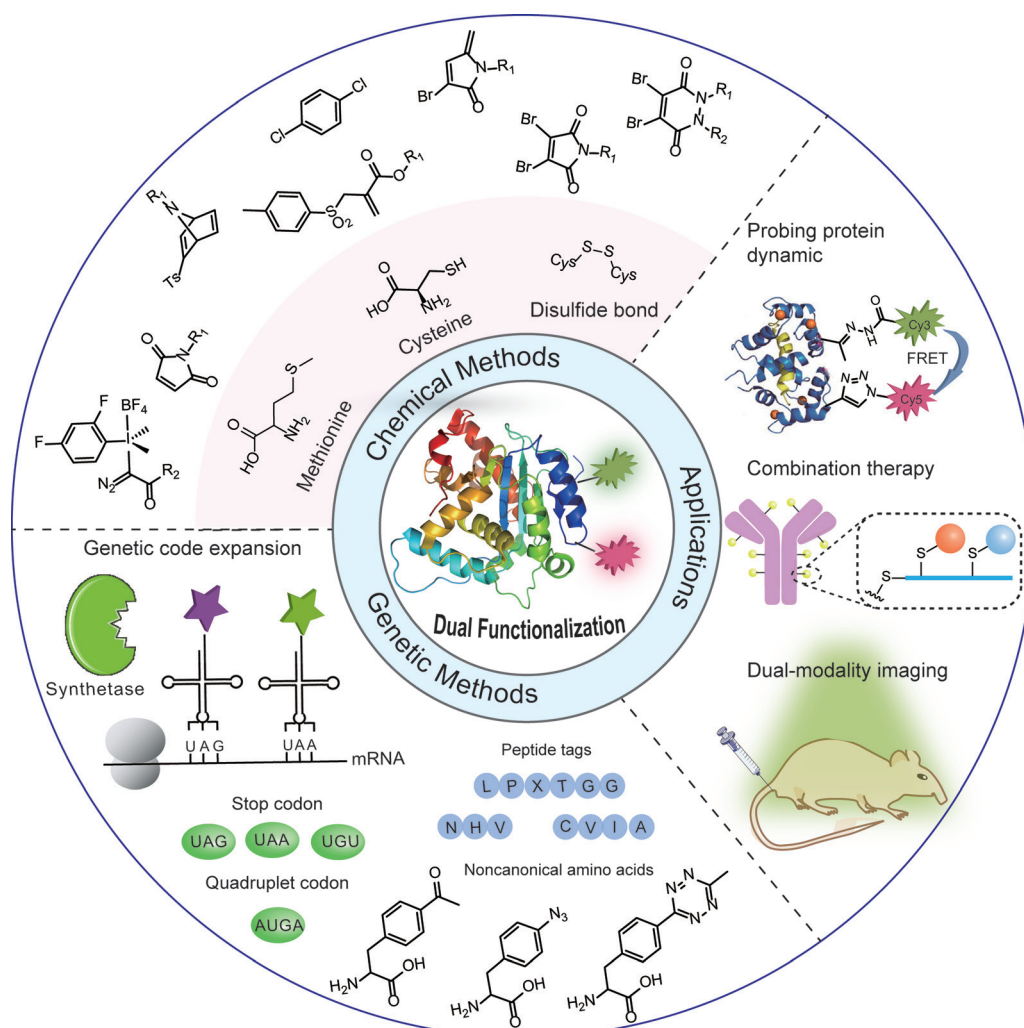


*Seah Ling Kuan received her B.Sc (Hons) in 2003 and completed her PhD (Chemistry) at the NUS in 2009. She joined Prof. Tanja Weil in NUS from 2008–2010 as a research fellow. In 2011, she began her independent research as an Alexander von Humboldt fellow in UUlM. She was a group leader in the Institute of Organic Chemistry III at UUlM from 2013–2016. She joined the MPIP in 2016 as a group leader (Protein Therapeutics). Her research focuses on the development of chemical approaches for the innovation of precision protein therapeutics to address devastating diseases such as cancer, inflammation or infections.*



*Tanja Weil received her PhD from Mainz University in 2002. She held several leading positions at Merz Pharmaceuticals GmbH (Frankfurt, 2002–2008). She was an Associate Professor at the National University of Singapore (NUS, 2008–2010) and since 2010, director at the Institute of Organic Chemistry III (OC III) at Ulm University (Ulm). In 2016, she was appointed as director of the Department of Synthesis of Macromolecules at the Max Planck Institute for Polymer Research (MPIP) in Mainz, Germany. She has received an ERC Synergy*

*Grant. Her scientific interests focus on innovative synthesis concepts to achieve functional macromolecules, hybrid materials and life-like systems to solve current challenges in biomedicine and material science.*



**Figure 2.** Overview of dual functionalization methodologies for proteins and biomedical applications of the resultant protein conjugates (protein dual functionalization at two different amino acid residues is selected as an example).

benefitting from the additive or even synergistic features are highlighted.

## 2. Chemical toolbox for site-selective functionalization of proteins

Site-selective protein functionalization either relies on the functionalization of canonical AAs on the protein surface or genetically encoded noncanonical AA (ncAA) bearing the appropriate bioorthogonal groups.<sup>[28]</sup> Different methodologies have been reported to achieve site-selective mono-functionalization of proteins with high efficiency in a residue specific manner. Bioconjugations have been accomplished at the N-terminus, tyrosine, cysteine or serine residues, to list just a few examples. A brief summary of the commonly used chemical methods for mono-functionalization of natural AA residues is given in Table 1.<sup>[3,6,7,29–32]</sup> For more detailed discussions on protein mono-functionalization, we refer the readers to other excellent reviews on this topic.<sup>[3,6,7,29–31]</sup> Among the 20 AAs, unpaired cysteines have become the

primary choice for functionalization owing to their low abundance at the protein surface as well as the unique nucleophilicity and versatile reactivity profile of the thiol groups.<sup>[7,33–35]</sup>

Besides the unpaired cysteine residues, targeting other side chain residues, such as disulfide bonds and methionine, have also emerged as valuable alternatives to the more commonly used cysteine-based strategies for protein functionalization (Table 1).

All these available methods potentially offer a versatile bioconjugation toolbox to achieve dual functionalization of proteins at two different AA sites. For instance, cysteine and methionine residues are exploited for dual modification with maleimide and hypervalent iodine reagents that proceed in a sequential manner without cross reactivity.<sup>[59]</sup> Despite the simplicity and convenience to achieve dual modification

at two different AA residues, there are still only few reports that address two different AA residues at the protein surface. Furthermore, judicious selection and combination of the two mono-functionalization methods is crucial as the bioconjugation reactions need to be orthogonal and also compatible with each other, to ensure high modification efficiency. Therefore, although other methods involving serine, tyrosine or tryptophan potentially offer a versatile bioconjugation toolbox that is in principle suitable for protein dual functionalization at two different AA residues, these reactions have not been reported yet in this context.

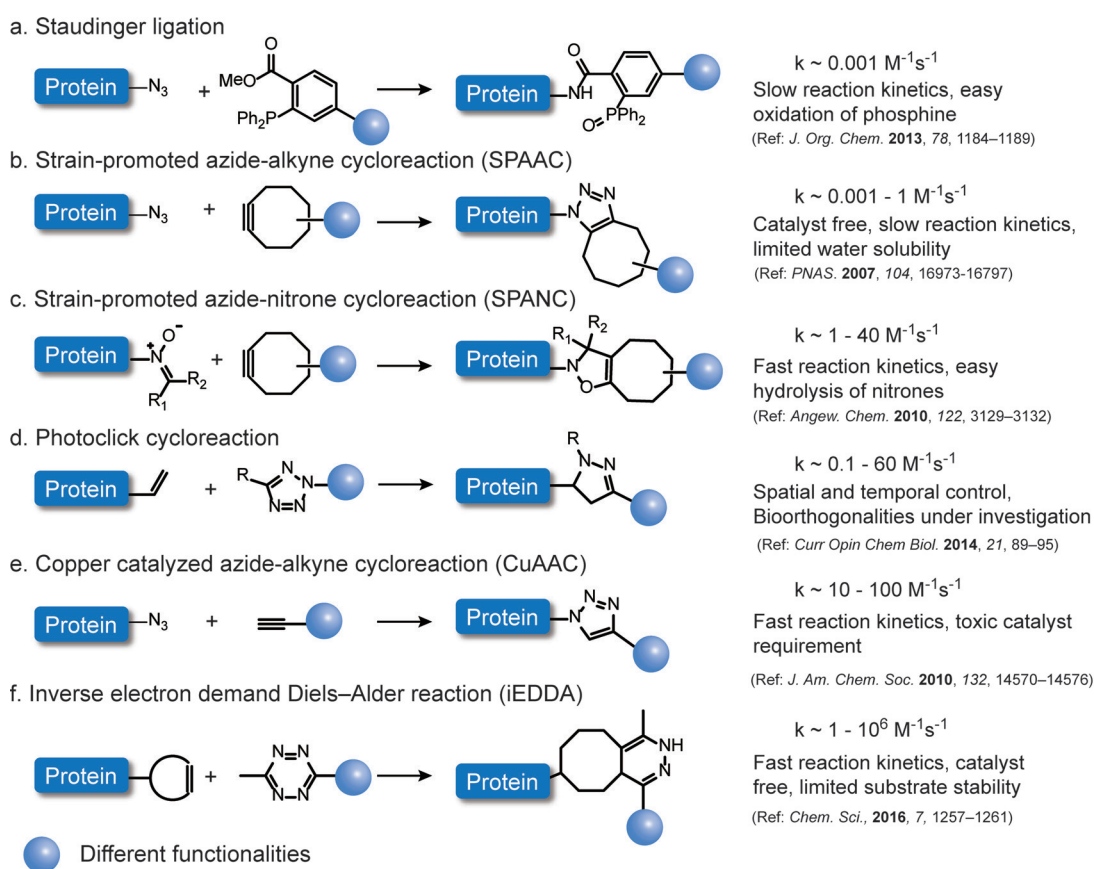
Alternatively, the attachment of a multifunctional linker containing two reactive orthogonal groups to a single site on the protein surface, usually an exposed cysteine or a disulfide bond, also provides straightforward access for the incorporation of two different functionalities. These bioconjugation reagents include maleimide derivatives,<sup>[47]</sup> sulfone derivatives<sup>[37,38]</sup> and other novel reagents,<sup>[44,46]</sup> which are discussed in Section 3.

The recent advances in bioorthogonal chemistry have greatly promoted the progress of site-selective protein

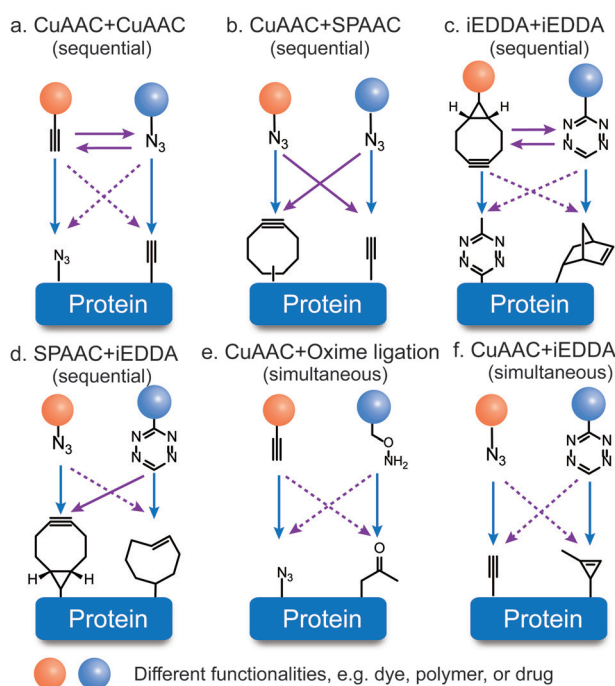
**Table 1:** Brief summary of the currently available chemical toolbox for site-selective protein functionalization.

AA residues	Methods
Cysteine	Halocarbonyl, <sup>[9]</sup> Maleimide derivatives, <sup>[33,36]</sup> Sulfone derivatives, <sup>[37,38]</sup> Metal-mediated arylation <sup>[39–41]</sup> Thiol–ene/yne reactions, <sup>[42,43]</sup> Dichlorotetrazine <sup>[44]</sup> Vinylphosphonothioate, <sup>[45]</sup> Ethynylbenziodoxolones <sup>[46]</sup>
Disulfide	Allyl sulfone, <sup>[38]</sup> Dibromomaleimide, <sup>[47]</sup> Oxetane, <sup>[48]</sup> Divinylpyridine, <sup>[49]</sup> Dibromopyridazinediones <sup>[50]</sup>
N-terminus	2-Formylphenylboronic derivatives, <sup>[51,52]</sup> 2-Pyridinecarboxyaldehyde, <sup>[53]</sup> 2-Ethynylbenzaldehydes, <sup>[54]</sup> o-Aminophenols, <sup>[55]</sup> 2-Cyanobenzothiazole, <sup>[56]</sup> Pictet–Spengler reaction, <sup>[57]</sup> Ketenes <sup>[58]</sup>
Methionine	Hypervalent iodine reagent, <sup>[59]</sup> Oxaziridines, <sup>[60]</sup> Epoxide <sup>[61]</sup>
Tyrosine	Diazonium, <sup>[62]</sup> Mannich-type reaction, <sup>[63]</sup> Pd-mediated alkylation, <sup>[64]</sup> Trizoline-diones <sup>[65]</sup>
Tryptophan	Rhodium-carbenoid, <sup>[66]</sup> Keto-ABNO, <sup>[67]</sup> Metal-mediated arylation/alkynylation <sup>[68–71]</sup>
Serine	Salicylaldehyde ester <sup>[72]</sup> Phosphorus–sulfur incorporation reagents <sup>[73]</sup>
Arginine	Glyoxal reagents <sup>[74]</sup>

modification. Nowadays, various bioorthogonal reactions have been reported with optimized reaction parameters, such as reaction rate, catalyst type, and substrate stability, to impart the desired functionalities for the envisaged applications.<sup>[75,76]</sup> Some commonly used bioorthogonal chemistries as well as selected characteristics are summarized in Figure 3. The copper-catalyzed azide-alkyne cycloaddition (CuAAC) bears the advantage of fast reaction kinetics but suffers from the usage of toxic copper catalyst.<sup>[77]</sup> Strain-promoted azide-alkyne cycloaddition (SPAAC) utilizes the strained alkyne derivatives as reactive partner to avoid the toxic catalyst.<sup>[78]</sup> Nonetheless, the major drawbacks of SPAAC are the limited water solubility of strained alkyne and slow reaction kinetics.<sup>[83]</sup> Strained cyclooctyne can also react with other 1,3-dipoles, for example, nitrones, which is termed as “strain-promoted alkyne-nitrone cycloaddition (SPANC)”. SPANC reactions proceed rapidly with second order rate constants of up to  $39 \text{ M}^{-1} \text{ s}^{-1}$ , which is about 30 times faster than the SPAAC reaction.<sup>[84]</sup> However, the fast reactivity is associated with the instability of the reactive nitrones, which are prone to hydrolysis in aqueous media.<sup>[84]</sup> Photoclick reactions offer the advantages of operational simplicity as well as spatial and temporal control due to the usage of light.<sup>[85]</sup> But recent evidences have shown that photoclick reactions could be limited by potential cross reactivity with, for example, amine residues, and its bioorthogonality still remains controversial.<sup>[86]</sup> Among all the existing bioorthogonal reactions, the

**Figure 3.** Commonly used bioorthogonal reactions for protein modification and summary of some selected features.

inverse electron demand Diels–Alder reaction (iEDDA) stands out because of the fast reaction kinetics (rate constant of up to  $10^6 \text{ M}^{-1} \text{ s}^{-1}$ ), high bioorthogonality, catalyst-free conditions and good biocompatibility.<sup>[87]</sup> Because of its very fast reaction kinetics, it revolutionized bioconjugation of various biomolecules and stimulated labeling in living systems with rate constants comparable to biological reactions.<sup>[87]</sup> The currently available toolbox of diverse bioorthogonal reactions provides the basis for protein dual functionalization ensuring no cross-reactivity and high yields of the bioconjugates. Figure 4 summarizes the different combinations of the bioorthogonal reactions that are commonly employed in the literature. Dual modification proceeds either in a sequential or simultaneous fashion. For example, if two identical CuAAC (or iEDDA) reactions are combined, protein dual modification has to be executed in a sequential manner to prevent cross-reactions, which otherwise would result in a mixture of products. In contrast, the combination of CuAAC and iEDDA introduces two bioorthogonal tags on proteins simultaneously. Since both reactions allow quantitative conversions, dual-modified bioconjugates can be obtained in a one-pot reaction without the need for purification of the single modified product.<sup>[82]</sup>



**Figure 4.** Bioorthogonal chemistry combinations for dual functionalization of proteins (a) CuAAC + CuAAC.<sup>[79]</sup> (b) CuAAC + SPAAC.<sup>[50]</sup> (c) CuAAC + iEDDA.<sup>[25]</sup> (d) SPAAC + iEDDA.<sup>[80]</sup> (e) CuAAC + Oxime ligation.<sup>[81]</sup> (f) CuAAC + iEDDA.<sup>[82]</sup> (Solid line refers to two moieties reacting with each other; dashed line refers to their orthogonal reactivities).

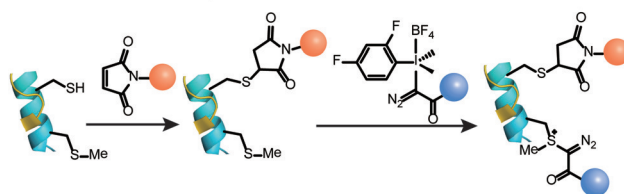
### 3. Synthetic strategies for dual functionalization of proteins

Most synthetic strategies for protein dual modification target canonical AAs exposed at the protein surface. These residues are immediately accessible without the need for tedious genetic engineering of recombinant protein variants and also mitigate the risk of negative effects on proteins folding and function. In this section, dual functionalization at two different sites as well as at a single site is discussed.

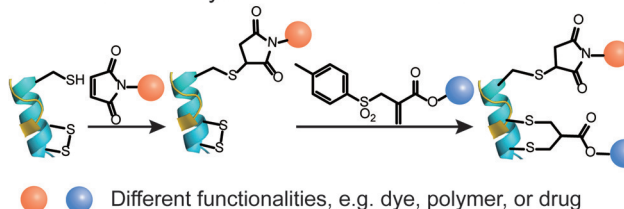
#### 3.1. Dual functionalization at two different sites

Direct modification of two rare AA residues on the protein surface is a straightforward approach to achieve dual modification of proteins. The combination of two different mono-functionalization methods requires stringent selection to ensure orthogonality, compatibility and preferably high modification efficiency. For example, Gaunt and co-workers developed a method based on chemoselective labeling of a single methionine residue with a hypervalent iodine reagent.<sup>[59]</sup> This hypervalent iodine reagent selectively reacted with the moderate nucleophilic methionine residue in the presence of other competitive nucleophilic AA residues, which make it compatible and complementary to other bioconjugation strategies targeting other AA residues. This has been demonstrated by first modifying the unpaired cysteine residue of GTP-binding protein fragment  $G\alpha$  with a maleimide derivative in a Michael reaction to form a thioether bond. Subsequent modification with hypervalent iodine reagent selectively addresses the thioether in a methionine residue yielding the dual modified protein bioconjugate (Figure 5a). Notably, the methionine modification showed high site selectivity without any cross reactivity with the thiol–maleimide conjugation. Besides that, Paavola and co-workers have also achieved dual functionalization by combining cysteine modification and an N-terminal transamination

#### a. Modification at cysteine and methionine sites



#### b. Modification at cysteine and disulfide sites



**Figure 5.** Site-selective protein dual modification at different AA residues. (a) Cysteine and methionine sites.<sup>[59]</sup> (b) Cysteine and disulfide sites.<sup>[89]</sup>

reaction mediated by pyridoxal 5-phosphate.<sup>[88]</sup> The periplasmic glutamine binding protein was site-selectively functionalized by a FRET pair, in which the ligand-induced conformational movements were monitored via changes in FRET efficiency.

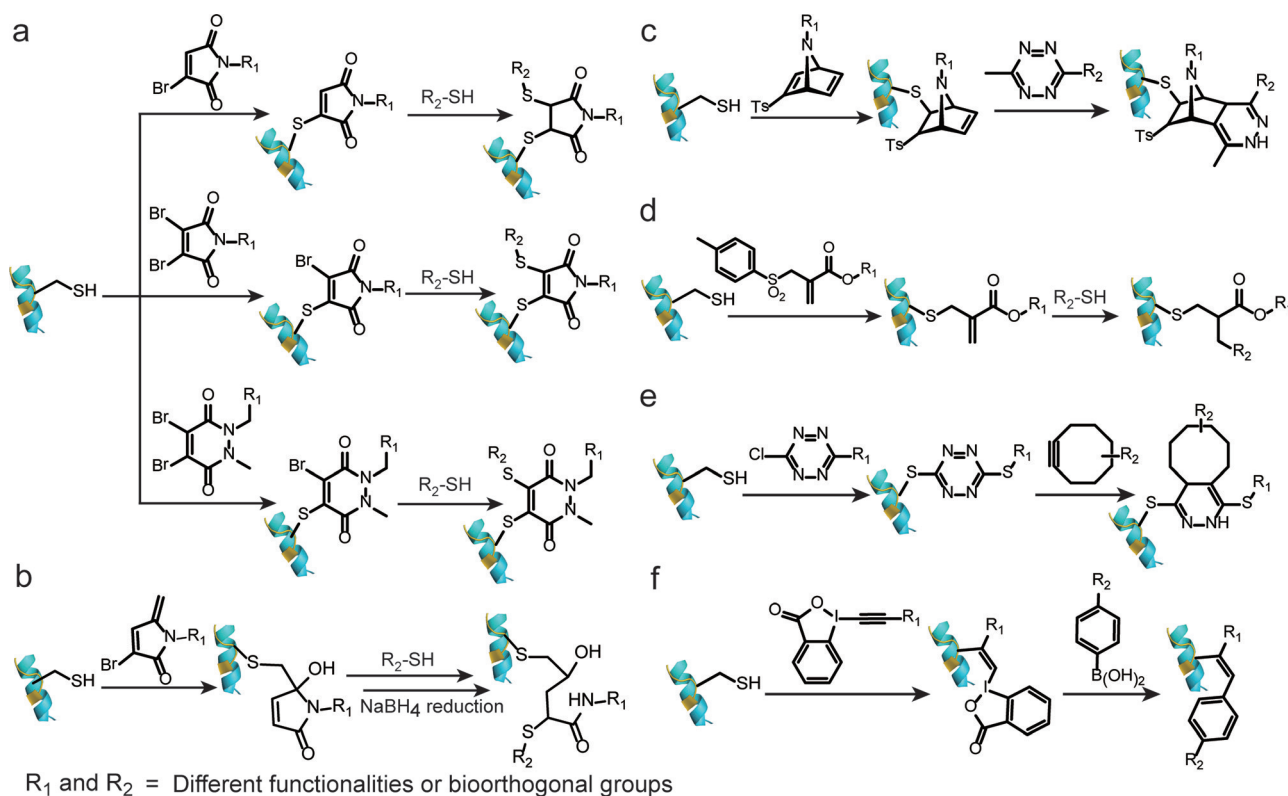
An alternative strategy for site-selective protein dual modification is based on the differences in the reactivity of cysteines in their free (thiol) and oxidized (disulfide) forms.<sup>[89]</sup> Disulfide bonds could be considered as protected thiols, which require activation to form the reduced free thiols for subsequent functionalization. Therefore, site-selective dual modification of native proteins proceeds stepwise by modifying an unpaired cysteine with a maleimide reagent as the initial step, followed by the disulfide reduction to liberate two additional free thiols, which will react with a disulfide rebridging reagent, for example, an allyl sulfone (Figure 5b). It is essential that the thiol–maleimide reaction should be applied first to functionalize the unpaired cysteine residue, followed by the disulfide functionalization using allyl sulfone reagents. Otherwise, the allyl sulfone reagents will also react with the unpaired cysteine, resulting in a heterogeneous product mixture.

### 3.2. Dual functionalization at a single site

Despite the simplicity of dual modification at two distinct sites, the availability of proteins with two different AA

residues with orthogonal reactivities is rather limited. Therefore, alternative methods have been developed to target one specific AA residue with a multifunctional bioconjugation reagent.

Baker, Caddick, and co-workers have demonstrated protein dual modification using mono- and dibromomaleimide reagents.<sup>[47]</sup> The first functionality is introduced by reacting these reagents with an accessible thiol group through an addition-elimination reaction. Subsequently, an additional thiol conjugation introduces the second functionality (Figure 6a). A conceptually similar strategy utilizing dibromopyridazinediones is depicted in Figure 6a.<sup>[90]</sup> Interestingly, the native protein, for example, Grb2 adaptor protein, could be regenerated from the dual-modified conjugate after addition of phosphine or a large excess of thiols, which opens access to the reversible modulation of proteins function or controlled release of the attached cargos, such as drug molecules.<sup>[90]</sup> In a recent example, another maleimide analogue, 3-bromo-5-methylene pyrrolones (3Br-5MPs), was reported for cysteine-specific dual modification of proteins, which has comparable modification efficiency but higher cysteine specificity than the traditional maleimide reagents (Figure 6b).<sup>[91]</sup> The dual modification was achieved by two sequential Michael reactions. First, a Michael reaction of cysteine and 3Br-5MPs generated the bioconjugate that is amenable to a second Michael addition with another thiol, allowing protein dual functionalization at a cysteine site. Due to the slow release of the second functionality, a reducing reagent, for example



**Figure 6.** Site-selective protein dual modification at a single cysteine site to introduce multifunctional bioconjugation reagents. (a) Mono and dibromomaleimide, dibromopyridazinediones (from top to down)<sup>[47,90]</sup> (b) 3-Bromo-5-methylene pyrrolones (3Br-5MPs)<sup>[91]</sup> (c) Azabicyclic vinyl sulfone<sup>[37]</sup> (d) Allyl sulfone<sup>[38]</sup> (e) Dichloro-1,2,4,5-tetrazine<sup>[44]</sup> (f) Ethynylbenziodoxolones (EBXs).<sup>[46]</sup>

$\text{NaBH}_4$ , was required to retard the elimination reaction to generate a stable and bioactive conjugate for subsequent applications.<sup>[91]</sup>

Alternatively, vinyl sulfones are also commonly explored Michael acceptors for protein modification due to its high electrophilic properties that enable their reaction with nucleophiles on the protein surface.<sup>[92]</sup> Nevertheless, their application for dual functionalization was hampered by the cross-reactivity with e.g., amino or imidazole groups generating heterogeneous products.<sup>[93]</sup> Recently, Bernardes and co-workers combined the strained [2.2.1]bicyclic systems with the vinyl sulfone systems and developed the azabicyclic vinyl sulfone reagents for dual functionalization (Figure 6c).<sup>[37]</sup> Such combination results in a fast chemoselective protein modification at the cysteine site, while the dienophile in the azabicyclic strained moiety concomitantly offers an opportunity for further bioorthogonal modification via iEDDA to liberate the energy stored in the strain systems. The second functionalization could even proceed inside living cells for selective apoptosis imaging. Besides vinyl sulfone, allyl sulfone reagents with enhanced water solubility and higher reactivity have also been proposed as a viable strategy for dual modification in a stepwise fashion.<sup>[38]</sup> By simply adjusting the pH, allyl sulfone reagents first reacted in a Michael reaction at pH 6 to attach the first functionality, yielding a conjugated ester system that reacted with the second thiol-containing moiety at pH 8 to achieve dual functionalization of proteins (Figure 6d).<sup>[38]</sup>

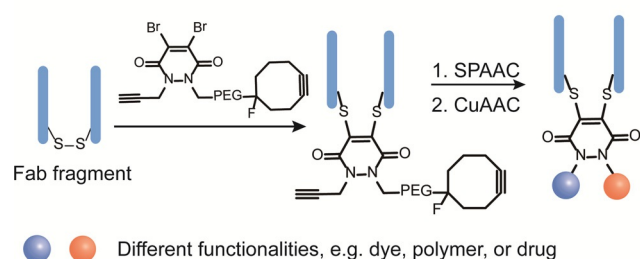
In addition to maleimide analogues and sulfone derivatives, other strategies have also been developed for dual-modification. For example, Goncalves and co-workers reported that dichloro-1,2,4,5-tetrazine can undergo two successive nucleophilic aromatic substitutions to introduce the thiol-containing payload at a cysteine residue with excellent selectivity (Figure 6e).<sup>[44]</sup> The tetrazine linkage could serve as the second handle for subsequent bioorthogonal iEDDA reaction, allowing the preparation of site-selective dual-modified protein conjugates. The feasibility of this strategy has been shown by the dual labeling of the human serum albumin with a macrocyclic chelator for nuclear imaging and a fluorescent probe for fluorescence imaging.<sup>[44]</sup> Despite the simplicity of this method, it could suffer from lower yield if bulky and hydrophobic functionalities need to be incorporated. Furthermore, a different strategy utilizing the inherent reactivity of the hypervalent bond was also reported. Waser and co-workers showed the dual modification of proteins with ethynylbenziodoxolones (EBXs) in high efficiency and chemoselectivity by introducing two reactive groups, i.e., an azide and a hypervalent iodine (Figure 6f).<sup>[46]</sup> Dual modification was achieved via a strain-release-driven cycloaddition and Suzuki–Miyaura cross-coupling of the vinyl hypervalent iodine bond with using palladium diacetate complex as catalyst.

Due to the emergence of ADCs for targeted cancer therapy, dual functionalization at disulfide site has also gained growing interest because of the presence of accessible disulfide bonds in antibodies and the antigen-binding fragment (Fab).<sup>[11,12]</sup> Previously, dibromopyridazinediones have been extensively employed as versatile reagents for dual

modification at the single cysteine site.<sup>[90]</sup> Further reports revealed that it can also serve as a disulfide rebridging reagent to introduce two bioorthogonal tags into disulfide-containing proteins, for example, antibodies or antibody fragments. Chudasama et al. exploited the insertion of dibromopyridazinediones bearing two bioorthogonal tags into the disulfide bonds in full antibody and antibody Fab fragments (the antigen-binding fragment) without perturbing the internal disulfide bonds that are vital for activity (Figure 7).<sup>[50]</sup> Such a plug-and-play platform allowed the introduction of two functionalities via two sequential bioorthogonal reactions in a modular and efficient way, paving the way for the next-generation ADCs.

Besides addressing cysteines and disulfide bonds, other modification strategies at less-explored AA residues have also been reported to expand the existing protein functionalization toolkit. For example, the hypervalent iodine reagents developed by Gaunt and co-workers can be combined with maleimide reagents to achieve dual functionalization at cysteine and methionine sites, which is described in Figure 5a. In addition, the hypervalent iodine reagents have also been demonstrated to show multifaceted reactivity.<sup>[59]</sup> The electrophilicity of the diazo sulfonium conjugate enables a photoredox radical cross-coupling reaction with C-4 benzylation Hantzsch ester derivatives to attach the second functionality yielding dual functionalized conjugates with high conversion.<sup>[59]</sup>

Achieving dual modification at one single AA site has less restriction in terms of the choice of reagents compared to the functionalization at two AA sites. One possible limitation of the single site strategy is that the two orthogonal groups can be sterically hindered due to close proximity on a relatively small multifunctional linker. This could prevent bulkier groups such as polymers to be attached onto the protein. In addition, modification besides cysteine residues is relatively unexplored, thus this field would greatly benefit from further investigations of modification strategies at other low abundant AA residues such as tyrosine, serine or the N-terminus.



**Figure 7.** Site-selective dual modification of proteins at the disulfide site of antibody Fab fragment.<sup>[50]</sup>

#### 4. Genetic engineering for dual functionalization of proteins

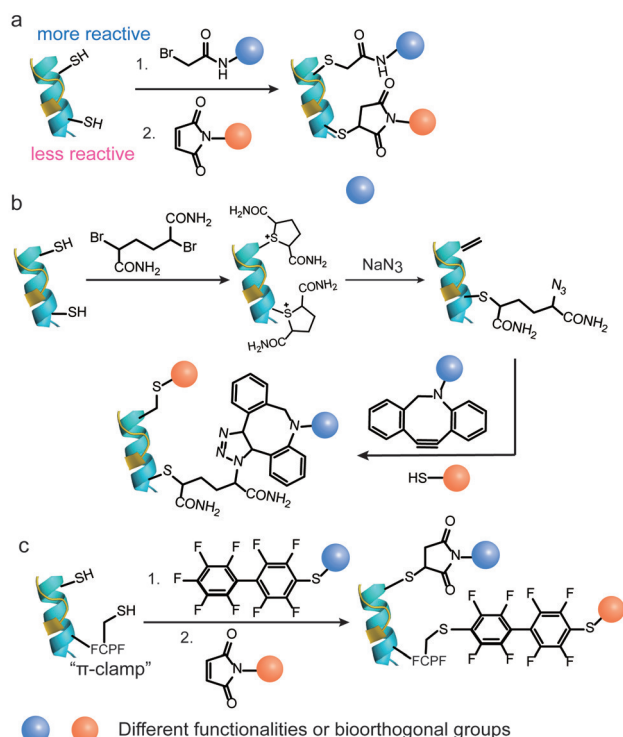
To expand the reactivity beyond what is offered in native proteins, reactive groups for functionalization through genetic incorporation of new AAs (canonical or noncanonical ones)



has emerged as an indispensable tool for site-selective protein dual functionalization. In this section, genetic methods to integrate new reactive canonical AAs, ncAAs or peptide tags for protein dual functionalization are summarized.

#### 4.1. Incorporation of canonical AAs

The expression of recombinant proteins containing one or more point mutations is a straightforward and well-established technology.<sup>[84]</sup> Early attempts to achieve dual functionalization of proteins via genetic methods focused on the preparation of dual cysteine mutants of the target protein. Two cysteine mutations were incorporated at two distinct locations, and both thiols exhibited different reactivity toward different thiol-reactive reagents.<sup>[94]</sup> For instance, Caddick et al. reported the two-cysteine insertions to genetically engineered antibody mimetic proteins, so-called designed ankyrin repeat proteins (DARPs).<sup>[94]</sup> Both thiols are carefully selected and revealed different nucleophilicity, which may originate from their different solvent accessibility. This allowed for the dual functionalization to be executed in a stepwise fashion with high overall labeling efficiency (Figure 8a).<sup>[94]</sup> The authors proposed that the less reactive reagent (bromoacetamide) was reacted first with the more nucleophilic cysteine and subsequently, the more reactive reagent (maleimide) was applied to the second, less nucleophilic cysteine. They successfully demonstrated that the



**Figure 8.** (a) Genetic encoding of two cysteine mutations with different nucleophilicity.<sup>[94]</sup> (b) Genetic encoding of two cysteine mutations possessing different protein local microenvironment.<sup>[95]</sup> (c) Genetic encoding of a cysteine mutation and a "π-clamp" FCPF peptide sequence creates a new microenvironment for the cysteine residue.<sup>[96]</sup>

differences in thiol nucleophilicity afforded a homogeneous product with quantitative conversion at each reaction step, and no purification was needed. Although this concept for dual modification of proteins is very elegant, the delicate balance of thiol reactivity and the selection of the most appropriate cysteine mutation site to prevent heterogeneous product formation (single and dual modified products) could be very challenging. Caddick and co-workers also succeeded in protein dual modification by reacting two cysteines with the same thiol-reactive reagent generating two identical sulfoniums.<sup>[95]</sup> Due to the different accessibilities of the  $\alpha$ -protons at the two cysteine mutations, one of the sulfonium groups, which had a good solvent accessible  $\alpha$ -proton on adjacent position, underwent a  $\beta$ -elimination reaction affording dehydroalanine (Figure 8b). In contrast, the other sulfonium group, which has a shielded  $\alpha$ -proton next to it, remained intact because of the different protein local microenvironment. The functionalities were incorporated via two different chemoselective reactions, offering the site-selective dual-modified protein conjugate in high yield.

The examples mentioned above took advantage of the different local microenvironment offered by the native protein. Inspired by Nature's elegance, Pentelute and co-workers have developed a strategy to create a specific local chemical environment for the cysteine residue by a newly developed, fine-tuned four-amino-acid peptide sequence (FCPF), which is termed as "π-clamp" (Figure 8c).<sup>[96]</sup> This π-clamp enables the conjugation exclusively at this cysteine site with perfluoroaromatic reagents with almost quantitative conversion. The reaction proceeds even in the presence of other competing thiols, thus rendering this approach compatible and complementary to other thiol-conjugation strategies.<sup>[96]</sup> Dual functionalization was demonstrated on a model protein substrate bearing a cysteine and a π-clamp mutation, which was functionalized with a perfluoroaryl probe first based on the π-clamp-mediated conjugation and followed by the thiol-maleimide conjugation reaction.<sup>[96]</sup>

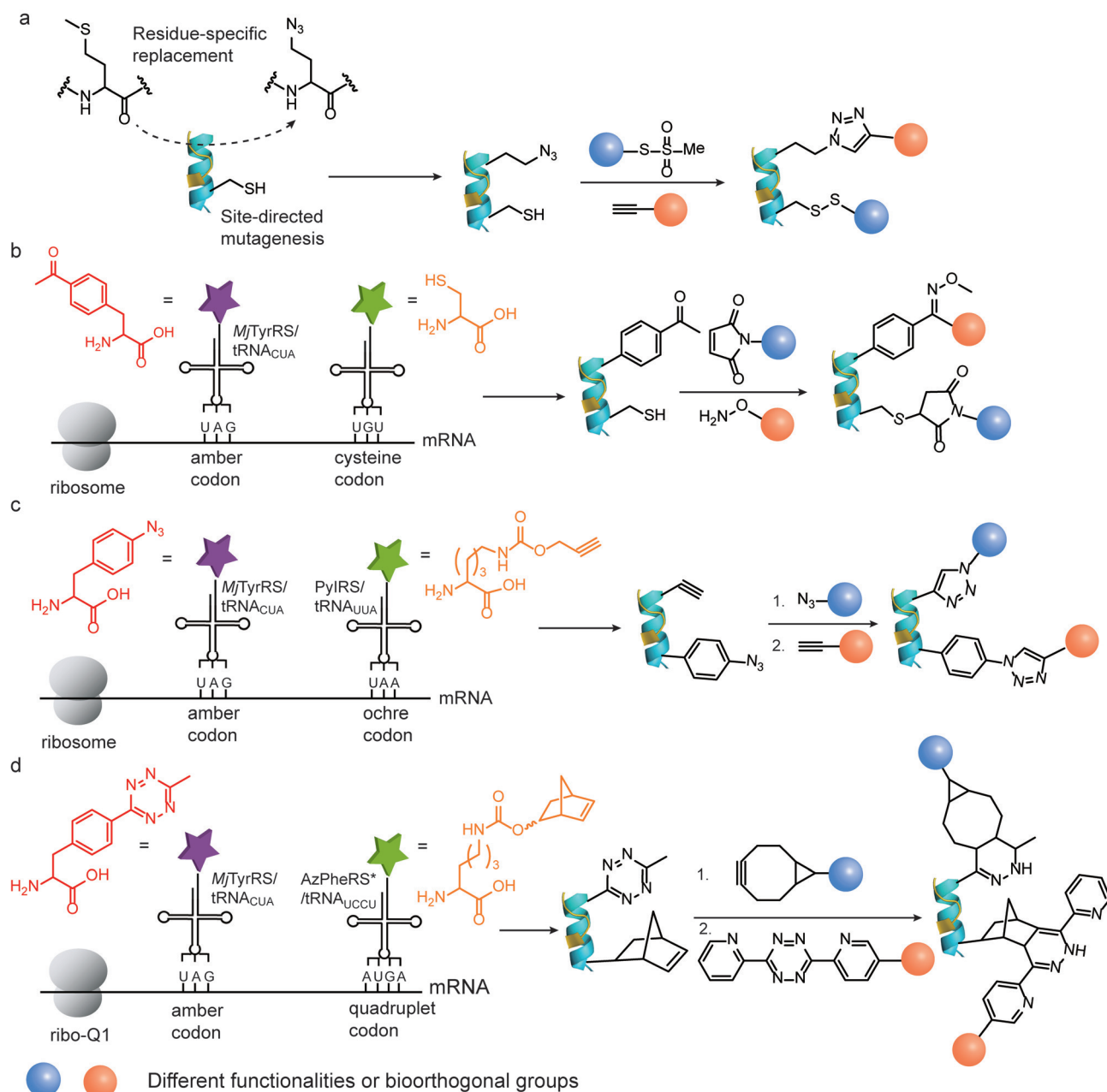
#### 4.2. Incorporation of ncAAs

The introduction of point mutations has certain limitations as the available functionalities and their respective reactivities could only be selected from the pool of the 20 canonical AAs. However, the cellular biosynthetic machinery can be manipulated to incorporate ncAAs that often represent structurally similar derivatives of the canonical AAs. These ncAAs allow protein labeling with unprecedented molecular precision.<sup>[97]</sup> To date, a diverse set of ncAAs with various functionalities or bioorthogonal groups has been reported for genetic incorporation into proteins.<sup>[3,84]</sup> In the auxotrophic strain, organism such as *E. coli* are applied that are not able to synthesize a certain amino acid required for its growth. The ncAAs need to structurally resemble the natural AA to allow binding to the respective endogenous aminoacyl-tRNA synthetase and to replace the natural AA in the polypeptide sequence.<sup>[84]</sup> This strategy has been mainly employed to introduce azide- or alkyne-containing methionine analogues in a methionine-auxotrophic *E. coli* strain.<sup>[84]</sup>

For example, the Davis group incorporated the methionine analogue, azidohomoalanine (AHA), and a cysteine mutation to the target protein based on the combination of site-directed gene mutagenesis and the residue-specific replacement of methionine by its analogues (Figure 9 a).<sup>[98]</sup> Dual modification was accomplished through the initial cysteine conjugation with methanethiosulfonates derivatives, followed by the CuAAC reaction of the azido group in AHA and ethynyl functionalities. This approach offers the benefit of the established AHA incorporation and its efficient modification by cycloaddition reactions. However, the set of available

nCAA is limited as they have to bind to their respective tRNA synthetase efficiently, using auxotrophic strains and all the AA residues within a sequence will be replaced by the respective nCAA analogue.

The genetic code expansion technique has been developed as another technique that allows the insertion of a broad variety of nCAAs with spatial precision at virtually any desired site.<sup>[97]</sup> It is accomplished by using an orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pair, which is capable of charging a designed nCAA in response to a non-sense codon, such as the amber stop codon (UAG), due to



**Figure 9.** (a) Site-selective incorporation of one cysteine mutation via site-directed mutagenesis and one methionine analogue via residue-specific replacement experiment.<sup>[98]</sup> (b) Genetic encoding of one cysteine mutation via site-directed mutagenesis and one nCAA via genetic code expansion in response to UAG codon.<sup>[60]</sup> (c) Genetic encoding of two noncanonical AAs via genetic code expansion in response to UAG and UAA codon.<sup>[100]</sup> (d) Genetic encoding of two noncanonical AAs via genetic code expansion in response to UAG and AUGA codon.<sup>[25]</sup>

their minimal occurrence in most organisms.<sup>[97,99]</sup> This strategy allows the genetic encoding of more than 150 ncAAs containing various reactive handles and functionalities.<sup>[84]</sup> Therefore, the incorporation of ncAA via genetic code expansion in combination with the site-directed canonical AA mutation provides a versatile strategy for protein dual functionalization. Representative work was reported by Deniz and co-workers, in which they utilized an engineered tyrosyl-tRNA synthetase (MjTyrRS)/tRNA<sub>CUA</sub> pair (MjTyrRS/tRNA<sub>CUA</sub> pair) derived from *Methanococcus jannaschi* in response to the amber (UAG) stop codon to insert *p*-acetylphenylalanine, a ketone bearing ncAA, into T4 lysozyme.<sup>[101]</sup> In combination with a single cysteine mutation, dual labeling of the T4 lysozyme mutant was demonstrated by modification with a FRET dye pair through the thiol–maleimide reaction and oxime ligation (Figure 9b).<sup>[101]</sup>

Despite the simplicity of introducing a cysteine mutation as one of the target sites for protein dual modification, this approach could become problematic if the protein consists of native cysteines that play important structural or functional roles. Furthermore, cysteine residues could form disulfide bonds, which could limit expression yields and the reaction reversibility of the thiol–maleimide conjugation could also cause stability problems of the resulting protein bioconjugates.<sup>[36]</sup> Hence, there have been much efforts to introduce two different ncAAs with bioorthogonal tags that could be functionalized independently. Liu and co-workers applied two mutually orthogonal aaRS/tRNA pairs in response to two blank codons.<sup>[100]</sup> They mutated the pyrrolysyl-tRNA synthetase (PylRS)/tRNA<sub>CUA</sub> pair (PylRS/tRNA<sub>CUA</sub> pair) to produce a new variant, PylRS/tRNA<sub>UUA</sub>, which can suppress the ochre (UAA) stop codon. In combination with the evolved MjTyrRS/tRNA<sub>CUA</sub> pair, these two orthogonal aaRS/tRNA pairs were capable of recognizing and inserting two ncAAs, *p*-azido-*L*-phenylalanine and N<sup>ε</sup>-propargyloxycarbonyl-*L*-lysine, into a glutamine-binding protein in response to the amber UAG codon and the ochre UAA codon (Figure 9c).<sup>[100]</sup> Two sequential CuAAC reactions were employed to install two chromophores for the protein dual modification. In a subsequent report, the authors incorporated the azide- and ketone-bearing ncAAs into GFP with the use of an evolved MjTyrRS (AzFRS)/tRNA<sub>CUA</sub> pair and PylRS (AcKRS)/tRNA<sub>UUA</sub> pair, to eliminate protein aggregation and oxidation induced by the copper catalyst.<sup>[81]</sup> Notably, dual functionalization was accomplished by SPAAC and an oxime ligation in one-pot and catalyst-free fashion. In addition, besides in *E. coli*, Schultz and co-workers have also successfully incorporated two different ncAAs, which contained azido and ketone groups, in mammalian cells by utilizing the two orthogonal *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (MbPylRS)/*Methanosarcina mazei* pyrrolysyl tRNA (MbPylRS/MmtRNA<sub>UUA</sub> pair) and tyrosyl-tRNA synthetase (EcTyrRS)/tRNA<sub>CUA</sub> pair (*EcTyrRS*/tRNA<sub>CUA</sub> pair).<sup>[102]</sup> The dual-tagged antibody was subsequently functionalized with a toxic drug payload and a fluorophore with high conversion.

Alternatively, instead of reassignment of the triplet stop codons, Chin and co-workers have exploited the two orthogonal aaRS/tRNA pairs in response to a quadruplet blank codon (four-base codon) and a stop codon for the incorpo-

ration of two ncAAs.<sup>[103]</sup> However, natural ribosomes suffer from very low efficiency in decoding the quadruplet codon. In this context, Chin et al. have synthetically evolved an orthogonal lysozyme (ribo-Q1), which was not responsible for synthesizing the proteome as natural lysozyme, for selectively decoding the quadruplet codon.<sup>[79]</sup> By the combination of ribo-Q1 with two orthogonal aaRS/tRNA pairs, AzPheRS\*/tRNA<sub>UCCU</sub> (a derivative of MjTyrRS/tRNA) and PylRS/tRNA<sub>CUA</sub>, two ncAAs were site-selectively introduced into Calmodulin, forming a triazole intramolecular crosslink through the subsequent CuAAC reaction.<sup>[79]</sup> Nonetheless, the major drawback of this system lies in the low efficiency and specificity of the AzPheRS\*/tRNA<sub>UCCU</sub> pair in directing the corresponding ncAA incorporation. The efficiency of the original system was substantially improved based on the evolution of the PylRS/tRNA<sub>CUA</sub> to obtain an optimized quadruplet decoding variant, PylRS/tRNA<sub>UACU</sub> pair.<sup>[25]</sup> As a proof-of-concept, the evolved PylRS/tRNA<sub>UACU</sub> pair was combined with the AzPheRS\*/tRNA<sub>CUA</sub> pair to site-selectively introduce two bioorthogonal tags, norbornene and tetrazine, into Calmodulin (Figure 9d).<sup>[25]</sup> Dual modification was successfully accomplished via two sequential iEDDA reactions.<sup>[25]</sup> With the established technologies, in the latter example, the authors incorporated the alkyne and cyclopropane handles into Calmodulin by using the orthogonal aaRS-tRNA pair described above, permitting the simultaneous dual functionalization in one pot.<sup>[82]</sup>

Genetic code expansion has witnessed incredible achievements in recent years, which greatly promotes and revolutionizes the field of site-selective protein dual functionalization. The advantages of this strategy lie in the small size of the ncAAs, flexible incorporation sites but high site-selectivity, and the diverse reactive tags available for incorporation. However, some important challenges still remain in this rapidly evolving field, for example, low catalytic efficiency of engineered aaRS which require tedious evaluation and optimization to improve their performance, relative low expression yields, repetitive optimization of the protocols for effective protein engineering, the scope and compatibilities of the functional groups capable to be inserted.<sup>[84]</sup> Besides developing the recombinant engineering techniques, the bioorthogonal chemistry toolbox can also be expanded so that there are more choices for compatible bioorthogonal reaction pairs in terms of reaction rate, catalyst type, and substrate solubility for their intended applications.

### 4.3. Incorporation of peptide tags

Besides the incorporation of ncAAs bearing bioorthogonal tags for dual modification of proteins, the insertion of two artificial peptide sequences that can be recognized by two distinct enzymes for the subsequent covalent labeling with user-defined probes also serves as a straightforward approach for dual modification of proteins. This enzyme-mediated peptide labeling combines the advantage of high specificity of the enzyme towards the peptide tags and excellent labeling efficiency with minimal perturbation to proteins structure and function.<sup>[104]</sup> A brief summary of the commonly used peptide

tags and their respective enzymes is given in Table 2 and relevant examples for dual modification are further discussed in this section. In a recent example, two distinct enzymes, Sortase A and Butelase 1, which demonstrate orthogonal specificity to LPXTGG and HNV motif, respectively, have been combined for dual modification of IgG antibody (Figure 10a).<sup>[118]</sup> Notably, a simple centrifugation process is sufficient to obtain the pure dual-modified protein conjugate, demonstrating the high selectivity and excellent conversion of the enzyme-mediated labeling approach. In addition, peptide tags have also been utilized in combination with nCAAC<sup>[105]</sup> or “ $\pi$ -clamp” sequence<sup>[106]</sup> for site-selective dual-modification of proteins. For example, Chen and co-workers have demonstrated the incorporation of a LAP peptide (LpIA acceptor

peptide), which can be recognized by a lipoic acid ligase (LpIA) to ligate a lipoic acid derivative, and an pyrrolysine analogues which bear an azido group for the site-selective dual labeling of epidermal growth factor receptor (EGFR) on living cells.<sup>[105]</sup>

Despite the simplicity of a one-step introduction of the desired functionalities by two independent reactions, the labeling efficiency may be greatly influenced by the steric effects and hydrophobicity of the introduced probes. This limitation can be overcome by the incorporation of a multi-functional scaffold possessing the bioorthogonal tags. Distefano et al. used farnesyl transferase, which can catalyze the transfer of an isoprenoid group from farnesyl diphosphate to the cysteine site of a tetrapeptide (CVIA, the letter code is the abbreviation of a specific AA), to introduce two reactive tags to the model protein, GFP (Figure 10b).<sup>[107]</sup> The bifunctional alkyne-aldehyde modified protein can undergo two independent bioorthogonal reactions simultaneously, offering the dual-modified conjugate in good conversion.

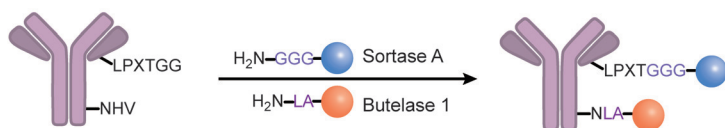
In some instances, the enzyme-mediated dual labeling of proteins can also be achieved on the native protein surface in a site-selective manner without the need to insert artificial peptide tags. For example, in Alabi and coworkers' work, microbial transglutaminase, which catalyzes the formation of interprotein isopeptide bonds between glutamine and lysine residues, can recognize the glutamine 295 on the aglycosylated human IgGs and thus allowed the incorporation of two bioorthogonal tags, azido group and methyltetrazine group (Figure 10c).<sup>[80]</sup> The combination of SPAAC and iEDDA enables the generation of a dual-modified antibody therapeutic containing a cytotoxic payload along with a hydro-

**Table 2:** Commonly used peptide tags and their respective enzymes.

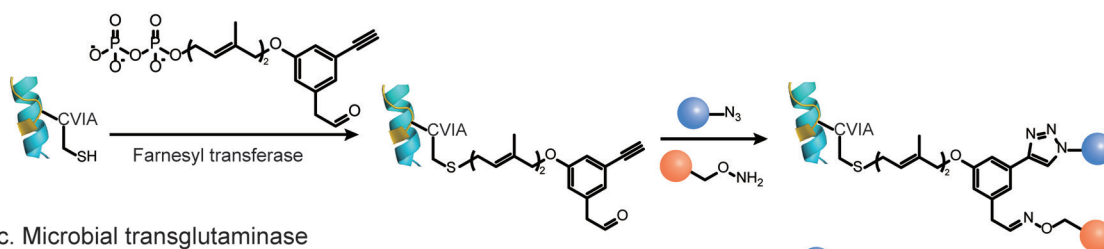
Tag sequence <sup>[a]</sup>	Enzyme
GLNDIFEAQKIEWHE <sup>[108]</sup>	Biotin ligase <sup>[108]</sup>
LPXTGG <sup>[109]</sup>	Sortase A <sup>[109]</sup>
NHV <sup>[110]</sup>	Butelase A <sup>[110]</sup>
CaaX <sup>[b][111]</sup>	Farnesyl transferase <sup>[111]</sup>
LLQG <sup>[112]</sup>	Microbial transglutaminase <sup>[112]</sup>
GFEIDKVWYDLDA <sup>[113]</sup>	Lipoate acid ligase A <sup>[113]</sup>
CXPXR <sup>[114]</sup>	Formylglycine Generating Enzyme <sup>[114]</sup>
GDSLWLLRLLN <sup>[115]</sup>	Phosphopantetheinyl transferase <sup>[115]</sup>
VDSVEGEGEEEGEE <sup>[116]</sup>	Tubulin tyrosine ligase <sup>[116]</sup>
AHIVMVDAYKPTK (Spytag) <sup>[117]</sup>	SpyCatcher <sup>[117]</sup>

[a] The capital letters are the abbreviation of different amino acids. “X” represents any amino acids. [b] “a” represents small aliphatic amino acids and “X” denotes one of Ala, Ser, Met or Glu residues amino acids in CaaX motif.

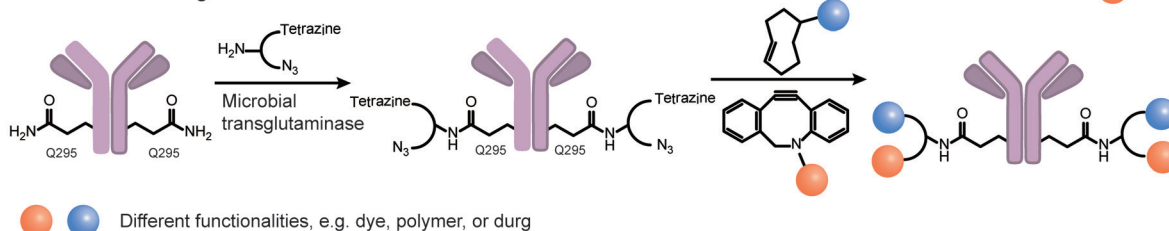
#### a. Combination of Sortase A and Butelase 1



#### b. Farnesyl transferase



#### c. Microbial transglutaminase



● ● Different functionalities, e.g. dye, polymer, or drug

**Figure 10.** Dual modification of proteins by using (a) the combination of Sortase A and Butelase A.<sup>[118]</sup> (b) Farnesyl transferase through the incorporation of a dual-functional scaffold bearing an alkyne and an aldehyde tag.<sup>[107]</sup> (c) Microbial transglutaminase through the incorporation of a dual-functional scaffold bearing a tetrazine and an azido tag.<sup>[80]</sup>

phobicity-masking PEG side chain in a mix-and-match manner. Nevertheless, this strategy is greatly limited by the protein substrate, which requires the recognition motif to be present on the native protein surface at a specific site.

### 5. Higher-level functionalization of proteins

Driven by exciting developments in the discovery of new synthetic procedures as well as scientific curiosity, achieving higher level of protein modification in a site-selective manner is considered crucial for advancing different fields in chemistry, biology and material science. Compared to mono- and dual functionalization of proteins, triple functionalization strategies allow the incorporation of three different functionalities into proteins, thus offering well-defined protein bioconjugates with further expanded structural and functional diversities that is beyond what current synthetic strategies can accomplish. These endeavors will greatly boost our capability to investigate, modulate and re-design the chemical and physical properties of proteins.

However, triple functionalization of proteins is even more demanding than mono- and dual functionalization as the combination of three orthogonal chemistries have to be applied in aqueous solution and in the presence of many reactive proteinogenic groups. Furthermore, by increasing the number of functionalities attached to the protein surface, it may lead to potential detrimental influence to their functional and structural integrity. In addition, the selection of the reactive groups requires more deliberate chemical design as the increase in the hydrophobicity of the bioconjugation reagents may cause protein aggregation. The labeling efficiency may also be compromised due to the steric hindrance from the insertion of three different functionalities.

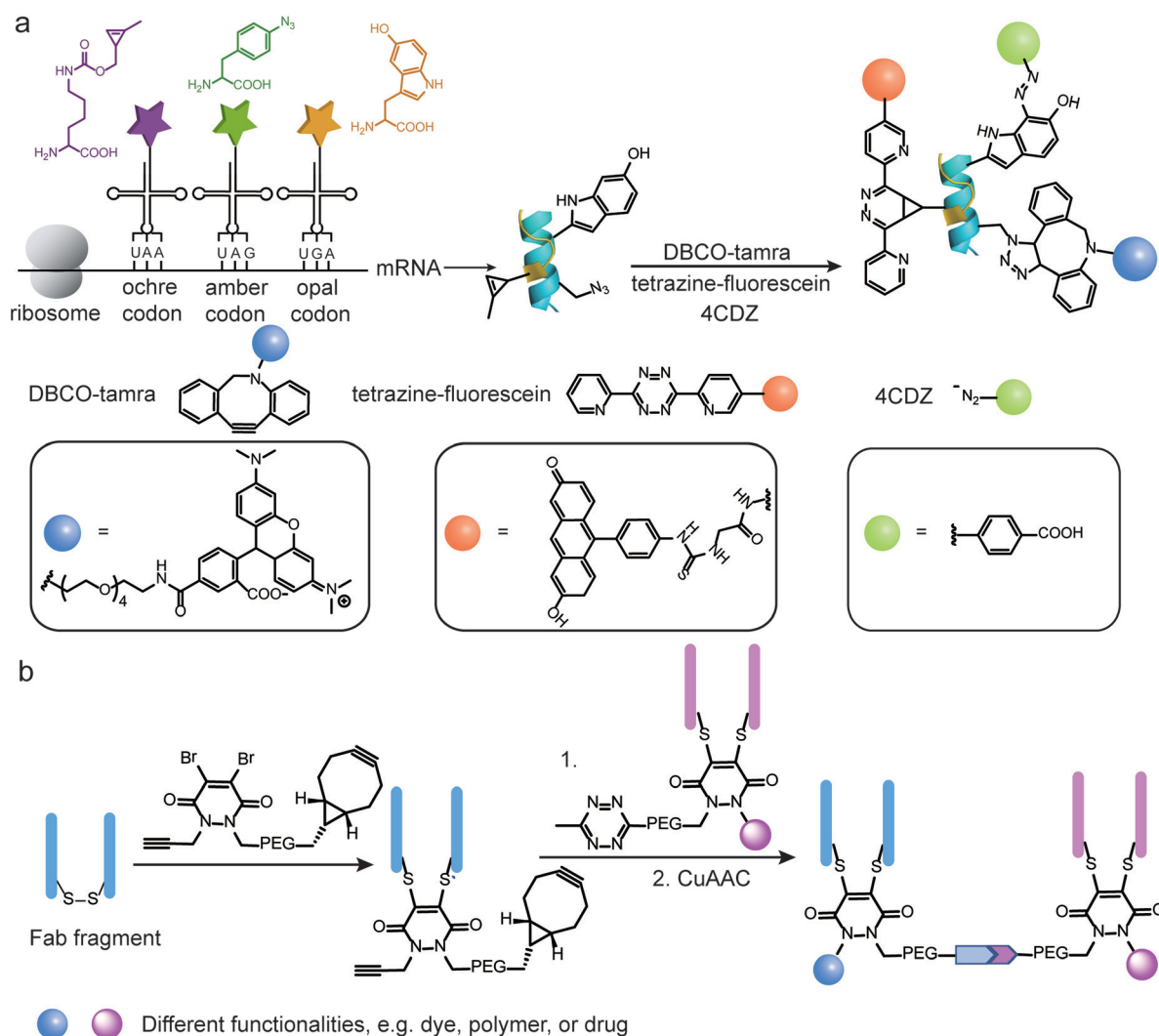
Recent examples have demonstrated the successful triple functionalization of proteins. Chatterjee and co-workers have reported the introduction of three ncAAs into proteins via genetic code expansion. In this work, they assigned the EcTrp, MjTyr, and Pyl pairs to suppress UGA, UAG and UAA codons for the incorporation of three ncAAs, 5-hydroxytryptophan, *p*-azidophenylalanine, and cyclopropene-lysine, in the engineered *E. coli* strain ATMW1 (Figure 11 a).<sup>[119]</sup> The triple functionalization of the protein was successfully achieved by three mutually compatible reactions, SPAAC, iEDDA, and chemoselective rapid azo-coupling reaction (CRACR), in which the electron-rich 5-hydroxyindole ring in 5-hydroxytryptophan reacts with electron-deficient aromatic diazonium ions with fast reaction kinetics and high conversion.<sup>[120]</sup> This is the first example of triple modification of proteins where three ncAAs were incorporated into the target protein in living cells, which further expands the simultaneous ncAAs coding capacity via genetic code expansion. However, this strategy utilized three different aaRS/tRNA pairs and three stop codons, which leaves no codon for termination of endogenous genes and therefore greatly interferes with the translation termination inside cells. Shortly after, Chin and co-workers have demonstrated the genetic encoding of three distinct ncAAs into proteins with three aaRS/tRNA pairs in a different strategy.<sup>[121]</sup> In their work, they identified new

$\Delta$ NPyIRS/<sup>AN</sup>Pyl-tRNA pairs, which lack the N-terminal domains, and revealed that they can be assigned into two classes, class A and B, based on their sequences. Specifically, class A  $\Delta$ NPyIRS preferentially function with class A <sup>AN</sup>Pyl-tRNA and vice versa. Next, they discovered a *Mm*PyIRS/*Spe*<sup>Pyl</sup>-tRNA pair, in which *Spe*<sup>Pyl</sup>-tRNA is orthogonal to both class A and B  $\Delta$ NPyIRS. In this context, the triply orthogonal aaRS/tRNA pairs were identified, which contained the *Mm*PyIRS/*Spe*<sup>Pyl</sup>-tRNA pair, an evolved class A  $\Delta$ NPyIRS/<sup>AN</sup>Pyl-tRNA pair and an evolved class B  $\Delta$ NPyIRS/<sup>AN</sup>Pyl-tRNA pair. Three ncAAs, N $\epsilon$ -((tert-butoxy)-carbonyl)-L-lysine, 3-methyl-L-histidine and N $\epsilon$ -(carbobenzyloxy)-L-lysine, were genetically encoded into proteins in response to the UAG, AGGA and AGUA codons. However, subsequent functionalizations have not demonstrated yet. Compared to the reassignment of three stop codons, this work utilizes the combination of a stop codon and quadruplet codons, which do not interfere with the translation termination and serve as a more general and applicable strategy to achieve higher-level protein functionalization. Even though there is no application shown and there are still obvious limitations, for example, low encoding efficiency, these examples represent the first proof-of-concept for protein triple modifications. They still represent significant technological advancements for the development of multifunctional protein conjugates, which holds immense potential for drug delivery, bioimaging, and material science.

Besides increasing the number of functionalities that can be incorporated into proteins, more complex protein bioconjugates have also been reported. For example, Chudasama et al. described an efficient and modular strategy for the generation of a bispecific antibodies as well as the dual functionalization of the resulting bioconjugate (Figure 11 b).<sup>[122]</sup> By fine-tuning the bioorthogonal chemistry employed, two dual-modified antibodies were first conjugated together via the iEDDA reaction of the two bioorthogonal handles from each antibody (Figure 11 b). Subsequently, two other bioorthogonal handles from each antibody allow for the dual functionalization of the chemically constructed bispecific antibodies. Despite these seminal studies, the preparation of protein conjugates with higher level of modification or structural complexity remains relatively unexplored. We envision that with technical breakthrough in methodology, exciting yet unexplored applications will be discovered, which will undoubtedly revolutionize fundamental studies and applications of proteins.

### 6. Applications

With the emergence of protein dual functionalization techniques, homogeneous protein bioconjugates with remarkable functional complexity have been achieved and studied for diverse applications. In the following, three major fields of applications are highlighted that have greatly benefited from these novel synthesis opportunities.



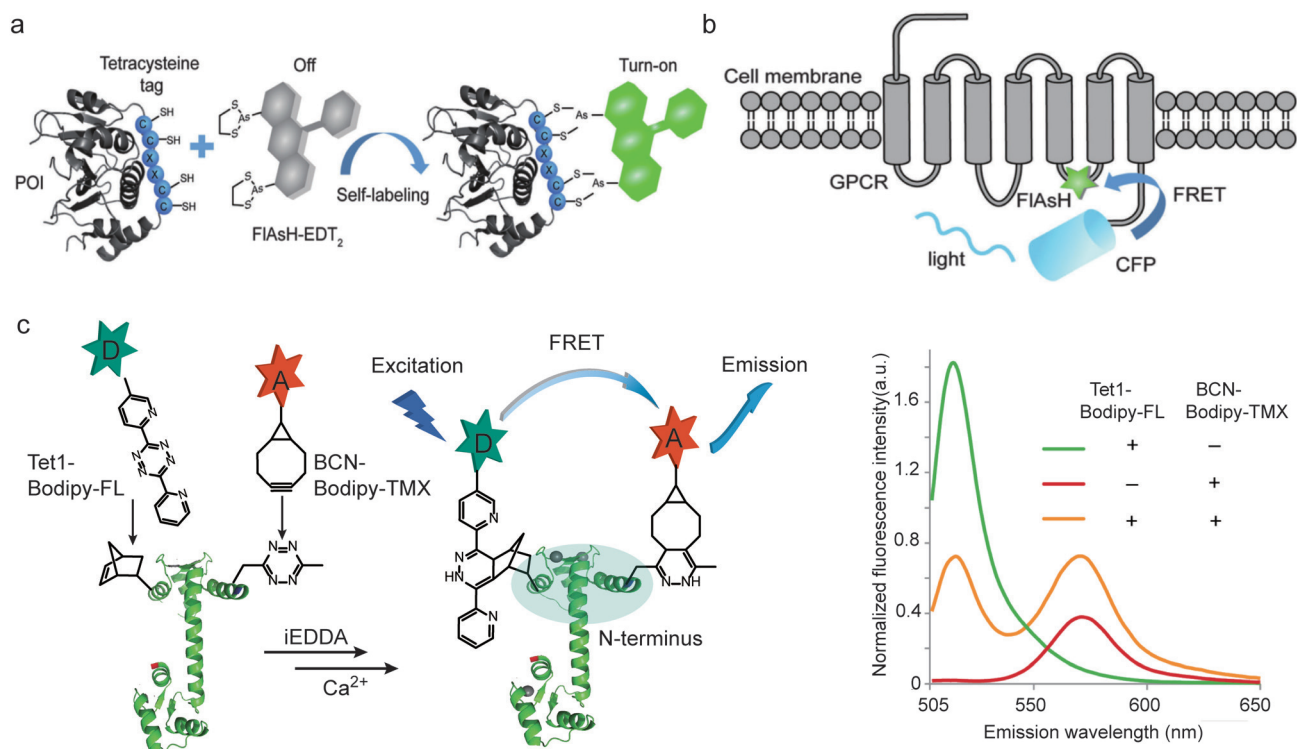
**Figure 11.** (a) Genetic incorporation of three nAAs into protein via genetic code expansion in response to the UAA, UAG and UGA codons for trifunctionalization of proteins.<sup>[119]</sup> (b) Preparation of antibody bispecifics by the conjugation of two dual-modified antibody Fab fragments.<sup>[122]</sup>

### 6.1. Probing protein dynamics through Förster resonance energy transfer

The excitation energy transfer from a donor to an acceptor chromophore, also termed Förster resonance energy transfer (FRET), allows assessing important parameters such as chromophore distances and their orientations in real time and in complex biological environments. Although cryo-electron microscopy is becoming a popular technique to determine structural snapshots of biomacromolecules at atomic resolution, samples need to be frozen and studies cannot be carried out in native environment. In contrast, FRET bears the advantage of easy accessibility and enables the real-time monitoring of the dynamics and conformational changes of proteins in their native environment, thus serving as an important complementary methodology for elucidating proteins structural dynamics.<sup>[123]</sup> A critical requirement for FRET studies is the site-selective incorporation of the respective donor and acceptor chromophores at pre-defined

locations on the protein surface without interfering with the proteins' structure and function.<sup>[123]</sup>

Fluorescent proteins (FPs), for example, cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP), can be genetically encoded into the target proteins and are widely utilized for the investigation of proteins inside living systems.<sup>[15]</sup> For example, agonist-induced activation of G-protein coupled receptor (GPCRs) is thought to cause a conformational rearrangement of their seven transmembrane  $\alpha$ -helices.<sup>[124]</sup> The insertion of CFP and YFP to the G-protein coupled receptors (GPCR) offered the dual tagged protein bioconjugate, which allows real-time monitoring of the activation switch of GPCRs in living cells.<sup>[125]</sup> However, due to the large size of FPs (molecular weight close to 30 kDa), the CFP-YFP FRET pair was found to perturb the proteins structure and function and impair the downstream signal transduction.<sup>[125]</sup> In order to overcome these limitations, the self-labeling peptide tags have been developed, which can bind or react with small-molecule reagents to attach the fluorophore to target proteins. For example, tetracysteine-



**Figure 12.** (a) The self-labeling peptide of tetracysteine reacting with FIAsh forming fluorescent conjugate. Adapted with permission from ref. [15]. Copyright (2015) Royal Society of Chemistry (b) Incorporating the chromophore and CFP as FRET pair to GPCR to probe protein dynamics. Adapted with permission from ref. [15]. Copyright (2015) Royal Society of Chemistry (c) Incorporation of FRET pair into Calmodulin (CaM) via genetic code expansion technique and the corresponding fluorescence spectra after addition of Ca<sup>2+</sup>. Adapted with permission from ref. [25]. Copyright (2014) Springer Nature.

containing motif CCXXCC (C is the abbreviation of cysteine and XX can be virtually any AA sequence), can selectively react with fluorescein arsenical hairpin binder (FIAsh) to form fluorescent product (Figure 12a).<sup>[126]</sup> Lohse and co-workers used the FIAsh-tetracysteine system to label the human adenosine A<sub>2A</sub> receptor in a site-specific fashion, which, in conjunction with a CFP, generated a FRET construct (A<sub>2A</sub>-FIAsh-CFP) that exhibited a 5-fold enhanced agonist-triggered FRET signal compared to the dual-tagged CFP and YFP conjugate (Figure 12b).<sup>[127]</sup>

In comparison to FPs, small molecule chromophores bear the advantage of small size, excellent photostability and high quantum yield.<sup>[128]</sup> So the incorporation of two chromophores into target proteins as FRET pair have represented as a more advantageous strategy to probe protein dynamics without interfering their structure and functions. One representative example is the introduction of two fluorophore into a calcium-binding protein, Calmodulin (CaM), via genetic code expansion to study its conformation changes as a function of the calcium ion (Ca<sup>2+</sup>) concentration. CaM consists of two domains, the C-terminal and N-terminal domain that contain two EF-hand motifs to bind Ca<sup>2+</sup>.<sup>[129]</sup> Upon Ca<sup>2+</sup> binding, the conformation of CaM changes, which affects its binding to different target enzymes.<sup>[129]</sup>

The structures of CaM with four Ca<sup>2+</sup> and without Ca<sup>2+</sup> have been resolved in detail via NMR or X-ray crystallography. However, the transition states that dynamically occur

during the Ca<sup>2+</sup> binding processes and that play an important role in modulating the interaction of CaM with its binding partners could not be resolved with conventional structure analysis by NMR and X-Ray.<sup>[130]</sup> Therefore, dual-tagged CaM was obtained by genetic code expansion and conjugated with the FRET pairs, BODIPY-TMR-X (Bodipy-tetramethylrhodamine, acceptor) and BODIPY-FL (4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionic acid, donor), respectively, to investigate the local conformation changes occurring in the N-terminal domain of CaM in response to the Ca<sup>2+</sup> concentration (Figure 12c).<sup>[25]</sup> After excitation at 485 nm, the dual chromophore tagged CaM reveals two distinct signal corresponding to the emission of the BODIPY-FL donor at 515 nm and the BODIPY-TMR-X acceptor at 570 nm. With increasing Ca<sup>2+</sup> concentrations, the emission intensity of the acceptor dye decreased while the emission of donor dye increased indicating that the FRET pair at the N-terminal domain moved apart as a response of Ca<sup>2+</sup> binding. The current applications utilizing FRET to follow protein dynamics are mainly applying a certain stimulus in test tubes rather than measurements inside cells. In the future, continuous advancements in dual modification of proteins could provide valuable tools to shed light on the dynamics of proteins in living cells or even in organisms. There is much interest and progress in probing protein labeling reactions proceeding in living environments. How-

ever, no dual labeling reactions have been achieved inside cells that would allow FRET studies until now.

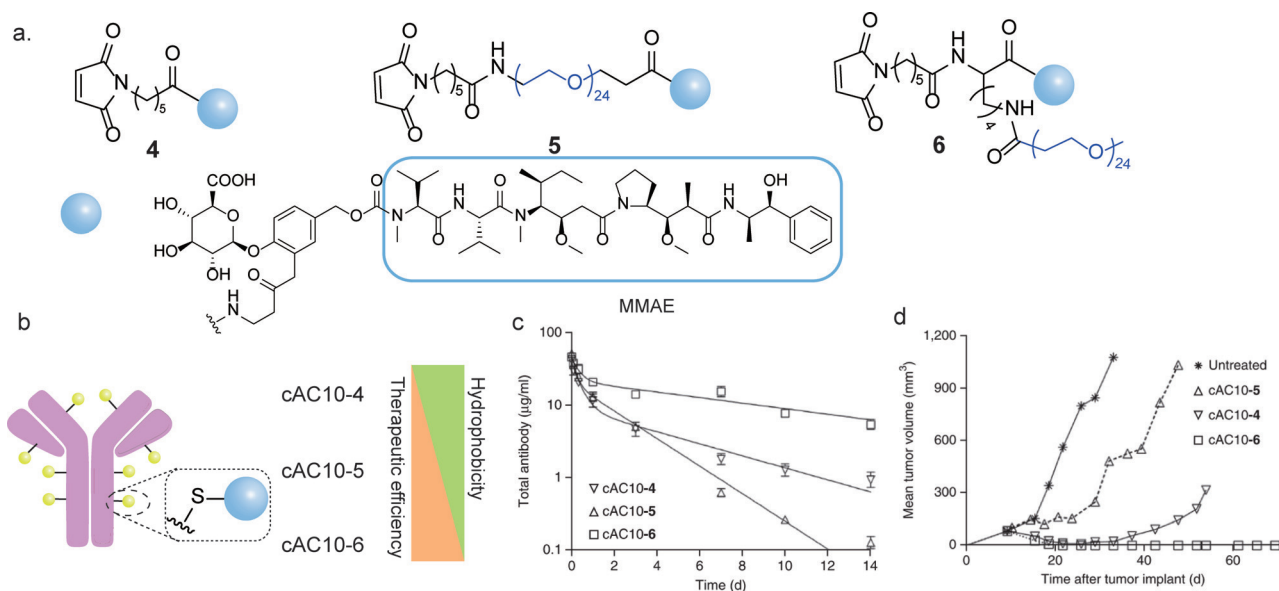
## 6.2. Combination therapy

ADCs have emerged as one of the most powerful and promising strategies for targeted cancer therapy.<sup>[12]</sup> This new class of biopharmaceuticals combines the exquisite target-specificity of antibodies with a highly potent cytotoxic payload via various conjugation technologies and enables the selective delivery of the payload to cancer cells with minimal off-target effects.<sup>[11,12]</sup> Owing to the elegant concept, there are already four ADCs approved by the US Food and Drug Administration on the market and over 60 ADCs are in clinical trials, demonstrating its great prospect for targeted therapeutic applications.<sup>[12]</sup>

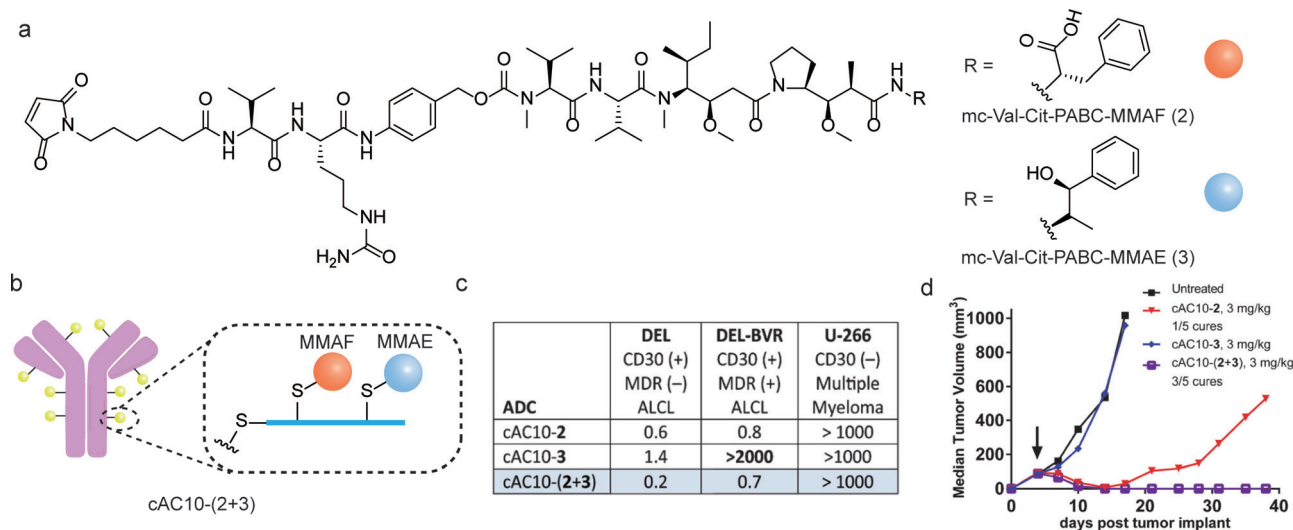
Despite the clinical success, most ADCs are still restricted to the conjugation with a single type of payload, for example a certain anticancer drug, to target a specific type of cancer cells. In fact, this kind of single-functionalized ADCs often suffer from the unfavorable pharmacokinetics and the inherent hydrophobicity of the chemotherapeutic drugs, which could significantly limit the *in vivo* therapeutic efficiency.<sup>[11]</sup> Dual modification of antibodies can serve as an elegant strategy to alleviate these drawbacks, allowing for the incorporation of two complementary modalities to achieve synergistic effects. This can be exemplified by the studies on extending the plasma circulation half-life and attenuating immunogenicity of proteins through the attachment of hydrophilic polyethylene glycol (PEG) chain to proteins. The hydrodynamic radius of the biomolecule also increases after attaching a long PEG chain, thereby mitigating

the renal filtration and circumventing its degradation by proteases and recognition by the immune system, which would otherwise lead to faster plasma clearance.<sup>[131]</sup> In this context, dual modification of antibodies allows the simultaneous introduction of a cytotoxic drug molecule and a PEG chain in a site-selective manner to afford well-defined dual-functional ADCs with improved pharmacokinetics. Senter and co-workers have reported that the accelerated plasma clearance of ADCs originated from their hydrophobicity, which could be reduced by introducing a PEG chain with different configuration.<sup>[132]</sup> As such, three different drug linkers were designed, in which the first drug linker **4** had no PEG chain while both the drug linker **5** and **6** consisted of a unfunctionalized PEG chain (PEG<sub>24</sub>, containing 24 repeating units) but with different configuration (Figure 13a).<sup>[132]</sup> Specially, drug linker **6** incorporated a branched scaffold. The cAC10 antibody (CD30-directed antibody) was conjugated with the three different drug linkers via thiol–maleimide chemistry offering three different ADCs, cAC10-4 (without PEG24), cAC10-5 (with PEG24) and cAC10-6 (with branched PEG24) as depicted in Figure 13b.<sup>[132]</sup> Both the *in vitro* and *in vivo* experiments revealed an inverse correlation between hydrophobicity and tumor volume (Hodgkin's lymphoma), namely that cAC10-6 exhibited slower plasma clearance and much slower tumor growth compared to cAC10-4 and cAC10-5 (Figure 13c,d).<sup>[132]</sup> In particular, the improved pharmacokinetics and therapeutic efficiency of the branched conjugate cAC10-6 compared to the linear conjugate cAC10-5 highlights the importance of the hydrophilic linkers for optimizing the pharmacokinetic parameters.

On the other hand, the therapeutic efficacy ADCs could also suffer from drug resistance after multiple treatments due to inherent and acquired resistance or tumor heterogenic-







**Figure 14.** (a) The chemical structure of two different drug linker (mc-Val-Cit-PABC-MMAF and mc-Val-Cit-PABC-MMAE). (b) ADCs with the incorporation of two different drug molecules: MMAE and MMAF. (c) *In vitro* data of antitumor activity of cAC10-(2 + 3) on MDR(+)-DEL-BVR cell lines. The values are reported as  $IC_{50\text{ng}}/m\text{LoF}$  ADC. ALCL = anaplastic large cell lymphoma. MDR = multidrug resistance. DEL-BVR is a cAC10-vc-MMAE resistant cell line, which is generated after prolonged exposure of the DEL anaplastic large cell lymphoma (ALCL) cell line to cAC10-3. (d) *In vivo* data of antitumor activity of cAC10-(2 + 3) on MDR(+)-DEL-BVR cells lines. (c, d) Adapted with permission from ref. [27] under the Creative Commons CC BY-NC license. Copyright (2017) Wiley-VCH.

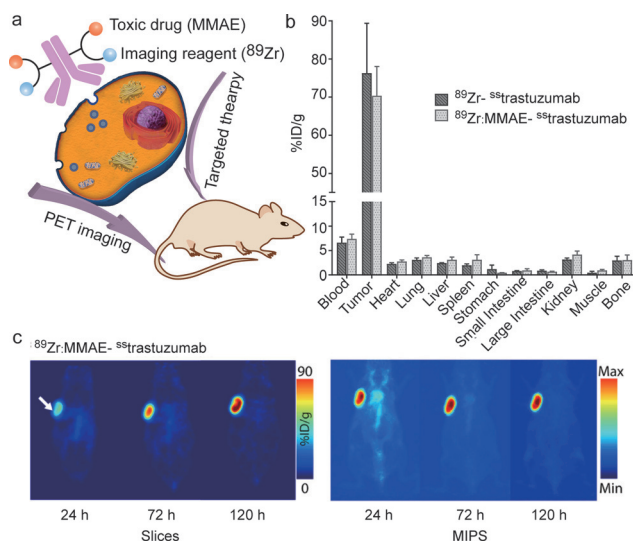
ity.<sup>[12]</sup> Although the exact mechanism of drug resistance is still under much investigation, current available clinical data indicate that malignant cells, which are resistant to a particular drug, could still respond to other drugs.<sup>[19]</sup> Therefore, the attachment of two drugs with different modes of action to an antibody can afford more complex ADCs. This could be co-delivered to cancer cells to overcome drug resistance, which represents an emerging field in targeted cancer therapy.<sup>[18]</sup> For example, Levengood and co-workers have reported a dual-conjugation strategy for the preparation of novel ADCs including two complementary drugs.<sup>[27]</sup> In their work, two different auristatin molecules, monomethyl auristatin E (MMAE) and monomethyl auristatin F (MMAF), were selected to conjugate to the cAC10 antibody due to their complementary physicochemical properties and anticancer activities (Figure 14a,b).<sup>[27]</sup> Specifically, MMAE is cell-permeable and exhibits bystander activity capable of killing neighboring antigen-negative cells.<sup>[133]</sup> However, MMAE is a substrate for multidrug efflux pump exporters, hence showing diminished activity on cells with high pump expression.<sup>[134]</sup> In contrast, MMAF is susceptible to drug export, but not cell-permeable and exhibits negligible bystander effect.<sup>[133]</sup> *In vitro* and *in vivo* experiments demonstrated that the dual-auristatin ADCs were active on cells and tumors that were refractory to treatment with either of the individual component drugs (Figure 14c,d). This work highlights the potential of dual modification of antibody for delivering two synergistic and complementary drug payloads for improved antitumor activities, which represents a notable advancement of the ADCs technology.

Furthermore, dual modification of antibodies also provides great opportunities for simultaneous therapy and diagnostic applications, so-called theranostics, in which the antibody is conjugated to an anticancer drug and a contrast

agent to allow tumor diagnosis and therapy at the same time. In one representative example, Zeglis and co-workers reported the dual labeling of a HER2-targeting trastuzumab with a toxin, MMAE, and a positron-emitting radiometal  $^{89}\text{Zr}$  for theranostic applications (Figure 15a).<sup>[26]</sup> The *in vivo* experiment indicated that the resulting  $^{89}\text{Zr}$ -trastuzumab-MMAE bioconjugate demonstrated excellent tumor targeting and therapeutic efficacy (Figure 15b). Importantly, the dual labeled ADCs represents a targeted drug delivery system that could be tracked *in vivo* using PET providing information of the *in vivo* biodistribution and real-time drug doses during the treatment (Figure 15c).

### 6.3. Dual-modality imaging

Molecular imaging is a powerful and invaluable tool for noninvasive visualization of physiological processes that occur in living organisms at cellular levels.<sup>[5]</sup> Until now, various modern imaging technologies, including optical imaging (OI), magnetic resonance imaging (MRI), positron emission tomography (PET) and computed tomography (CT), have been developed and widely used to monitor the structural, functional and dynamic changes in cancer tissues.<sup>[135]</sup> Each imaging modality has its own unique strength and intrinsic limitations. Consequently, combining two modal imaging methods has received increasing attention, as it allows the collection of complementary imaging data, thereby improving the reliability and accuracy of the diagnosis. For example, PET can provide real-time images of tumor lesions as well as monitor their whole-body distribution and migration, while optical imaging can provide high-resolution imaging to support surgeons in identifying tumor margins during surgical resection.<sup>[136,137]</sup> By combining both modalities

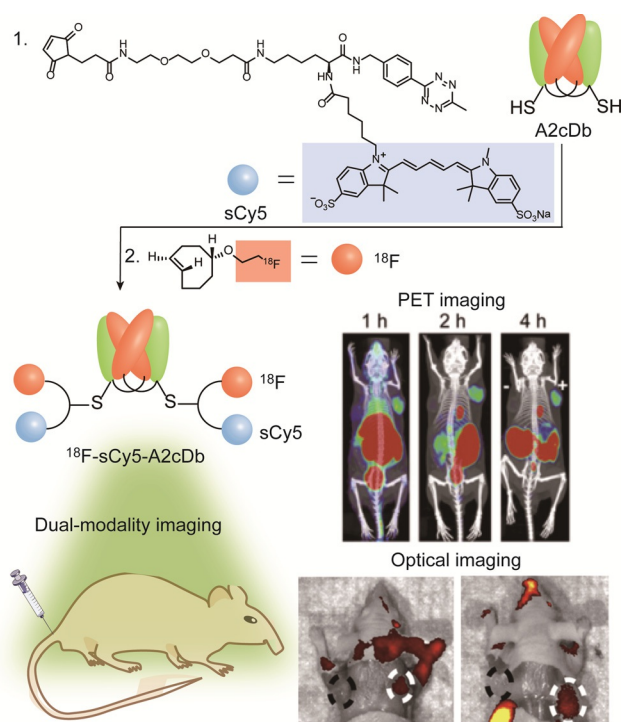


**Figure 15.** (a) Incorporation of a toxic drug (MMAE) and imaging reagent ( $^{89}\text{Zr}$ ) for targeted theranostic applications; (b) Biodistribution data of athymic nude mice bearing HER2-expressing BT474 breast cancer xenografts after 120 h of administration of the corresponding bioconjugate; (c) PET images of athymic nude mice bearing HER2-expressing BT474 breast cancer xenografts after the injection of  $^{89}\text{Zr}$ -trastuzumab-MMAE bioconjugate. (b, c) Adapted with permission from ref. [26]. Copyright (2018) American Chemical Society.

into a single imaging agent, doctors are able to assess the extent of the disease before and after surgery by PET imaging, meanwhile fluorescence imaging can be utilized for image-guided surgery.<sup>[135]</sup> Site-selective dual modification of monoclonal antibodies provides an elegant chemical platform to implement the combination of a radionuclide and a fluorescent dye for dual modality PET and fluorescence imaging. One representative example focused on the conjugation of  $^{18}\text{F}$  and far-red dye sulfonate cyanine 5 (sCy5) to anti-prostate stem cell antigen (PSCA) cysteine mutated diabody A2 via a dual-modality linker, offering the dual-modified imaging probe ( $^{18}\text{F}$ -sCy5-A2cDb) (Figure 16).<sup>[138]</sup>  $^{18}\text{F}$ -immuno-PET showed fast and specific tumor uptake of prostate cancer xenografts, suggesting high-contrast whole-body images with organ-level biodistribution as early as 1 h after injection (Figure 16).<sup>[138]</sup> Postmortem optical imaging confirmed high-contrast fluorescence in the PSCA-expressing tumors and excellent delineation of cancerous cells from surrounding tissue (Figure 16). The dual-modality imaging provides complementary data, significantly contributing to the reliable and accurate diagnostic applications.

## 7. Conclusion and Outlook

The technical breakthroughs in protein bioconjugation chemistry has served as a major driving force for the elucidation of protein trafficking and interactions in living cells as well as the emergence of protein therapeutics. With the increasing need for personalized treatments to address great challenges in biomedical research such as efficient cell or organ targeting, overcoming drug resistance and reducing



**Figure 16.** Incorporation of radiometal and  $^{18}\text{F}$  and sCy5 into A2cDb diabodies for PET imaging and optical imaging. Adapted with permission from ref. [138]. Copyright (2019) Society of Nuclear Medicine and Molecular Imaging.

systemic toxicity, there is a high demand for protein therapeutics with improved features. Multifunctional protein conjugates provide more than one functionality, and they are expected to enhance in vivo performance of therapeutic proteins by combining different functional groups possessing therapy, diagnostics and imaging properties in a single protein system that is tailored to the requirements of the patient. Nevertheless, dual functionalization of proteins remains much more challenging relative to single functionalization owing to the plethora of reactive functional groups on the protein surface and the requirement for the optimal combination of different orthogonal reactions with high efficiencies. In this review, we have summarized the remarkable progress in both synthetic and genetic engineering strategies that have overcome these significant hurdles and thus allow covalent functionalization of proteins with two different functionalities at distinct sites.

Undoubtedly, dual functionalization of proteins has witnessed significant advancements over the past ten years which offers a new arsenal of functional protein conjugates with advantages over singly-functionalized proteins in probing protein dynamics, combination therapy and bioimaging. We envision that tremendous efforts will continue to be devoted to enriching the current methodology toolkit by exploring new chemistries to improve the specificity as well as efficiency of dual functionalization or to achieve a higher level of functionalization, for example, triple functionalization. Moreover, the proper combination of different sophisticated strategies, for example, chemical methods, genetic methods, to capitalize the advantage of different approaches

from the currently expanding toolbox will provide new insights for the preparation of a myriad of functional nanomaterials. Ultimately, we believe that this will pave the way towards “smart” and “intelligent” protein conjugates that can self-adapt to the microenvironment of diseases and provide a self-feedback loop to achieve an output or decision, for example, termination/activation. In this way, current bottlenecks and challenges in therapeutic applications can be addressed with a completely new perspective through rationale chemical design.

### Acknowledgements

The authors are grateful to the Max Planck Society and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project number 316249678-SFB 1279 (Projects A05, C01), for the funding. L.X. is grateful to the China Scholarship Council for a scholarship. We thank Colette J. Whitfield, Nicole Kirsch-Pietz and Zhixuan Zhou for proof-reading. Open access funding enabled and organized by Projekt DEAL.

### Conflict of interest

The authors declare no conflict of interest.

- [1] M. Karplus, J. Kuriyan, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6679–6685.
- [2] O. Shlyk-Kerner, I. Samish, D. Kaftan, N. Holland, P. S. Sai, H. Kless, A. Scherz, *Nature* **2006**, *442*, 827–830.
- [3] C. D. Spicer, B. G. Davis, *Nat. Commun.* **2014**, *5*, 4740.
- [4] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, Jr., *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372; *Angew. Chem.* **2005**, *117*, 7508–7539.
- [5] N. Krall, F. P. da Cruz, O. Boutureira, G. J. L. Bernardes, *Nat. Chem.* **2016**, *8*, 103–113.
- [6] E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, *3*, 147–171.
- [7] O. Boutureira, G. J. L. Bernardes, *Chem. Rev.* **2015**, *115*, 2174–2195.
- [8] X. Chen, Y. W. Wu, *Org. Biomol. Chem.* **2016**, *14*, 5417–5439.
- [9] N. C. Reddy, M. Kumar, R. Molla, V. Rai, *Org. Biomol. Chem.* **2020**, *18*, 4669–4691.
- [10] S. L. Kuan, M. Raabe, *ChemMedChem* **2021**, *16*, 94–104.
- [11] P. Akkapeddi, S. A. Azizi, A. M. Freedy, P. Cal, P. M. P. Gois, G. J. L. Bernardes, *Chem. Sci.* **2016**, *7*, 2954–2963.
- [12] V. Chudasama, A. Maruani, S. Caddick, *Nat. Chem.* **2016**, *8*, 114–119.
- [13] C. Chen, D. Y. W. Ng, T. Weil, *Prog. Polym. Sci.* **2020**, *105*, 101241.
- [14] I. Cobo, M. Li, B. S. Sumerlin, S. Perrier, *Nat. Mater.* **2015**, *14*, 143–159.
- [15] G. Zhang, S. Zheng, H. Liu, P. R. Chen, *Chem. Soc. Rev.* **2015**, *44*, 3405–3417.
- [16] Q. Y. Hu, F. Berti, R. Adamo, *Chem. Soc. Rev.* **2016**, *45*, 1691–1719.
- [17] L. S. Witus, M. B. Francis, *Acc. Chem. Res.* **2011**, *44*, 774–783.
- [18] R. X. Zhang, H. L. Wong, H. Y. Xue, J. Y. Eoh, X. Y. Wu, *J. Controlled Release* **2016**, *240*, 489–503.
- [19] S. García-Alonso, A. Ocaña, A. Pandiella, *Cancer Res.* **2018**, *78*, 2159–2165.
- [20] N. Crawley, M. Thompson, A. Romaschin, *Anal. Chem.* **2014**, *86*, 130–160.
- [21] S. L. Kuan, S. Fischer, S. Hafner, T. Wang, T. Syrovets, W. Liu, Y. Tokura, D. Y. W. Ng, A. Riegger, C. Förtsch, D. Jäger, T. F. E. Barth, T. Simmet, H. Barth, T. Weil, *Adv. Sci.* **2018**, *5*, 1701036.
- [22] S. Chakraborty, B. K. Agrawalla, A. Stumper, N. M. Vegi, S. Fischer, C. Reichardt, M. Kogler, B. Dietzek, M. Feuring-Buske, C. Buske, S. Rau, T. Weil, *J. Am. Chem. Soc.* **2017**, *139*, 2512–2519.
- [23] S. Hafner, M. Raabe, Y. Wu, T. Wang, Z. Zuo, V. Rasche, T. Syrovets, T. Weil, T. Simmet, *Adv. Ther.* **2019**, *2*, 1900084.
- [24] A. Maruani, D. A. Richards, V. Chudasama, *Org. Biomol. Chem.* **2016**, *14*, 6165–6178.
- [25] K. Wang, A. Sachdeva, D. J. Cox, N. M. Wilf, K. Lang, S. Wallace, R. A. Mehl, J. W. Chin, *Nat. Chem.* **2014**, *6*, 393–403.
- [26] P. Adumeau, D. Vivier, S. K. Sharma, J. Wang, T. Zhang, A. Chen, B. J. Agnew, B. M. Zeglis, *Mol. Pharm.* **2018**, *15*, 892–898.
- [27] M. R. Levengood, X. Zhang, J. H. Hunter, K. K. Emmerton, J. B. Miyamoto, T. S. Lewis, P. D. Senter, *Angew. Chem. Int. Ed.* **2017**, *56*, 733–737; *Angew. Chem.* **2017**, *129*, 751–755.
- [28] N. Stephanopoulos, M. B. Francis, *Nat. Chem. Biol.* **2011**, *7*, 876–884.
- [29] J. N. deGruyter, L. R. Malins, P. S. Baran, *Biochemistry* **2017**, *56*, 3863–3873.
- [30] C. B. Rosen, M. B. Francis, *Nat. Chem. Biol.* **2017**, *13*, 697–705.
- [31] D. G. Rawale, K. Thakur, S. R. Adusumalli, V. Rai, *Eur. J. Org. Chem.* **2019**, 6749–6763.
- [32] L. Xu, M. Raabe, M. M. Zegota, J. C. F. Nogueira, V. Chudasama, S. L. Kuan, T. Weil, *Org. Biomol. Chem.* **2020**, *18*, 1140–1147.
- [33] P. Ochdrop, C. P. R. Hackenberger, *Curr. Opin. Chem. Biol.* **2020**, *58*, 28–36.
- [34] J. M. Chalker, G. J. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630–640.
- [35] S. B. Gunnoo, A. Madder, *ChemBioChem* **2016**, *17*, 529–553.
- [36] J. Ravasco, H. Faustino, A. Trindade, P. M. P. Gois, *Chem. Eur. J.* **2019**, *25*, 43–59.
- [37] E. Gil de Montes, E. Jiménez-Moreno, B. Oliveira, C. D. Navo, P. M. S. D. Cal, G. Jiménez-Osés, I. Robina, A. J. Moreno-Vargas, G. J. L. Bernardes, *Chem. Sci.* **2019**, *10*, 4515–4522.
- [38] T. Wang, A. Riegger, M. Lamla, S. Wiese, P. Oeckl, M. Otto, Y. Wu, S. Fischer, H. Barth, S. L. Kuan, T. Weil, *Chem. Sci.* **2016**, *7*, 3234–3239.
- [39] E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute, S. L. Buchwald, *Nature* **2015**, *526*, 687–691.
- [40] J. M. Antos, M. B. Francis, *Curr. Opin. Chem. Biol.* **2006**, *10*, 253–262.
- [41] C. Zhang, E. V. Vinogradova, A. M. Spokoyny, S. L. Buchwald, B. L. Pentelute, *Angew. Chem. Int. Ed.* **2019**, *58*, 4810–4839; *Angew. Chem.* **2019**, *131*, 4860–4892.
- [42] C. E. Hoyle, C. N. Bowman, *Angew. Chem. Int. Ed.* **2010**, *49*, 1540–1573; *Angew. Chem.* **2010**, *122*, 1584–1617.
- [43] C. E. Hoyle, A. B. Lowe, C. N. Bowman, *Chem. Soc. Rev.* **2010**, *39*, 1355–1387.
- [44] C. Canovas, M. Moreau, C. Bernhard, A. Oudot, M. Guillemin, F. Denat, V. Goncalves, *Angew. Chem. Int. Ed.* **2018**, *57*, 10646–10650; *Angew. Chem.* **2018**, *130*, 10806–10810.
- [45] A. L. Baumann, S. Schwagerus, K. Broi, K. Kemnitz-Hassanin, C. E. Stieger, N. Trieloff, P. Schmieder, C. P. R. Hackenberger, *J. Am. Chem. Soc.* **2020**, *142*, 9544–9552.
- [46] R. Tessier, J. Ceballos, N. Guidotti, R. Simonet-Davin, B. Fierz, J. Waser, *Chem* **2019**, *5*, 2243–2263.

- [47] M. E. B. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick, J. R. Baker, *J. Am. Chem. Soc.* **2010**, *132*, 1960–1965.
- [48] N. Martinez-Sáez, S. Sun, D. Oldrini, P. Sormanni, O. Boutur-eira, F. Carboni, I. Compañón, M. J. Deery, M. Vendruscolo, F. Corzana, R. Adamo, G. J. L. Bernardes, *Angew. Chem. Int. Ed.* **2017**, *56*, 14963–14967; *Angew. Chem.* **2017**, *129*, 15159–15163.
- [49] S. J. Walsh, S. Omarjee, W. Galloway, T. T. Kwan, H. F. Sore, J. S. Parker, M. Hyvonen, J. S. Carroll, D. R. Spring, *Chem. Sci.* **2019**, *10*, 694–700.
- [50] A. Maruani, M. E. Smith, E. Miranda, K. A. Chester, V. Chudasama, S. Caddick, *Nat. Commun.* **2015**, *6*, 6645.
- [51] K. Li, W. Wang, J. Gao, *Angew. Chem. Int. Ed.* **2020**, *59*, 14246–14250; *Angew. Chem.* **2020**, *132*, 14352–14356.
- [52] H. Faustino, M. Silva, L. F. Veiros, G. J. L. Bernardes, P. M. P. Gois, *Chem. Sci.* **2016**, *7*, 5052–5058.
- [53] J. I. MacDonald, H. K. Munch, T. Moore, M. B. Francis, *Nat. Chem. Biol.* **2015**, *11*, 326–331.
- [54] J.-R. Deng, N. C.-H. Lai, K. K.-Y. Kung, B. Yang, S.-F. Chung, A. S.-L. Leung, M.-C. Choi, Y.-C. Leung, M.-K. Wong, *Commun. Chem.* **2020**, *3*, 67.
- [55] A. C. Obermeyer, J. B. Jarman, M. B. Francis, *J. Am. Chem. Soc.* **2014**, *136*, 9572–9579.
- [56] G. Liang, H. Ren, J. Rao, *Nat. Chem.* **2010**, *2*, 54–60.
- [57] S. Pomplun, M. Y. H. Mohamed, T. Oelschlaegel, C. Wellner, F. Bergmann, *Angew. Chem. Int. Ed.* **2019**, *58*, 3542–3547; *Angew. Chem.* **2019**, *131*, 3580–3585.
- [58] A. O. Chan, C. M. Ho, H. C. Chong, Y. C. Leung, J. S. Huang, M. K. Wong, C. M. Che, *J. Am. Chem. Soc.* **2012**, *134*, 2589–2598.
- [59] M. T. Taylor, J. E. Nelson, M. G. Suero, M. J. Gaunt, *Nature* **2018**, *562*, 563–568.
- [60] A. H. Christian, S. Jia, W. Cao, P. Zhang, A. T. Meza, M. S. Sigman, C. J. Chang, F. D. Toste, *J. Am. Chem. Soc.* **2019**, *141*, 12657–12662.
- [61] J. R. Kramer, T. J. Deming, *Biomacromolecules* **2012**, *13*, 1719–1723.
- [62] M. W. Jones, G. Mantovani, C. A. Blindauer, S. M. Ryan, X. Wang, D. J. Brayden, D. M. Haddleton, *J. Am. Chem. Soc.* **2012**, *134*, 7406–7413.
- [63] D. W. Romanini, M. B. Francis, *Bioconjugate Chem.* **2008**, *19*, 153–157.
- [64] S. D. Tilley, M. B. Francis, *J. Am. Chem. Soc.* **2006**, *128*, 1080–1081.
- [65] H. Ban, M. Nagano, J. Gavriljuk, W. Hakamata, T. Inokuma, C. F. Barbas 3rd, *Bioconjugate Chem.* **2013**, *24*, 520–532.
- [66] J. M. Antos, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2009**, *131*, 6301–6308.
- [67] Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki, M. Kanai, *J. Am. Chem. Soc.* **2016**, *138*, 10798–10801.
- [68] A. J. Reay, T. J. Williams, I. J. Fairlamb, *Org. Biomol. Chem.* **2015**, *13*, 8298–8309.
- [69] H. R. Alexandra Schischko, N. Kaplaneris, L. Ackermann, *Angew. Chem. Int. Ed.* **2017**, *56*, 1576–1580; *Angew. Chem.* **2017**, *129*, 1598–1602.
- [70] Z. Ruan, N. Saueremann, E. Manoni, L. Ackermann, *Angew. Chem. Int. Ed.* **2017**, *56*, 3172–3176; *Angew. Chem.* **2017**, *129*, 3220–3224.
- [71] M. B. Hansen, F. Hubalek, T. Skrydstrup, T. Hoeg-Jensen, *Chem. Eur. J.* **2016**, *22*, 1572–1576.
- [72] X. Li, H. Y. Lam, Y. Zhang, C. K. Chan, *Org. Lett.* **2010**, *12*, 1724–1727.
- [73] J. C. Vantourout, S. R. Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padial, A. Istrate, K. Maziarz, J. N. deGruyter, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes, P. S. Baran, *J. Am. Chem. Soc.* **2020**, *142*, 17236–17242.
- [74] Y. Gong, D. Andina, S. Nahar, J. C. Leroux, M. A. Gauthier, *Chem. Sci.* **2017**, *8*, 4082–4086.
- [75] C. R. Bertozzi, *Acc. Chem. Res.* **2011**, *44*, 651–653.
- [76] E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998; *Angew. Chem.* **2009**, *121*, 7108–7133.
- [77] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599; *Angew. Chem.* **2002**, *114*, 2708–2711.
- [78] J. C. Jewett, C. R. Bertozzi, *Chem. Soc. Rev.* **2010**, *39*, 1272–1279.
- [79] H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, *464*, 441–444.
- [80] J. A. Walker, J. J. Bohn, F. Ledesma, M. R. Sorkin, S. R. Kabaria, D. N. Thornlow, C. A. Alabi, *Bioconjugate Chem.* **2019**, *30*, 2452–2457.
- [81] B. Wu, Z. Wang, Y. Huang, W. R. Liu, *ChemBioChem* **2012**, *13*, 1405–1408.
- [82] A. Sachdeva, K. Wang, T. Elliott, J. W. Chin, *J. Am. Chem. Soc.* **2014**, *136*, 7785–7788.
- [83] R. K. Lim, Q. Lin, *Chem. Commun.* **2010**, *46*, 1589–1600.
- [84] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764–4806.
- [85] C. P. Ramil, Q. Lin, *Curr. Opin. Chem. Biol.* **2014**, *21*, 89–95.
- [86] Z. Li, L. Qian, L. Li, J. C. Bernhammer, H. V. Huynh, J. S. Lee, S. Q. Yao, *Angew. Chem. Int. Ed.* **2016**, *55*, 2002–2006; *Angew. Chem.* **2016**, *128*, 2042–2046.
- [87] B. L. Oliveira, Z. Guo, G. J. L. Bernardes, *Chem. Soc. Rev.* **2017**, *46*, 4895–4950.
- [88] A. P. Crochet, M. M. Kabir, M. B. Francis, C. D. Paavola, *Biosens. Bioelectron.* **2010**, *26*, 55–61.
- [89] B. K. Agrawalla, T. Wang, A. Riegger, M. P. Domogalla, K. Steinbrink, T. Dorfler, X. Chen, F. Boldt, M. Lamla, J. Michaelis, S. L. Kuan, T. Weil, *Bioconjugate Chem.* **2018**, *29*, 29–34.
- [90] V. Chudasama, M. E. Smith, F. F. Schumacher, D. Papaioannou, G. Waksman, J. R. Baker, S. Caddick, *Chem. Commun.* **2011**, *47*, 8781–8783.
- [91] Y. Zhang, C. Zang, G. An, M. Shang, Z. Cui, G. Chen, Z. Xi, C. Zhou, *Nat. Commun.* **2020**, *11*, 1015.
- [92] J. Morales-Sanfrutos, J. Lopez-Jaramillo, M. Ortega-Munoz, A. Megia-Fernandez, F. Perez-Balderas, F. Hernandez-Mateo, F. Santoyo-Gonzalez, *Org. Biomol. Chem.* **2010**, *8*, 667–675.
- [93] J. Morales-Sanfrutos, F. J. Lopez-Jaramillo, F. Hernandez-Mateo, F. Santoyo-Gonzalez, *J. Org. Chem.* **2010**, *75*, 4039–4047.
- [94] P. Moody, V. Chudasama, R. I. Nathani, A. Maruani, S. Martin, M. E. Smith, S. Caddick, *Chem. Commun.* **2014**, *50*, 4898–4900.
- [95] R. I. Nathani, P. Moody, V. Chudasama, M. E. Smith, R. J. Fitzmaurice, S. Caddick, *Chem. Sci.* **2013**, *4*, 3455–3458.
- [96] C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. Van Voorhis, B. L. Pentelute, *Nat. Chem.* **2016**, *8*, 120–128.
- [97] J. W. Chin, *Nature* **2017**, *550*, 53–60.
- [98] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105–1109.
- [99] A. R. Nodding, L. A. Spear, T. L. Williams, L. Y. P. Luk, Y. H. Tsai, *Essays Biochem.* **2019**, *63*, 237–266.
- [100] W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell, W. R. Liu, *Angew. Chem. Int. Ed.* **2010**, *49*, 3211–3214; *Angew. Chem.* **2010**, *122*, 3279–3282.
- [101] E. A. L. Eric, M. Brustad, P. G. Schultz, A. A. Deniz, *J. Am. Chem. Soc.* **2008**, *130*, 17664–17665.
- [102] H. Xiao, A. Chatterjee, S. H. Choi, K. M. Bajjuri, S. C. Sinha, P. G. Schultz, *Angew. Chem. Int. Ed.* **2013**, *52*, 14080–14083; *Angew. Chem.* **2013**, *125*, 14330–14333.
- [103] K. Wang, W. H. Schmied, J. W. Chin, *Angew. Chem. Int. Ed.* **2012**, *51*, 2288–2297; *Angew. Chem.* **2012**, *124*, 2334–2344.
- [104] E. M. Milczek, *Chem. Rev.* **2018**, *118*, 119–141.

- [105] Y. Yang, S. Lin, W. Lin, P. R. Chen, *ChemBioChem* **2014**, *15*, 1738–1743.
- [106] M. D. Lee, W. Y. Tong, T. Nebl, L. A. Pearce, T. M. Pham, A. Golbaz-Hagh, S. Puttick, S. Rose, T. E. Adams, C. C. Williams, *Bioconjugate Chem.* **2019**, *30*, 2539–2543.
- [107] M. Rashidian, S. C. Kumarapperuma, K. Gabrielse, A. Fegan, C. R. Wagner, M. D. Distefano, *J. Am. Chem. Soc.* **2013**, *135*, 16388–16396.
- [108] I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, *2*, 99–104.
- [109] L. Schmohl, D. Schwarzer, *Curr. Opin. Chem. Biol.* **2014**, *22*, 122–128.
- [110] G. K. T. Nguyen, Y. Cao, W. Wang, C. F. Liu, J. P. Tam, *Angew. Chem. Int. Ed.* **2015**, *54*, 15920–15924; *Angew. Chem.* **2015**, *127*, 15920–15924.
- [111] Y. Zhang, M. J. Blanden, C. Sudheer, S. A. Gangopadhyay, M. Rashidian, J. L. Houglund, M. D. Distefano, *Bioconjugate Chem.* **2015**, *26*, 2542–2553.
- [112] L. Deweid, L. Neureiter, S. Englert, H. Schneider, J. Deweid, D. Yanakieva, J. Sturm, S. Bitsch, A. Christmann, O. Avrutina, H. L. Fuchsbaauer, H. Kolmar, *Chem. Eur. J.* **2018**, *24*, 15195–15200.
- [113] J. G. Plaks, R. Falatach, M. Kastantin, J. A. Berberich, J. L. Kaar, *Bioconjugate Chem.* **2015**, *26*, 1104–1112.
- [114] P. Wu, W. Shui, B. L. Carlson, N. Hu, D. Rabuka, J. Lee, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3000–3005.
- [115] J. Grünwald, H. E. Klock, S. E. Cellitti, B. Bursulaya, D. McMullan, D. H. Jones, H. P. Chiu, X. Wang, P. Patterson, H. Zhou, J. Vance, E. Nigoghossian, H. Tong, D. Daniel, W. Mallet, W. Ou, T. Uno, A. Brock, S. A. Lesley, B. H. Geierstanger, *Bioconjugate Chem.* **2015**, *26*, 2554–2562.
- [116] D. Schumacher, J. Helma, F. A. Mann, G. Pichler, F. Natale, E. Krause, M. C. Cardoso, C. P. Hackenberger, H. Leonhardt, *Angew. Chem. Int. Ed.* **2015**, *54*, 13787–13791; *Angew. Chem.* **2015**, *127*, 13992–13996.
- [117] S. C. Reddington, M. Howarth, *Curr. Opin. Chem. Biol.* **2015**, *29*, 94–99.
- [118] T. J. Harmand, D. Bousbaine, A. Chan, X. Zhang, D. R. Liu, J. P. Tam, H. L. Ploegh, *Bioconjugate Chem.* **2018**, *29*, 3245–3249.
- [119] J. S. Italia, P. S. Addy, S. B. Erickson, J. C. Peeler, E. Weerapana, A. Chatterjee, *J. Am. Chem. Soc.* **2019**, *141*, 6204–6212.
- [120] P. S. Addy, S. B. Erickson, J. S. Italia, A. Chatterjee, *J. Am. Chem. Soc.* **2017**, *139*, 11670–11673.
- [121] D. L. Dunkelmann, J. C. W. Willis, A. T. Beattie, J. W. Chin, *Nat. Chem.* **2020**, *12*, 535–544.
- [122] A. Maruani, P. A. Szijj, C. Bahou, J. C. F. Nogueira, S. Caddick, J. R. Baker, V. Chudasama, *Bioconjugate Chem.* **2020**, *31*, 520–529.
- [123] B. Schuler, W. A. Eaton, *Curr. Opin. Struct. Biol.* **2008**, *18*, 16–26.
- [124] K. L. Pierce, R. T. Premont, R. J. Lefkowitz, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 639–650.
- [125] J.-P. Villardaga, M. Bunemann, C. Krasel, M. Castro, M. J. Lohse, *Nat. Biotechnol.* **2003**, *21*, 807–812.
- [126] B. A. Griffin, R. Y. Tsien, *Science* **1998**, *281*, 269–271.
- [127] C. Hoffmann, G. Gaietta, M. Bunemann, S. R. Adams, S. Oberdorff-Maass, B. Behr, J. P. Villardaga, R. Y. Tsien, M. H. Ellisman, M. J. Lohse, *Nat. Methods* **2005**, *2*, 171–176.
- [128] L. D. Lavis, R. T. Raines, *ACS Chem. Biol.* **2014**, *9*, 855–866.
- [129] J. L. Fallon, F. A. Quijoch, *Structure* **2003**, *11*, 1303–1307.
- [130] K. P. Hoeflich, M. Ikura, *Cell* **2002**, *108*, 739–742.
- [131] G. Pasut, F. M. Veronese, *Adv. Drug Delivery Rev.* **2009**, *61*, 1177–1188.
- [132] R. P. Lyon, T. D. Bovee, S. O. Doronina, P. J. Burke, J. H. Hunter, H. D. Neff-LaFord, M. Jonas, M. E. Anderson, J. R. Setter, P. D. Senter, *Nat. Biotechnol.* **2015**, *33*, 733–735.
- [133] F. Li, K. K. Emmerton, M. Jonas, X. Zhang, J. B. Miyamoto, J. R. Setter, N. D. Nicholas, N. M. Okeley, R. P. Lyon, D. R. Benjamin, C. L. Law, *Cancer Res.* **2016**, *76*, 2710–2719.
- [134] R. Chen, J. Hou, E. Newman, Y. Kim, C. Donohue, X. Liu, S. H. Thomas, S. J. Forman, S. E. Kane, *Mol. Cancer Ther.* **2015**, *14*, 1376–1384.
- [135] M. Wu, J. Shu, *Contrast Media Mol. Imaging* **2018**, *2018*, 1382183.
- [136] S. Hernandez Vargas, S. C. Ghosh, A. Azhdarinia, *J. Nucl. Med.* **2019**, *60*, 459–465.
- [137] B. M. Zeglis, C. B. Davis, D. Abdel-Atti, S. D. Carlin, A. Chen, R. Aggeler, B. J. Agnew, J. S. Lewis, *Bioconjugate Chem.* **2014**, *25*, 2123–2128.
- [138] K. A. Zettlitz, C. M. Waldmann, W. K. Tsai, R. Tavares, J. Collins, J. M. Murphy, A. M. Wu, *J. Nucl. Med.* **2019**, *60*, 1467–1473.

Manuscript received: September 3, 2020

Accepted manuscript online: December 1, 2020

Version of record online: February 26, 2021