

Dual modification of biomolecules

*Antoine Maruani**, Daniel A. Richards and Vijay Chudasama*

* *Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK.
Tel: +44 (0)20 7679 2077; E-mails: antoine.maruani.12@ucl.ac.uk and v.chudasama@ucl.ac.uk*

Biographies

Dr Antoine Maruani obtained his Master's degree in Chemistry from École Normale Supérieure of Lyon (France) in 2011. He then joined Prof. Stephen Caddick's group at University College London (UK) where he obtained his PhD in 2015 with his thesis focusing on the site-selective dual modification of proteins. He is currently working as a postdoctoral fellow under the supervision of Dr Vijay Chudasama on the development of novel methodologies for bioconjugation.

Dr Daniel Richards began his studies at the University Of York, obtaining an MChem in 2011 with his thesis focused on the study of metal-halogen exchange reactions in nitrogen containing heterocycles. He subsequently joined the lab of Dr James Baker at University College London (UCL) to study for his PhD, focusing on the development and application of novel biocompatible photochemical reactions. He is currently working as a postdoctoral fellow in the group of Dr Vijay Chudasama, developing novel methods for the selective functionalisation of nanoparticles.

Dr Vijay Chudasama obtained his MSci degree and PhD from University College London in 2008 and 2011, respectively. Following post-doctoral studies under the supervision of Prof. Stephen Caddick, Vijay obtained a Ramsay Memorial Fellowship. During this time, he was made Technical Director of a biotechnology spin-out (ThioLogics). In April 2015, he was awarded a lectureship at UCL for him to focus on the research areas of aerobic C-H activation and various aspects of Chemical Biology. Vijay's research has recently been highlighted by Forbes, Scientific American, CNN, Nature Chemistry and the Royal Society of Chemistry.

Abstract

With the advent of novel bioorthogonal reactions and “click” chemistry, an increasing number of strategies for the single labelling of proteins and oligonucleotides have emerged. Whilst several methods exist for the site-selective introduction of a single chemical moiety, site-selective and bioorthogonal dual modification of biomolecules remains a challenge. The introduction of multiple modules enables a plethora of permutations and combinations and can generate a variety of bioconjugates with many potential applications. From *de novo* approaches on oligomers to the post-translational functionalisation of proteins, this review will highlight the main strategies to dually modify biomolecules.

Review

1. Introduction

a. Bioorthogonal reactions and click chemistry on biomolecules

The last decade has witnessed the rapid and sustained development of novel bioorthogonal reactions that display exquisite chemoselectivity in the presence of biomolecules. In this burgeoning chemical scene of site-selective functionalisation, various approaches for labelling different classes of biomolecules (*e.g.* oligonucleotides, proteins) with a single entity have been developed. The amount of research being conducted in this field and related areas is growing and a vast number of papers and reviews have been published in recent years.¹⁻⁹

There are two main strategies for the selective and bioorthogonal chemical labelling of oligonucleotides. The first consists of incorporating a nucleic acid analogue (*e.g.* a nucleoside phosphoramidite), usually under automated solid-phase synthesis conditions.^{10, 11} Importantly, this method is restricted to labels or chemical handles that can withstand the harsh conditions of oligonucleotide synthesis and deprotection.¹² The second is post-synthetic labelling in solution, which allows for incorporation of labels too sensitive for the first approach.¹³

Selectivity in the modification of proteins is also commonly achieved *via* two distinct approaches.

Either the reactivity of a chemical reagent is adjusted to be selective for a single natural amino acid or an unnatural amino acid (uAA) is incorporated into the target protein, which offers a unique reactivity that is orthogonal to the surrounding groups. Thus, using these approaches, proteins can be manipulated at the single amino acid level to allow numerous applications both *in vitro* and *in vivo*.

Regardless of the approach, or the class of biomolecule to be functionalised, the chemoselective reaction(s) required for site-selective modification often fall under the category of “click” reactions. Click chemistry has been defined as a set of reactions that are modular, stereospecific, wide in scope, high yielding, generate only inoffensive by-products, have simple reaction conditions and generate products that are easy to purify.¹⁴ Click chemistry and bioorthogonal chemistry therefore often reflect similar chemical principles.¹⁵ Indeed, click reactions generally offer water compatible chemistry and their potential for bioorthogonal, regioselective and high-yielding conjugation under mild conditions make them ideal candidates for the chemical functionalisation of biomolecules. For these reasons, among the numerous functionalisation strategies for biomolecules, the use of “click” chemistry has received particular attention. Due to its ever-growing importance, “click” chemistry on proteins and oligonucleotides has been reviewed extensively in the past few years, although they are mainly focused on the single modification of biomolecules.¹⁵⁻²¹

b. Dual modification of biomolecules

Whilst several methods exist for the site-selective introduction of a single chemical moiety, site-selective and bioorthogonal dual modification remains a challenge. One early method used to incorporate two functionalities into biomolecules is the *de novo* total synthesis of oligomers, though this method is still suboptimal. Whilst bis-heterolabelled oligonucleotides could be prepared successfully,^{12,22-24} this approach was not amenable to long peptides and full-length proteins and, despite recent developments in this field, remains limited to short fragments.^{25, 26}

c. Multifunctional scaffolds – Applications

The use of dual modification is essential for certain applications that require at least two functional groups to be introduced into a protein, *e.g.* Förster resonance energy transfer (FRET),²⁷ experiments for folding studies,²⁸ targeted theranostic biopharmaceuticals,²⁹⁻³¹ or protein–drug therapeutics stabilised with conjugated half-life extension chemistry.^{31,32} Furthermore, attachment of fluorescent probes or affinity tags to oligonucleotides is critical to research in nanotechnology and modern life sciences.^{22, 33, 34} For example, site-specific functionalisation with two or more distinct FRET pairs allows for the development of oligonucleotide probes for imaging purposes.^{35, 36} These systems attract high interest since they are essential for techniques such as flow cytometry,³⁷ real-time polymerase chain reaction,³⁸ and advanced fluorescence microscopy.³⁹ Indeed, multifunctional scaffolds have a myriad of uses and can find significant utility in a wide range of areas. It is therefore crucial to develop more straightforward, practical and versatile methods for the introduction of two modalities into biomolecules.

The aim of this review is to give an overview of the current existing strategies for the dual modification of DNA, RNA and proteins. As recent research has shown a clear shift towards site-selective methodologies, these approaches will be given particular emphasis. We hope this review will allow the reader to grasp the main strategies and uses of this emerging yet rapidly expanding field.

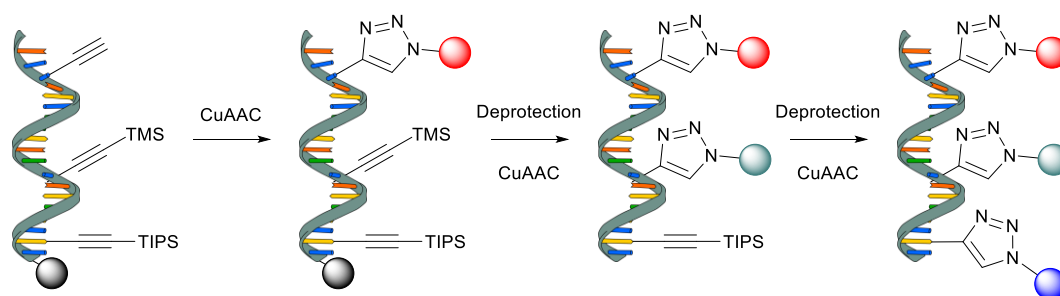
2. Dual modification of oligonucleotides

Due to their remarkable efficiency, “click” reactions, in particular the Cu^I-catalysed [3+2] alkyne–azide cycloaddition (CuAAC), have enabled a plethora of novel ideas and applications to be readily appraised in the nucleic acid field.²¹ It is well documented that, in addition to the high efficiency of these reactions and the mild conditions required, alkynes and azides can be incorporated into oligomers without greatly affecting their biophysical properties.

a. DNA

i. Sequential approaches

Gramlich *et al.* were the first to report a CuAAC-based method for the multiple, sequential labelling of small DNAs.¹² By using orthogonal protection/deprotection strategies with alkyne-bearing nucleotides capped with trimethylsilyl and triisopropylsilyl groups, highly efficient and sequential conjugations of a variety of azide-containing molecules (*e.g.* sugars, fluorophores) were demonstrated (Scheme 1).

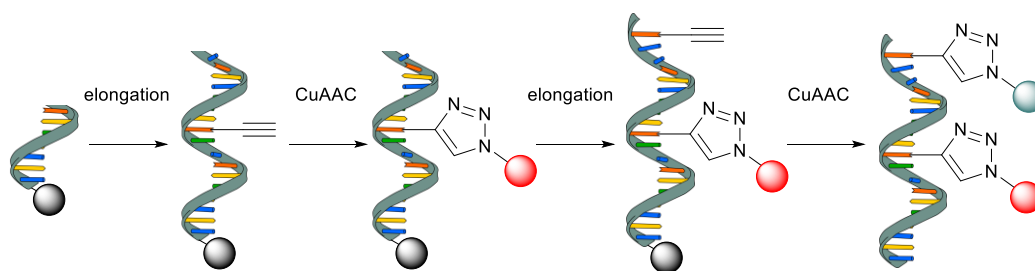


Scheme 1: Protection/deprotection strategy for DNA dual modification.

Interestingly, this triple click strategy was shown to be amenable both on solid support and in solution with variable efficiency, and the labels could also be introduced outside the nucleobases by using alkyne-bearing non-nucleoside DNA modifiers.

Another approach to dually modify DNA is to first incorporate a “clickable” nucleotide, then pause the elongation of the chain to react it, add a second reactive nucleotide and continue elongation before reacting the second reactive nucleotide; thus avoiding the need for additional protection/deprotection steps. The first functionalisation can occur during automated oligonucleotide synthesis and the second can either take place on solid support or post synthesis in solution to allow for the incorporation of more sensitive labels.^{22, 23}

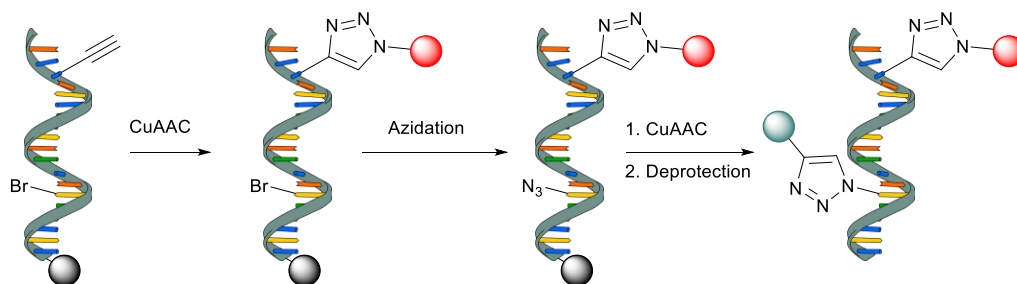
Using this approach, Astakhova *et al.* optimised CuAAC conditions and developed a strategy for internal labelling of DNA by combining solid-phase and postsynthetic CuAAC reactions (Scheme 2).²²



Scheme 2: Sequential double CuAAC on oligonucleotide with elongation interruption.

This method was exemplified through the preparation of novel fluorescent probes containing a variety of fluorophores in moderate to excellent yields. The main challenge of this strategy was the incompatibility of oligonucleotide synthesis with the aqueous conditions required for CuAAC which resulted in additional washing and drying steps.

Alternatively, instead of interrupting the elongation at each step of incorporation of a chemical handle (*e.g.* alkyne), Pourceau *et al.* showed that azide functions could be incorporated post-synthetically on a DNA that possessed reacted alkynes and pendant alkyl bromides, thus allowing for specific and sequential dual modification. Using an alkyne-bearing and a bromohexyl-bearing phosphoramidite derivative as a precursor to azide functionalisation, a series of glyco oligonucleotide conjugates were successfully prepared (Scheme 3).²³ The carbohydrates could be introduced either at the 3'- or 5'-end of oligonucleotides using microwave-assisted CuAAC in high conversion and moderate isolated yields. Interestingly, reactions on solid support and in solution were compared and led to similar purity profiles and isolated yields. The use of bromohexyl as a precursor to azide functionalisation was necessary as azides can react with P^{III} via a Staudinger reaction during phosphoramidite-based oligonucleotide synthesis.⁴⁰



Scheme 3: Sequential double CuAAC on DNA with post-synthetic azidation.

ii. One pot modification

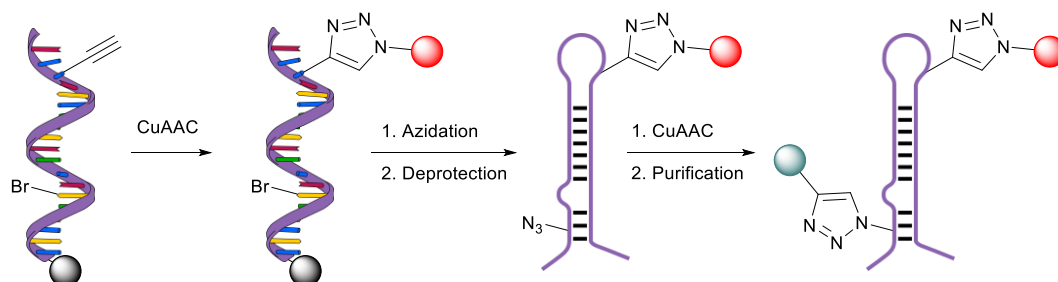
In an attempt to simplify the procedure for the dual modification of DNA, it was demonstrated that two orthogonal functional groups can also be added to the oligonucleotide which can then be simultaneously dual-modified. Indeed, by combining CuAAC click reaction and inverse electron-demand Diels–Alder reaction³⁶, or strain-promoted azide–alkyne [3+2] cycloaddition (SPAAC) and NHS ester coupling⁴¹, oligonucleotide FRET probes could be prepared in a one-pot, site specific and high yielding process.

b. RNA

Pradère *et al.* adapted a solid phase synthesis-based method for double hetero-labelling of DNA to long pre-microRNAs.²² In this approach, a single alkyne-modified nucleoside is attached and the synthesis is interrupted to allow for selective reaction of this first reactive handle. Synthesis is subsequently resumed and the second alkyne-modified nucleoside is attached and reacted with a second label (Scheme 2).²⁴ Although this afforded the desired modified RNAs, incomplete CuAAC resulted in cross-labelling, affording complex mixtures which proved inseparable in some cases.

Santner *et al.* described a method to synthesise siRNA with a 3' terminal azide and a primary amine using phosphoramidite nucleoside. These two handles can be sequentially reacted with an alkyne and an activated acid respectively to enable access to hetero bifunctional RNA.⁴² Although this approach alleviates undesirable cross-labelling, the use of the azido-modified nucleoside is restricted to the 3'-end of RNA due to the incompatibility of azides and the phosphoramidite method.⁴⁰ However, taking inspiration from a method for bis-heterolabelling of short DNA that uses an alkyne- and a bromide-bearing phosphoramidite²³, Pradère *et al.* recently developed a method for double hetero-labelling of pre-microRNAs that circumvents this restriction.⁴³ By incorporating and reacting an alkyne and a bromide-bearing nucleoside sequentially, a series of dually modified pre-microRNAs labelled at various positions were prepared in high purity (Scheme 4). The prepared oligonucleotides bearing

fluorescent dyes and forming a fluorescence resonance energy transfer pair were then used as probes for miRNA maturation in an *in vitro* assay.



Scheme 4: Bis-heterolabelling of pre-miRNA.

3. Dual modification of peptides and proteins

The main approaches to dually modify peptides and proteins can be classified in two separate categories (Figure 1).⁴⁴ The first one is the modification of two amino acids on the proteins using various methods that include: exploiting the difference of accessibility or reactivity of amino acids,⁴⁵⁻⁴⁷ functionalisation of the C- and N-terminus of proteins,⁴⁸ use of unnatural amino acids,⁴⁹⁻⁵¹ use of specific enzymes⁵⁰ or a combination of these strategies⁵². The second category involves conjugation of a multifunctional linker to the protein of interest. The multifunctional scaffold can then be reacted to give a dually modified conjugate.^{19, 31, 44, 53}

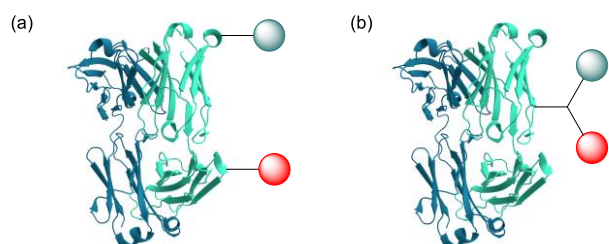


Figure 1: Different approaches for the introduction of two functional moieties to a protein. (a) *Via* two amino acids; (b) *Via* a multifunctional scaffold containing the two reactive handles.

a. Approach using two amino acids

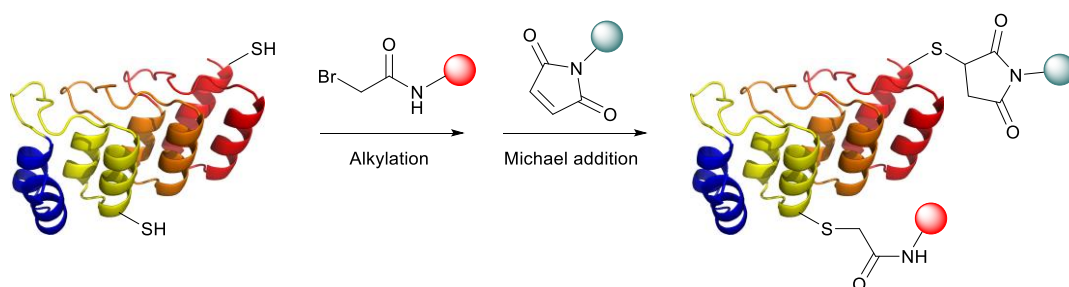
i. Native amino acids

1. Identical amino acids with different reactivity

By exploiting the difference of accessibility or reactivity of two identical amino acids, it is possible to site selectively dual-modify proteins. To demonstrate the feasibility of this approach, Ratner *et al.* engineered a dual cysteine mutant of *Escherichia coli* adenylate kinase and, by carefully selecting the position of the engineered cysteines, the reaction times and the stoichiometry of reagents, the authors could generate doubly labelled proteins.⁴⁵ This method relies on sufficient difference in reactivity of the cysteines so that the most reactive cysteine can be selectively reacted by addition of a weakly thiol-reactive reagent, and the less reactive cysteine subsequently modified with a highly thiol-reactive reagent. After purification of the mixture obtained, the resulting conjugates were isolated in *ca.* 70% yield and were further used for spectroscopic studies of protein folding and dynamics.

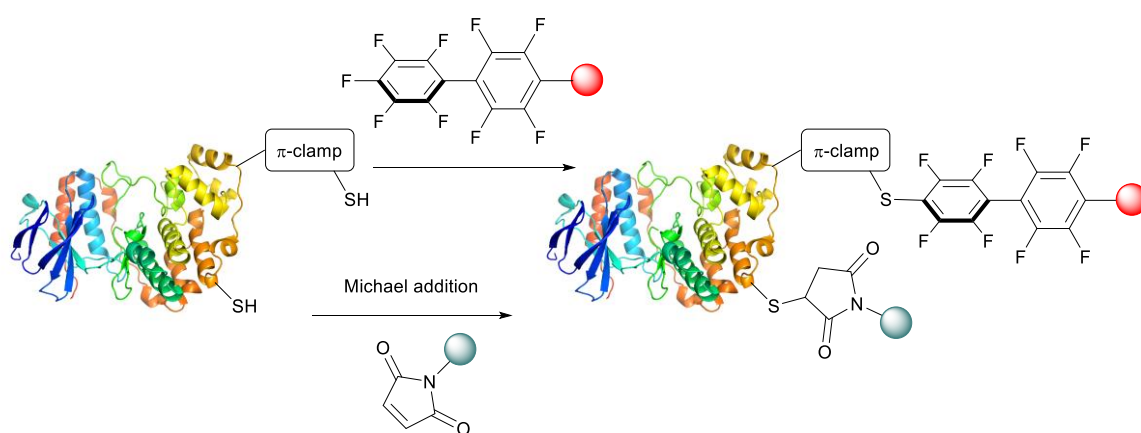
Although this is a powerful and useful concept, it heavily relies on identifying cysteines with suitable differences in reactivity, which is not straightforward or easily predicted. A reaction with moderate selectivity would lead to heterogeneous mixture of products and would require intermediate purification steps.

Nevertheless, using a similar approach and through identification of cysteine environments with significant differences in nucleophilicity, Moody *et al.* engineered cysteine mutants of designed ankyrin repeat proteins (DARPin)s that can be sequentially and quantitatively modified with appropriate thiol-reactive reagents.⁴⁷ Due to the high selectivity of the reactions, their method for the dual modification of DARPins allowed for clean and high-yielding sequential functionalisation without the need for intermediate purification (Scheme 5).



Scheme 5: Dual modification of DARPins

Recently, Zhang *et al.* showed that a four-amino-acid sequence (Phe-Cys-Pro-Phe), called the “ π -clamp” tunes the reactivity of its cysteine thiol for site-selective conjugation with perfluoroaromatic reagents.⁵⁴ The “ π -clamp” strategy was exemplified on a fusion protein of the anthrax toxin lethal factor 1–263 and diphtheria toxin domain A with an engineered N-terminal cysteine and a C-terminal π -clamp. Using a one-pot procedure, the protein was successfully dually modified with a small peptide and a fluorophore in a site-selective manner (Scheme 6).



Scheme 6: Dual modification by π -clamp approach.

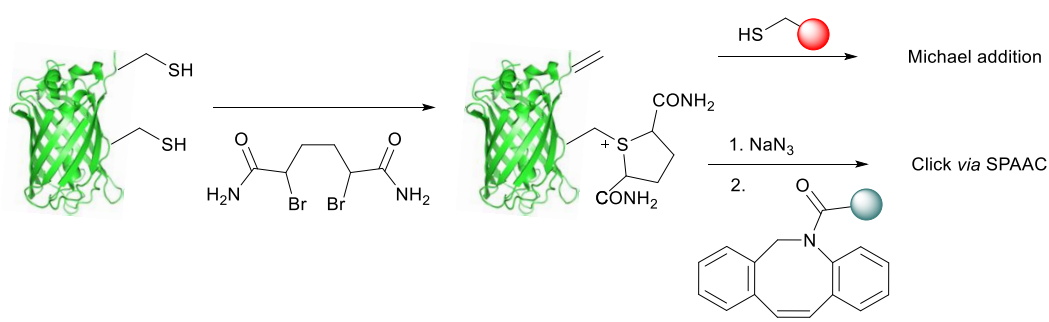
The aforementioned strategies require protein engineering to incorporate cysteines at specific positions which can lead to complications during expression or functionalisation (*e.g.* aggregation, scrambling).⁵⁵

The approaches described above for protein modification rely on the nucleophilicity of the target residue. Selectivity is achieved almost exclusively by using protein modification reagents carefully designed to exploit the differences in reactivity of the amino acid side chains under carefully controlled conditions.

Taking a different approach, Nathani *et al.* showed that the outcome of a reaction between cysteine residues and a single reagent (*i.e.* 2,5-dibromohexanediamide) could be controlled using only the local

protein microenvironment.⁴⁶ This site-selective dual functionalisation of a protein based on substrate control was exemplified on a double cysteine mutant of superfolder green fluorescent protein.

The protein was initially transformed into a bis-sulfonium through reaction with 2,5-dibromohexanediamide. Depending on the protein microenvironment, the sulfonium either persisted as a stable entity or eliminated to yield dehydroalanine. These functional groups were subsequently reacted independently to yield a site-selective dually modified protein (Scheme 7).

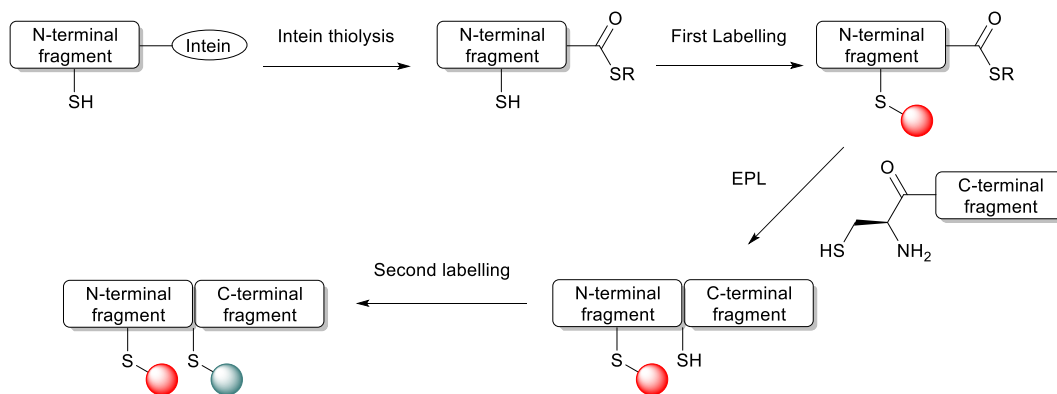


Scheme 7: Dual modification by substrate control using 2,5-dibromohexanediamide.

2. Identical amino acids sequentially

Instead of taking advantage of the differential reactivity of cysteine residues, it is also possible to independently prepare two functionalisable fragments of a protein, modify them independently and then join them together to generate the dually modified biomolecule. This can be achieved using expressed protein ligation (EPL) either on solid support⁵⁶ or in solution^{48, 57}.

De Rosa *et al.* described a homogeneous dual-labelling strategy where EPL, a method related to native chemical ligation which expands its applicability to recombinant protein fragments, was exploited.⁵⁷ A cysteine-containing N-terminal fragment of a protein was first generated and it was reacted with a probe. The second C-terminal fragment was then attached using EPL and finally modified with a second cysteine-reactive probe. This strategy was also applied to a full length protein by replacing the second fragment by a cysteine (Scheme 8). This method was exemplified with a series of repeat protein CTPR3 (*Consensus* TetratricoPeptideRepeat) functionalised with fluorophore pairs at different positions in the protein.



Scheme 8: Dual labelling of proteins using EPL.

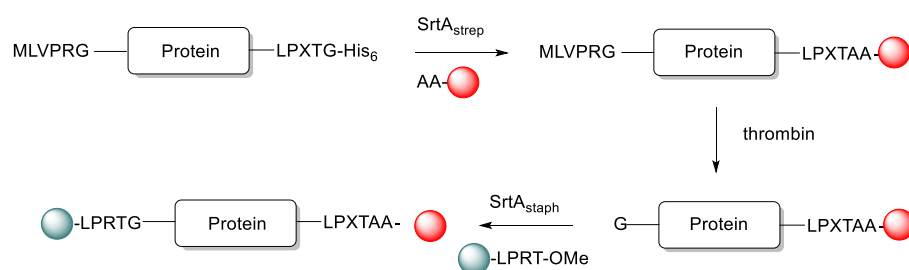
Yang *et al.* reported an alternative approach to cysteine labelling of fragments to afford a dual-labelled di-ubiquitin structure, thus facilitating FRET studies.⁵⁸ The approach relied upon generating protein fragments that contained a single cysteine for fluorophore conjugation, labelling each individual cysteine, and joining the fragments to generate a full-length protein through split intein-mediated protein splicing. Inteins are peptide sequences that mediate self-splicing when inserted within a larger precursor polypeptide. They are eliminated from the precursor sequence as they facilitate conjugation of the N- and C-terminal regions into a new protein.⁵⁹

Both EPL and protein splicing are powerful methods that have allowed studies in areas such as electrophysiology for the monitoring of ion channels,⁶⁰ NMR through the introduction of specific isotopes,⁶¹ and optical spectroscopy;^{62, 63} all with the aim of probing protein structure and function. Interestingly, both approaches could potentially be adapted to the introduction of more than two molecular probes.

Enzymatic approaches for the dual modification of proteins have also been developed.⁴⁸ These methods generally use sortases that catalyse transpeptidation reactions.⁶⁴ Target proteins are engineered to have, at both ends, the recognition sites necessary for each enzyme. In the presence of a labelled peptide, the transpeptidation reaction can then take place.

By exploiting the orthogonality of two sortase enzymes with unique reactivities (*i.e.* sortase A from *Staphylococcus aureus*, SrtA_{staph}, and sortase A from *Streptococcus pyogenes*, SrtA_{strep}), Antos *et al.*

developed a strategy for placing different chemical labels at the two termini of a protein. The approach consists of installing the first modification on the C-terminus using SrtA_{strep}, which results in the formation of a peptide sequence that cannot be cleaved by SrtA_{staph}. Following deprotection of the N-terminus of the protein using thrombin, the liberated glycine residue can then be coupled using SrtA_{staph} without disrupting the C-terminal modification, thus yielding a dually modified protein (Scheme 9).



Scheme 9: Dual labelling of proteins using orthogonal enzymes.

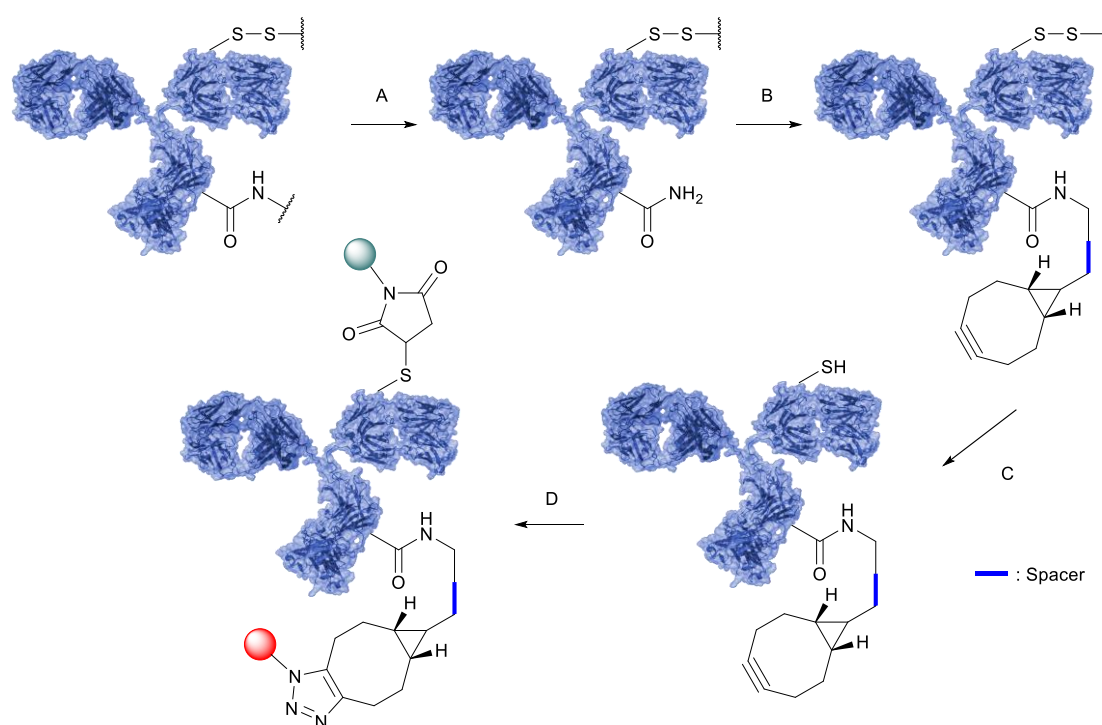
This strategy was exemplified on an eGFP mutant with five N-terminal glycines and ubiquitin carboxyl-terminal hydrolase isozyme L3 containing an N-terminal glycine. As a proof of concept, both proteins were modified in high yield and purity with two distinct fluorophores. Nevertheless, although highly specific, this approach is limited to the labelling of the terminal protein region.

3. Different amino acids

In recent years, numerous methods for the site-selective modification of antibodies have been developed.^{65, 66} By selecting two orthogonal reactions in this ever-increasing pool, Puthenveetil *et al.* generated dual-labelled, site-specific antibody and antibody-fragment conjugates.

This approach was demonstrated on a cysteine mutant of trastuzumab. The first reactive handle was installed by enzymatic conjugation using transglutaminase to couple a strained alkyne to the deglycosylated antibody. Following this, using a careful reduction/re-oxidation strategy allowed for the reaction of the liberated engineered cysteine to a maleimide-modified fluorophore. In a last step,

the strained alkyne installed in the first step was reacted with a suitable cyanine-based dye to generate the dually functionalised antibody in a homogeneous and site-selective manner (Scheme 10).⁵²



Scheme 10: Dual labelling of proteins by combining orthogonal methods. A. Deglycosylation; B. Enzymatic conjugation; C. Reduction/re-oxidation; D. Michael addition/SPAAC.

Although this method requires several steps, it is virtually compatible with all IgG1s as the incorporation of non-canonical cysteine has been shown to be compatible across antibodies and they all share the same glycosylation site.⁶⁷

ii. Non-canonical amino acids

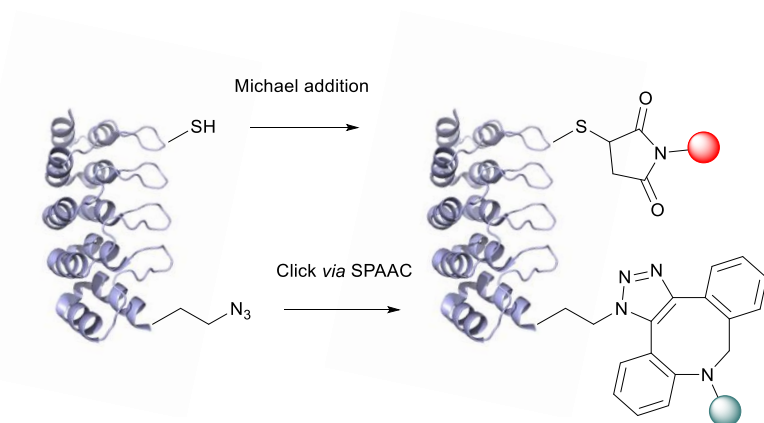
Dual functionalisation of biomolecules can be achieved by targeting different unnatural amino acids with orthogonal reactivity. The insertion of two uAAs in proteins is dependent on the availability of unnatural protein expression methods to incorporate orthogonal chemical handles for subsequent bioorthogonal reactions.^{5, 68, 69} A common strategy to introduce two mutually orthogonal chemical handles is to combine nonsense and quadruplet codon suppressions.⁷⁰⁻⁷³

Neumann *et al.* have reported the incorporation of two uAAs into a protein by synthetically evolving an orthogonal ribosome that efficiently decodes a series of quadruplet codons and the amber codon, providing several blank codons on an orthogonal mRNA, which it specifically translates. The authors created mutually orthogonal aminoacyl-tRNA synthetase-tRNA pairs and combined them with ribosome Q1 to effect incorporation of two uAAs into the protein sequence. At the same time, Wan *et al.* independently developed another method to introduce two non-natural amino acids in the same protein by using two stop codons, amber codon UAG and ochre codon UAA.⁷¹ This method was used to incorporate two fluorophores for FRET analysis of glutamate binding protein (QBP) and superfolder green fluorescent protein (GFP).

Although recombinant expression strains have been engineered to improve incorporation efficiency, this dual-labelling approach is often associated with low protein yields.⁷⁴⁻⁷⁶ The main reasons for these low yields are the competition of uAA incorporation with translational frame shifting or termination and low catalytic efficiency of the engineered aminoacyl-tRNA synthetases.⁷⁷ Thus, strategies often comprise unnatural protein expression in conjunction with the modification of canonical amino acids (*e.g.* cysteine).

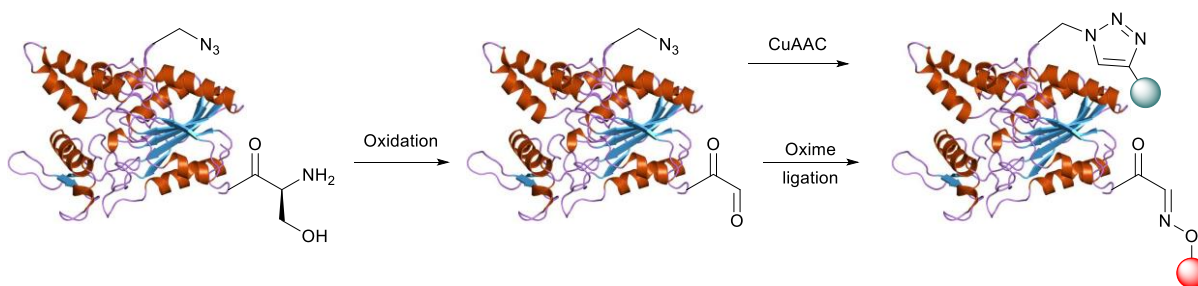
In 2007 and 2008, van Kasteren *et al.* and Brustad *et al.* paved the way for the development of dual-modification of proteins bearing a single non-canonical amino acid by successfully installing two distinct moieties in proteins using a combination of an uAA and cysteine. Van Kasteren *et al.* functionalised β -glycosidase from archaeon *Sulfolobus solfataricus* using a combination of uAA and cysteine modification. These handles were reacted further using sequential CuAAC and disulfide bond formation, respectively, to effectively mimic post-translational modifications.⁷⁸ This uAA/cysteine pair was also used by Brustad *et al.* to incorporate ketone-containing uAAs (*i.e.* *p*-acetylphenylalanine) and cysteines in a T4 lysozyme to generate mutants for the study of T4 lysozyme folding at single-molecule resolution by FRET using oxime ligation and Michael addition, respectively.⁷⁹

More recently, Simon *et al.* reported a dual-labelling approach for DARPin using a combination of uAA and cysteine functionalisation.⁵¹ Azidohomoalanine was introduced at the N-terminus of the protein and a cysteine was introduced at a second position of interest *via* site-directed mutagenesis. SPAAC reaction was utilised for orthogonal functionalisation of azidohomoalanine, followed by maleimide modification of the single cysteine residue to obtain homogeneous dually functionalised DARPin (Scheme 11).



Scheme 11: DARPin dual functionalisation using SPAAC followed by thiol modification.

To alleviate some issues associated with the incorporation and reaction of cysteines (*e.g.* disulfide bond formation, reaction reversibility, side-reactions), Mühlberg *et al.* recently described a dual-labelling strategy for proteins that combines residue-specific incorporation of uAA with chemical oxidative aldehyde formation at the N-terminus of a protein.⁴⁹ The authors selectively introduced two different functional moieties in thermophilic lipase from *Thermoanaerobacter thermohydrosulfuricus* *via* CuAAC and oxime ligation (Scheme 12). This method was applied to the conjugation of biotin and β -linked galactose residues to yield an enzymatically active thermophilic lipase.



Scheme 12: Lipase dual functionalisation using oxime ligation followed by CuAAC modification.

By engineering both azidohomoalanine and a pentapeptide that can be utilised for sortase A-mediated ligation (SML) at the C-terminus of a truncated thrombomodulin domain, Jiang *et al.* could combine SPAAC and SML to generate dual-modified proteins.⁵⁰ After optimisation of the reaction conditions, the protein of interest was functionalised with a PEG chain and a fluorophore in a one-pot/site-specific process without disrupting the activity of the recombinant thrombomodulin.

A similar approach that combines an uAA and a peptide sequence was also developed by Grünewald *et al.*⁸⁰ A pyrroline-carboxy-lysine and a dodecameric S6 peptide were engineered in various protein scaffolds and dually tagged proteins were site-selectively prepared using 2-aminobenzaldehyde-activated fluorophores and phosphopantetheinyl transferase-catalysed reaction in a one-pot process.

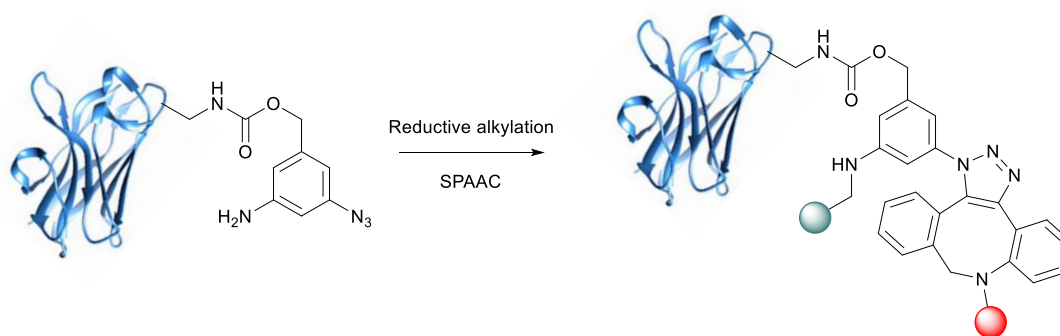
b. Approach using a multifunctional scaffold containing two reactive handles

If the two labels are not required to be on different amino acids, a more straightforward approach to achieve dual modifications on proteins is to combine the reactive handles onto a single multifunctional scaffold incorporated into the protein or that could be conjugated to proteins.¹⁹ The multifunctional tag can either contain the two desired labels⁸¹ or can be subsequently chemoselectively modified with the desired labels.

i. Endogenous multifunctional scaffold

As an alternative to the simultaneous incorporation of two unnatural amino acids, a single, well-designed uAA that possesses two orthogonal handles can also be used for direct dual functionalisation

of proteins. To this effect, Yamaguchi *et al.* recently designed a N^{ϵ} -benzyloxycarbonyl-L-lysine derivative with both amino and azido groups in the benzyl moiety.⁸² The phenylazide could undergo either alkyne–azide cycloaddition or a photo-induced reaction, while the phenylamine can undergo selective reductive alkylation under specific conditions. As a proof of concept, the bifunctional uAA was incorporated into the variable domain of a camelid heavy-chain antibody without affecting its binding properties. The resulting protein was subjected to an aldehyde linked to a 5 kDa polyethylene glycol and an azadibenzocyclooctyne linked to a fluorophore in a one-pot procedure (Scheme 13). After purification, the site-specific, dually modified conjugate was isolated in *ca.* 80% yield.



Scheme 13: Dual modification of protein using bifunctional uAA.

ii. Exogenous multifunctional scaffold

Although there are many available scaffolds that contain bifunctional linkers, there are relatively few that possess three orthogonal handles, thus enabling attachment to a protein followed by dual modification.¹⁹ Efforts have been made towards the development of such platforms but with limited success, due to the potentially complex synthesis of a branched structure with orthogonal handles.^{83, 84} Furthermore, these platforms have often been exemplified on small molecules or short peptides but rarely on proteins, which can represent an additional layer of difficulty.^{85, 86}

For example, trifunctional scaffold **1** was reacted with three small molecules (*i.e.* drug and fluorophores) and, clickable decapeptide scaffold **2** was attached to a glucose derivative and two short peptides, but no reaction was reported on proteins (Figure 2).

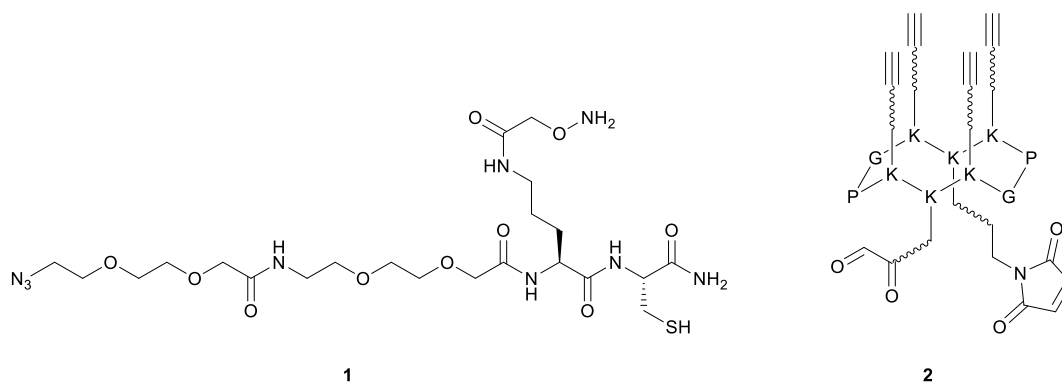
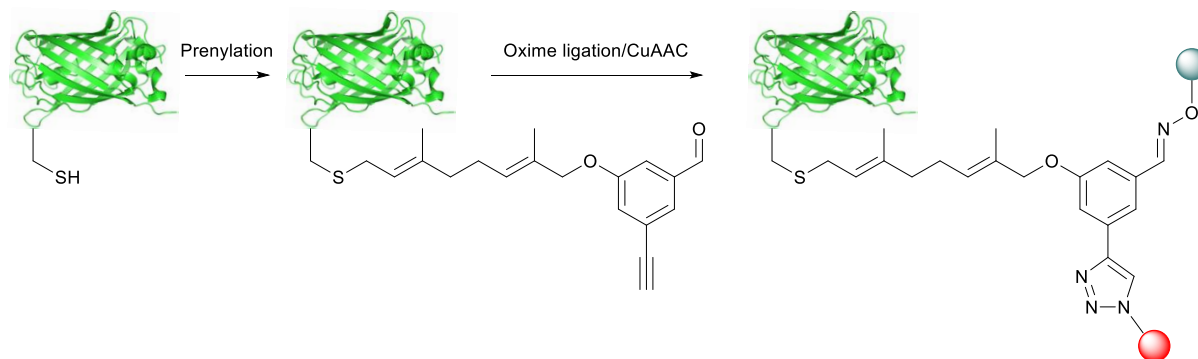


Figure 2: Multifunctional platforms with the potential to enable the dual functionalisation of proteins.

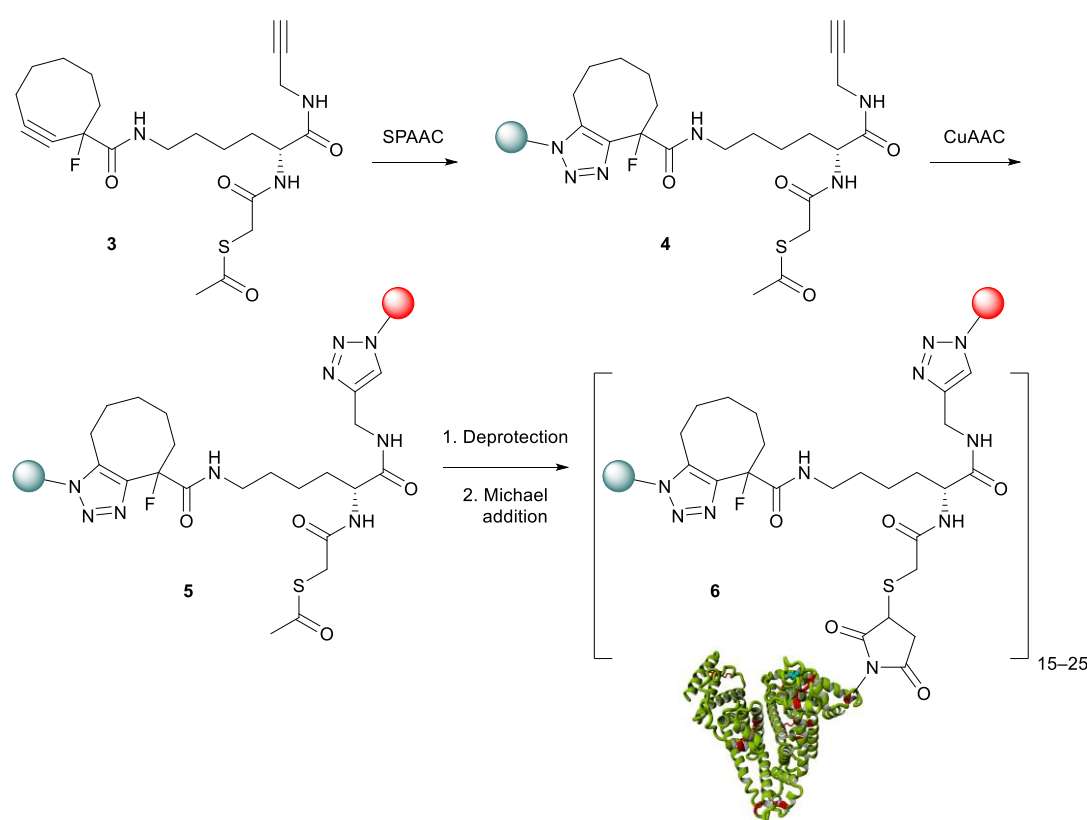
Rashidian *et al.* described a method to site-specifically prenylate proteins containing a C-terminal “CaaX-box” amino acid sequence (*i.e.* a sequence where “C” represents a cysteine, “a” an aliphatic amino acid and “X” can be M, S, Q, A or C) with various isoprenoids using Protein Farnesyl Transferase (PFTase). A triorthogonal reagent was synthesised to allow incorporation of an alkyne and an aldehyde group simultaneously into a protein.⁵³

To illustrate the capabilities of this approach, a model protein (*i.e.* GFP) and a therapeutically useful protein (*i.e.* ciliary neurotrophic factor) with engineered CaaX-box were enzymatically modified with a synthesised triorthogonal reagent that allowed incorporation of an alkyne and an aldehyde group simultaneously into the protein. Using both oxime ligation and CuAAC reaction to incorporate a series of “clickable” molecules, the authors successfully generated a series of dually modified conjugates in a one-pot procedure (Scheme 14).



Scheme 14: Dual modification using triorthogonal prenylation reagent.

Beal *et al.* also developed a heterotrifunctional template that utilises click reactions (both copper-free and copper-mediated) to effect non-enzymatic sequential biomolecule conjugations in a one-pot process. Following the preparation of substrate **3**, SPAAC followed by CuAAC was used to react its two alkynes handles and give dually modified template **5** which was purified by HPLC. Finally, **5** was reacted with a maleimide-functionalised BSA modified *via* its lysines with an average of 15 – 25 maleimides per protein to yield heterogeneous dually modified BSA **6** (Scheme 15). The scope of compatible substrates was illustrated through conjugation of peptide, sugar, lipid, fluoroalkane, biotin and fluorophore molecules.⁸⁴

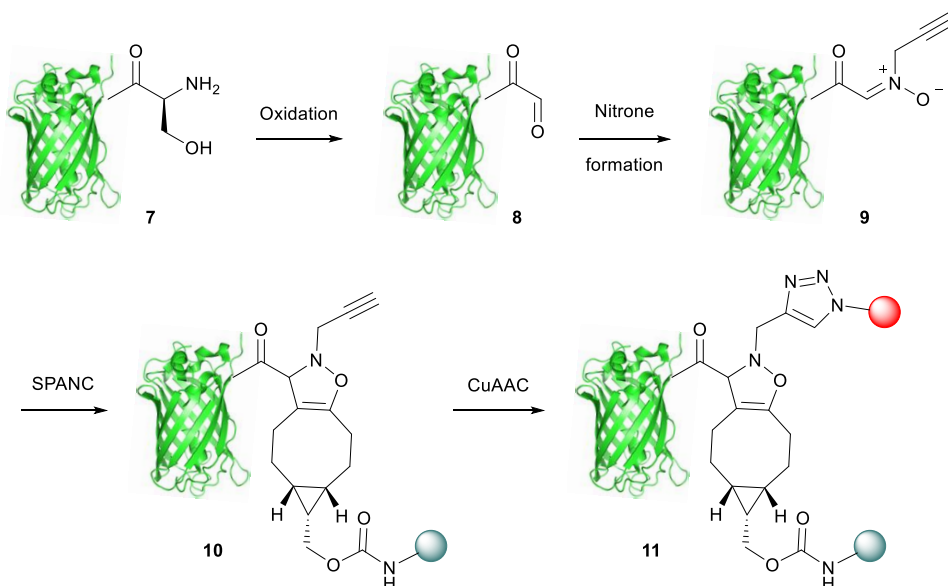


Scheme 15: Dual modification using heterotrifunctional template **3**.

Temming *et al.* described a dual-functionalisation method where the two handles were introduced directly by modifying a serine residue on the N-terminus of a peptide (*i.e.* cell-penetrating peptide human lactoferrin) or on proteins (*i.e.* enhanced GFP mutant and heat shock protein HspB2).⁴⁴ The authors rationalised that this approach potentially alleviates the need for lengthy syntheses of complex

platforms used in exogeneous modification strategies. The methods consists of a careful oxidation of the requisite N-terminal serine residue **7** with NaIO_4 as the oxidising agent followed by reaction of the resulting aldehyde **8** with a large excess of *N*-propargylhydroxylamine to yield dually “clickable” protein **9**.

Following this, the first functional label was introduced by strain-promoted alkyne–nitronc cycloaddition (SPANc) with a “clickable” BCN-biotin derivative. The terminal acetylene of **10** then served as the second anchor point for a CuAAC reaction with azido-containing fluorescein to give dually modified construct **11** (Scheme 16).

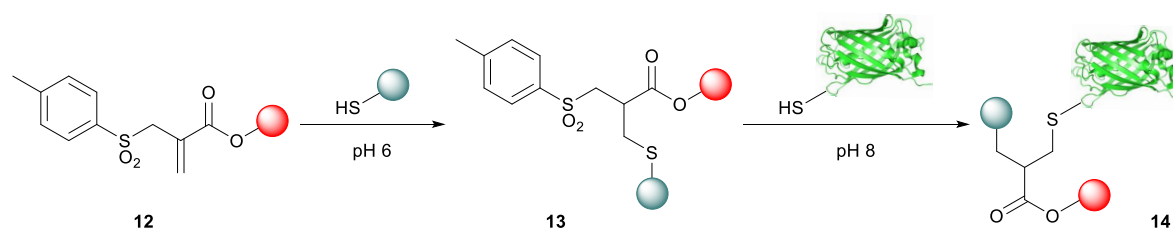


Scheme 16: Dual modification of GFP mutant using terminal serine modification.

At the same time, Munch *et al.* reported a similar site-selective three-component reaction for dual-functionalisation of peptides and proteins using a maleimide as the dipolarophile in the nitronc/alkene cycloaddition.⁸⁷

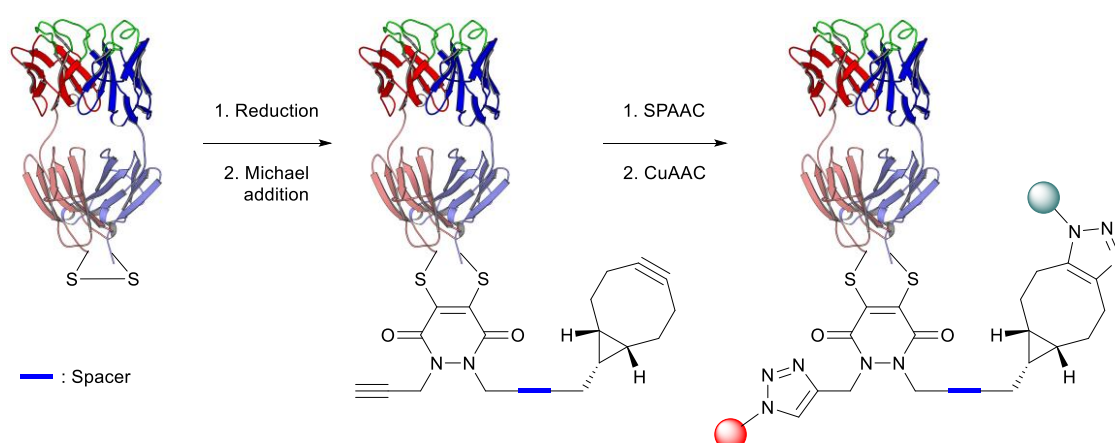
Recently, Wang *et al.* reported a pH-dependent method for the site-selective dual modification of cysteine-containing proteins and peptides based on water-soluble allyl sulfone trifunctional reagent **12**.⁸⁸ As allyl sulfones can undergo one Michael reaction with thiols at pH 6 and, after elimination, a

second one at pH 8 in a sequential fashion, bis-labelled allyl-sulfone **13** was prepared and successfully added to a recombinant cysteine-containing GFP (Scheme 17).



Scheme 17: Dual modification using water-soluble allyl sulfone trifunctional reagent.

Recently, Chudasama *et al.* reported that reagents based on pyridazinedione (PD) could be used as a platform to dually modify cysteine-containing biomolecules.^{31, 89-92} Indeed, in addition to the two cysteine-reactive centres, these compounds possess two additional points of attachment that can be functionalised. To this end, Maruani *et al.* prepared a series of PD-based platforms with orthogonal handles (*e.g.* strained alkyne, alkyne) that could easily be reacted with cysteines *via* Michael addition.^{31, 93} By the introduction of PD molecules bearing orthogonal “clickable” handles into the native disulfide bonds of an antibody and antibody fragment (*i.e.* trastuzumab and Fab fragment from trastuzumab), the authors generated a range of site-selectively dually-modified bioconjugates using two sequential click reactions (Scheme 18).



Scheme 18: Dual modification of disulfide-containing proteins using PDs.

The prepared conjugates bearing a drug and a fluorophore, or a drug and a half-life extension modality were assayed *in vitro* and demonstrated retention of both activity and stability.

4. Conclusion and outlook

In conclusion, a plethora of methods for the dual (and even tri-) modification of biomolecules in a selective manner has been developed, especially in recent years. However, further sophisticated/tailored strategies are required to overcome the remaining limitations of existing methodologies. Indeed, many challenges still remain; these include the stability and solubility of protein-expressed fragments, the harsh reaction conditions used in oligonucleotide synthesis, partial conversions or low yields, incomplete bioorthogonality, the lengthy syntheses of heterofunctional scaffolds and slow rates of reaction.

Even with further novel technologies ironing out these issues, it is our belief that it is unlikely that one particular method will prevail for the dual modification of proteins or oligomers. Instead it is envisaged that a general set of rules based on the accrued knowledge of the use of each platform will allow us to determine what is the most appropriate method to at least dually modify a certain biomolecule to enable a particular application. Thus these new technologies have a key role to play in this process whilst concurrently providing access to multi-functionalised conjugates for applications that were previously thought of as unobtainable.

5. Acknowledgements

We gratefully acknowledge EPSRC (EP/M01792X/1) and i-sense EPSRC IRC in Early Warning Sensing Systems for Infectious Diseases (EP/K031953/1) for funding AM and DAR, respectively.

References

1. K. Lang and J. W. Chin, *ACS Chem. Biol.*, 2014, **9**, 16-20.
2. N. Krall, F. P. da Cruz, O. Boutureira and G. J. Bernardes, *Nat. Chem.*, 2016, **8**, 103-113.
3. R. van Vught, R. J. Pieters and E. Breukink, *Comput. Struct. Biotechnol. J.*, 2014, **9**, e201402001.

4. T. L. Foley and M. D. Burkart, *Curr. Opin. Chem. Biol.*, 2007, **11**, 12-19.
5. J. M. Chalker, G. J. Bernardes and B. G. Davis, *Acc. Chem. Res.*, 2011, **44**, 730-741.
6. M. F. Debets, J. C. van Hest and F. P. Rutjes, *Org. Biomol. Chem.*, 2013, **11**, 6439-6455.
7. G. J. L. Bernardes, J. M. Chalker and B. G. Davis, in *Ideas in Chemistry and Molecular Sciences*, Wiley-VCH Verlag GmbH & Co. KGaA, 2010, DOI: 10.1002/9783527630516.ch3, pp. 59-91.
8. O. Boutureira and G. J. Bernardes, *Chem. Rev.*, 2015, **115**, 2174-2195.
9. S. B. Gunnoo and A. Madder, *Chembiochem*, 2016, **17**, 529-553.
10. C. Wagner and H. A. Wagenknecht, *Chemistry*, 2005, **11**, 1871-1876.
11. S. Kodama, S. Asano, T. Moriguchi, H. Sawai and K. Shinozuka, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 2685-2688.
12. P. M. Gramlich, S. Warncke, J. Gierlich and T. Carell, *Angew. Chem., Int. Ed.*, 2008, **47**, 3442-3444.
13. J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13-21.
14. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2001, **40**, 2004-2021.
15. C. S. McKay and M. G. Finn, *Chem. Biol.*, 2014, **21**, 1075-1101.
16. K. Nwe and M. W. Brechbiel, *Cancer Biother. Radiopharm.*, 2009, **24**, 289-302.
17. S. I. Presolski, V. P. Hong and M. G. Finn, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 153-162.
18. C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu and P. Wu, *Angew. Chem., Int. Ed.*, 2011, **50**, 8051-8056.
19. D. M. Beal and L. H. Jones, *Angew. Chem., Int. Ed.*, 2012, **51**, 6320-6326.
20. M. Zheng, L. Zheng, P. Zhang, J. Li and Y. Zhang, *Molecules*, 2015, **20**, 3190-3205.
21. A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388-1405.
22. I. K. Astakhova and J. Wengel, *Chemistry*, 2013, **19**, 1112-1122.
23. G. Pourceau, A. Meyer, J.-J. Vasseur and F. Morvan, *The Journal of Organic Chemistry*, 2009, **74**, 1218-1222.
24. U. Pradere, A. Brunschweiler, L. F. Gebert, M. Lucic, M. Roos and J. Hall, *Angew. Chem., Int. Ed.*, 2013, **52**, 12028-12032.
25. H. J. Musiol and L. Moroder, *Angew. Chem., Int. Ed.*, 2010, **49**, 7624-7626.
26. Y. Zhu and W. A. van der Donk, *Org. Lett.*, 2001, **3**, 1189-1192.
27. C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai and T. Ha, *Annu. Rev. Biochem.*, 2008, **77**, 51-76.

28. C. Cecconi, E. A. Shank, C. Bustamante and S. Marqusee, *Science*, 2005, **309**, 2057-2060.
29. D. Banerjee, A. P. Liu, N. R. Voss, S. L. Schmid and M. G. Finn, *ChemBiochem*, 2010, **11**, 1273-1279.
30. N. Stephanopoulos, G. J. Tong, S. C. Hsiao and M. B. Francis, *ACS nano*, 2010, **4**, 6014-6020.
31. A. Maruani, M. E. Smith, E. Miranda, K. A. Chester, V. Chudasama and S. Caddick, *Nat. Commun.*, 2015, **6**, 6645.
32. I. K. Oh, H. Mok and T. G. Park, *Bioconjugate Chem.*, 2006, **17**, 721-727.
33. E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf and J. Kjems, *Nature*, 2009, **459**, 73-76.
34. J. B. Lee, M. J. Campolongo, J. S. Kahn, Y. H. Roh, M. R. Hartman and D. Luo, *Nanoscale*, 2010, **2**, 188-197.
35. B. Juskowiak, *Anal. Bioanal. Chem.*, 2011, **399**, 3157-3176.
36. J. Schoch, M. Staudt, A. Samanta, M. Wiessler and A. Jaschke, *Bioconjugate Chem.*, 2012, **23**, 1382-1386.
37. H. Abe and E. T. Kool, *Proc. Natl Acad. Sci. U S A*, 2006, **103**, 263-268.
38. N. Jothikumar, T. L. Cromeans, V. R. Hill, X. Lu, M. D. Sobsey and D. D. Erdman, *Appl. Environ. Microbiol.*, 2005, **71**, 3131-3136.
39. D. I. Cherny, I. C. Eperon and C. R. Bagshaw, *Eur. Biophys. J.*, 2009, **38**, 395-405.
40. A. M. Jawalekar, N. Meeuwenoord, J. S. Cremers, H. S. Overkleeft, G. A. van der Marel, F. P. Rutjes and F. L. van Delft, *J. Org. Chem.*, 2008, **73**, 287-290.
41. M. Shelbourne, T. Brown, Jr., A. H. El-Sagheer and T. Brown, *Chem. Commun.*, 2012, **48**, 11184-11186.
42. T. Santner, M. Hartl, K. Bister and R. Micura, *Bioconjugate Chem.*, 2014, **25**, 188-195.
43. U. Pradere and J. Hall, *Bioconjugate Chem.*, 2016, **27**, 681-687.
44. R. P. Temming, L. Eggermont, M. B. van Eldijk, J. C. van Hest and F. L. van Delft, *Org. Biomol. Chem.*, 2013, **11**, 2772-2779.
45. V. Ratner, E. Kahana, M. Eichler and E. Haas, *Bioconjugate Chem.*, 2002, **13**, 1163-1170.
46. R. I. Nathani, P. Moody, V. Chudasama, M. E. Smith, R. J. Fitzmaurice and S. Caddick, *Chem. Sci.*, 2013, **4**, 3455-3458.
47. P. Moody, V. Chudasama, R. I. Nathani, A. Maruani, S. Martin, M. E. Smith and S. Caddick, *Chem. Commun.*, 2014, **50**, 4898-4900.
48. J. M. Antos, G. L. Chew, C. P. Guimaraes, N. C. Yoder, G. M. Grotenbreg, M. W. Popp and H. L. Ploegh, *J. Am. Chem. Soc.*, 2009, **131**, 10800-10801.

49. M. Muhlberg, M. G. Hoesl, C. Kuehne, J. Dervede, N. Budisa and C. P. Hackenberger, *Beilstein J. Org. Chem.*, 2015, **11**, 784-791.
50. R. Jiang, L. Wang, J. Weingart and X. L. Sun, *Chembiochem*, 2014, **15**, 42-46.
51. M. Simon, U. Zangemeister-Wittke and A. Pluckthun, *Bioconjugate Chem.*, 2012, **23**, 279-286.
52. S. Puthenveetil, S. Musto, F. Loganzo, L. N. Tumey, C. J. O'Donnell and E. I. Graziani, *Bioconjugate Chem.*, 2016, DOI: 10.1021/acs.bioconjchem.6b00054.
53. M. Rashidian, S. C. Kumarapperuma, K. Gabrielse, A. Fegan, C. R. Wagner and M. D. Distefano, *J. Am. Chem. Soc.*, 2013, **135**, 16388-16396.
54. C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. Van Voorhis and B. L. Pentelute, *Nat. Chem.*, 2016, **8**, 120-128.
55. A. Buchanan, V. Clementel, R. Woods, N. Harn, M. A. Bowen, W. Mo, B. Popovic, S. M. Bishop, W. Dall'Acqua, R. Minter, L. Jermutus and V. Bedian, *mAbs*, 2013, **5**, 255-262.
56. G. J. Cotton and T. W. Muir, *Chemistry & Biology*, 2000, **7**, 253-261.
57. L. De Rosa, A. L. Cortajarena, A. Romanelli, L. Regan and L. D. D'Andrea, *Org. Biomol. Chem.*, 2012, **10**, 273-280.
58. J. Y. Yang and W. Y. Yang, *J. Am. Chem. Soc.*, 2009, **131**, 11644-11645.
59. V. Muralidharan and T. W. Muir, *Nat. Methods*, 2006, **3**, 429-438.
60. S. Brenzel, M. Cebi, P. Reiss, U. Koert and H. D. Mootz, *Chembiochem*, 2009, **10**, 983-986.
61. Y. Minato, T. Ueda, A. Machiyama, I. Shimada and H. Iwai, *J. Biomol. NMR*, 2012, **53**, 191-207.
62. C. F. Becker, R. Seidel, M. Jahnz, K. Bacia, T. Niederhausen, K. Alexandrov, P. Schwille, R. S. Goody and M. Engelhard, *Chembiochem*, 2006, **7**, 891-895.
63. N. Xie, E. N. Elangwe, S. Asher and Y. G. Zheng, *Bioconjugate Chem.*, 2009, **20**, 360-366.
64. W. J. Bradshaw, A. H. Davies, C. J. Chambers, A. K. Roberts, C. C. Shone and K. R. Acharya, *FEBS J.*, 2015, **282**, 2097-2114.
65. V. Chudasama, A. Maruani and S. Caddick, *Nat. Chem.*, 2016, **8**, 114-119.
66. P. Agarwal and C. R. Bertozzi, *Bioconjugate Chem.*, 2015, **26**, 176-192.
67. S. Bhakta, H. Raab and J. R. Junutula, in *Antibody-Drug Conjugates*, ed. L. Ducry, Humana Press, Totowa, NJ, 2013, DOI: 10.1007/978-1-62703-541-5_11, pp. 189-203.
68. T. C. Lee, M. Kang, C. H. Kim, P. G. Schultz, E. Chapman and A. A. Deniz, *Chembiochem*, 2016, DOI: 10.1002/cbic.201500695.
69. L. Wang and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2004, **44**, 34-66.
70. H. Neumann, K. Wang, L. Davis, M. Garcia-Alai and J. W. Chin, *Nature*, 2010, **464**, 441-444.

71. W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell and W. R. Liu, *Angew. Chem., Int. Ed.*, 2010, **49**, 3211-3214.
72. B. Wu, Z. Wang, Y. Huang and W. R. Liu, *Chembiochem*, 2012, **13**, 1405-1408.
73. I. Nikic, T. Plass, O. Schraidt, J. Szymanski, J. A. Briggs, C. Schultz and E. A. Lemke, *Angew. Chem., Int. Ed.*, 2014, **53**, 2245-2249.
74. T. Mukai, A. Hayashi, F. Iraha, A. Sato, K. Ohtake, S. Yokoyama and K. Sakamoto, *Nucleic Acids Res.*, 2010, **38**, 8188-8195.
75. D. B. Johnson, J. Xu, Z. Shen, J. K. Takimoto, M. D. Schultz, R. J. Schmitz, Z. Xiang, J. R. Ecker, S. P. Briggs and L. Wang, *Nat. Chem. Biol.*, 2011, **7**, 779-786.
76. M. J. Lajoie, A. J. Rovner, D. B. Goodman, H. R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church and F. J. Isaacs, *Science*, 2013, **342**, 357-360.
77. S. Nehring, N. Budisa and B. Wiltschi, *PloS one*, 2012, **7**, e31992.
78. S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105-1109.
79. E. M. Brustad, E. A. Lemke, P. G. Schultz and A. A. Deniz, *J. Am. Chem. Soc.*, 2008, **130**, 17664-17665.
80. J. Grunewald, D. H. Jones, A. Brock, H. P. Chiu, B. Bursulaya, K. Ng, T. Vo, P. Patterson, T. Uno, J. Hunt, G. Spraggon and B. H. Geierstanger, *Chembiochem*, 2014, **15**, 1787-1791.
81. J. Morales-Sanfrutos, F. J. Lopez-Jaramillo, F. Hernandez-Mateo and F. Santoyo-Gonzalez, *J. Org. Chem.*, 2010, **75**, 4039-4047.
82. A. Yamaguchi, T. Matsuda, K. Ohtake, T. Yanagisawa, S. Yokoyama, Y. Fujiwara, T. Watanabe, T. Hoshaka and K. Sakamoto, *Bioconjugate Chem.*, 2016, **27**, 198-206.
83. J. A. Stewart, B. F. Piligian, S. R. Rundell and B. M. Swarts, *Chem. Commun.*, 2015, **51**, 17600-17603.
84. D. M. Beal, V. E. Albrow, G. Burslem, L. Hitchen, C. Fernandes, C. Laphorn, L. R. Roberts, M. D. Selby and L. H. Jones, *Org. Biomol. Chem.*, 2012, **10**, 548-554.
85. G. Clavé, H. Volland, M. Flaender, D. Gasparutto, A. Romieu and P. Y. Renard, *Org. Biomol. Chem.*, 2010, **8**, 4329-4345.
86. M. Galibert, O. Renaudet, P. Dumy and D. Boturyn, *Angew. Chem., Int. Ed.*, 2011, **50**, 1901-1904.
87. H. K. Munch, J. E. Rasmussen, G. Popa, J. B. Christensen and K. J. Jensen, *Chem. Commun.*, 2013, **49**, 1936-1938.
88. T. Wang, A. Riegger, M. Lamla, S. Wiese, P. Oeckl, M. Otto, Y. Wu, S. Fischer, H. Barth, S. L. Kuan and T. Weil, *Chem. Sci.*, 2016, DOI: 10.1039/C6SC00005C.
89. V. Chudasama, M. E. Smith, F. F. Schumacher, D. Papaioannou, G. Waksman, J. R. Baker and S. Caddick, *Chem. Commun.*, 2011, **47**, 8781-8783.

90. M. T. W. Lee, A. Maruani and V. Chudasama, *J. Chem. Res.*, 2016, **40**, 1-9.
91. A. Maruani, H. Savoie, F. Bryden, S. Caddick, R. Boyle and V. Chudasama, *Chem. Commun.*, 2015, **51**, 15304-15307.
92. M. T. W. Lee, A. Maruani, J. R. Baker, S. Caddick and V. Chudasama, *Chem. Sci.*, 2016, **7**, 799-802.
93. A. Maruani, S. Alom, P. Canavelli, M. T. Lee, R. E. Morgan, V. Chudasama and S. Caddick, *Chem. Commun.*, 2015, **51**, 5279-5282.