ACCOUNTS
of chemical research

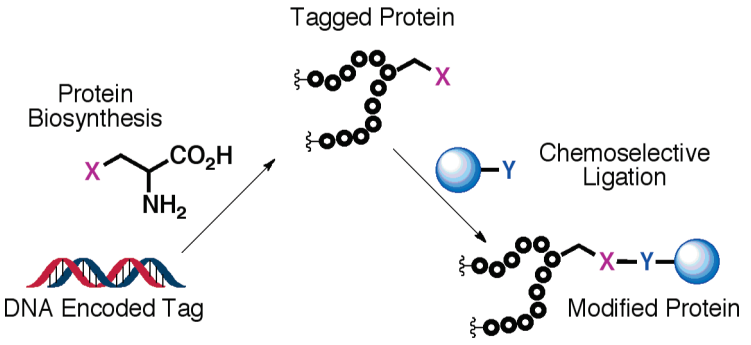
A "Tag-and-Modify" Approach to Site-Selective Protein Modification

JUSTIN M. CHALKER, GONÇALO J. L. BERNARDES, AND BENJAMIN G. DAVIS*

Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford OX1 3TA, United Kingdom

RECEIVED ON FEBRUARY 27, 2011

CONSPECTUS



Covalent modification can expand a protein's functional capacity. Fluorescent or radioactive labeling, for instance, allows imaging of a protein in real time. Labeling with an affinity probe enables isolation of target proteins and other interacting molecules. At the other end of this functional spectrum, protein structures can be naturally altered by enzymatic action. Protein–protein interactions, genetic regulation, and a range of cellular processes are under the purview of these post-translational modifications. The ability of protein chemists to install these covalent additions selectively has been critical for elucidating their roles in biology. Frequently the transformations must be applied in a site-specific manner, which demands the most selective chemistry. In this Account, we discuss the development and application of such chemistry in our laboratory. A centerpiece of our strategy is a "tag-and-modify" approach, which entails sequential installation of a uniquely reactive chemical group into the protein (the "tag") and the selective or specific modification of this group. The chemical tag can be a natural or unnatural amino acid residue.

Of the natural residues, cysteine is the most widely used as a tag. Early work in our program focused on selective disulfide formation in the synthesis of glycoproteins. For certain applications, the susceptibility of disulfides to reduction was a limitation and prompted the development of several methods for the synthesis of more stable thioether modifications. The desulfurization of disulfides and conjugate addition to dehydroalanine are two routes to these modifications. The dehydroalanine tag has since proven useful as a general precursor to many modifications after conjugate addition of various nucleophiles; phosphorylated, glycosylated, peptidylated, prenylated, and even mimics of methylated and acetylated lysine-containing proteins are all accessible from dehydroalanine.

While cysteine is a useful tag for selective modification, unnatural residues present the opportunity for bio-orthogonal chemistry. Azide-, arylhalide-, alkyne-, and alkene-containing amino acids can be incorporated into proteins genetically and can be specifically modified through various transformations. These transformations often rely on metal catalysis. The Cu-catalyzed azide–alkyne addition, Ru-catalyzed olefin metathesis, and Pd-catalyzed cross-coupling are examples of such transformations. In the course of adapting these reactions to protein modification, we learned much about the behavior of these reactions in water, and in some cases entirely new catalysts were developed. Through a combination of these bio-orthogonal transformations from the panel of tag-and-modify reactions, multiple and distinct modifications can be installed on protein surfaces. Multiple modifications are common in natural systems, and synthetic access to these proteins has enabled study of their biological role.

Throughout these investigations, much has been learned in chemistry and biology. The demands of selective protein modification have revealed many aspects of reaction mechanisms, which in turn have guided the design of reagents and catalysts that allow their successful deployment in water and in biological milieu. With this ability to modify proteins, it is now possible to interrogate biological systems with precision that was not previously possible.

Published on the Web 05/12/2011 www.pubs.acs.org/accounts
10.1021/ar200056q

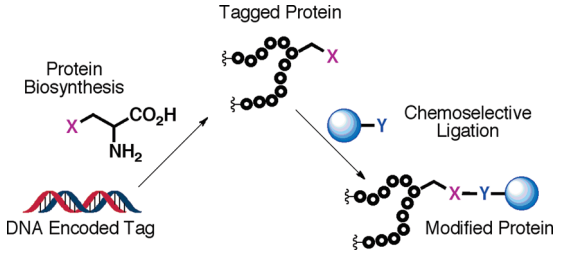
Introduction

The utility of covalent protein modification is manifold: analytical probes can be attached for protein tracking,[1] unnatural appendages such as artificial cofactors can impart novel activity,[2] and protein conjugates have gained headway in therapeutic contexts.[3] Proteins are also naturally modified after translation, dramatically expanding their functional capacity.[4] Synthetic installation of such modifications offers a way to access and study these macromolecules.[5]

There are many strategies for the chemical modification of proteins.[1,6−8] If the modified protein is to be further studied in a functional context, then a well-defined modification is essential. Achieving selective modification is often fraught with difficulty: a single residue must react in preference to hundreds of other side chains, the reaction must be efficient in buffered water, and the transformation must proceed at or near room temperature. In this Account, we describe our efforts to meet these challenges in accessing precisely modified proteins. While many biological methods are available to incorporate modified amino acids into proteins,[7,9−11] we have adopted a combined biological and chemical approach. The centerpiece of our program is a "tag-and-modify" strategy[12] that is inspired by prescient investigations of pioneers such as Koshland,[13−15] Bender,[16] Lowe,[17] Jones,[18] and Hilvert.[19] The tag-and-modify method relies on the installation of a uniquely reactive amino acid tag at a specific site in the protein and the selective chemical modification of that tag (Scheme 1).

The amino acid tag can be natural or unnatural, but the final chemical transformation must be highly selective. To meet the challenge of selectivity, we describe the development of novel reagents, catalysts, and reactions. We hope our reflection on this work will highlight the utility of these transformations and help frame the challenges and opportunities that await in the synthesis and study of modified proteins.

SCHEME 1. A Tag-and-Modify Strategy for Site-Selective Protein Modification



Protein Modification at Cysteine

Of the natural amino acids, cysteine is a convenient tag for selective modification. Cysteine is highly nucleophilic and is a classic target for selective reaction with electrophiles.[20] Cysteine is easily installed at a site of interest by site-directed mutagenesis and its relatively low natural abundance often enables the preparation of single cysteine mutants. We first describe the use of cysteine as a precursor to disulfides. We then describe how these studies led to nonclassical routes for cysteine modification: first, the conversion of disulfides to thioethers, and then the conversion of cysteine to dehydroalanine.

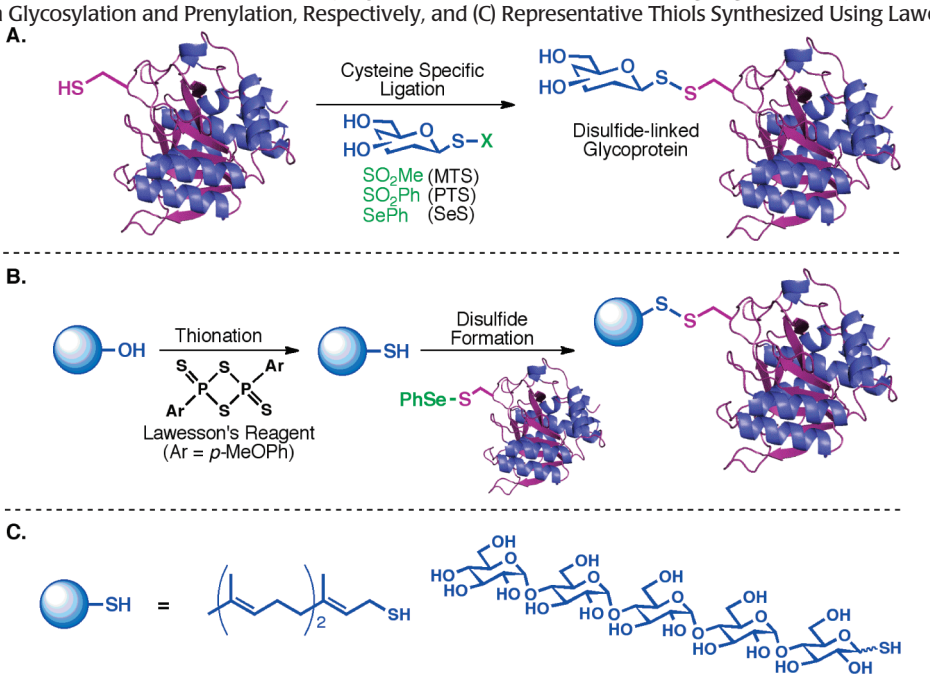
Site-Selective Disulfide Formation

We first demonstrated our tag-and-modify strategy in the construction of disulfide-linked glycoproteins. Glycoproteins mediate a variety of cellular processes[21] and access to homogeneous glycoproteins is critical for their study.[22,23] Attaching a glycan through a disulfide linkage provided an early opportunity to assess glycoprotein function.

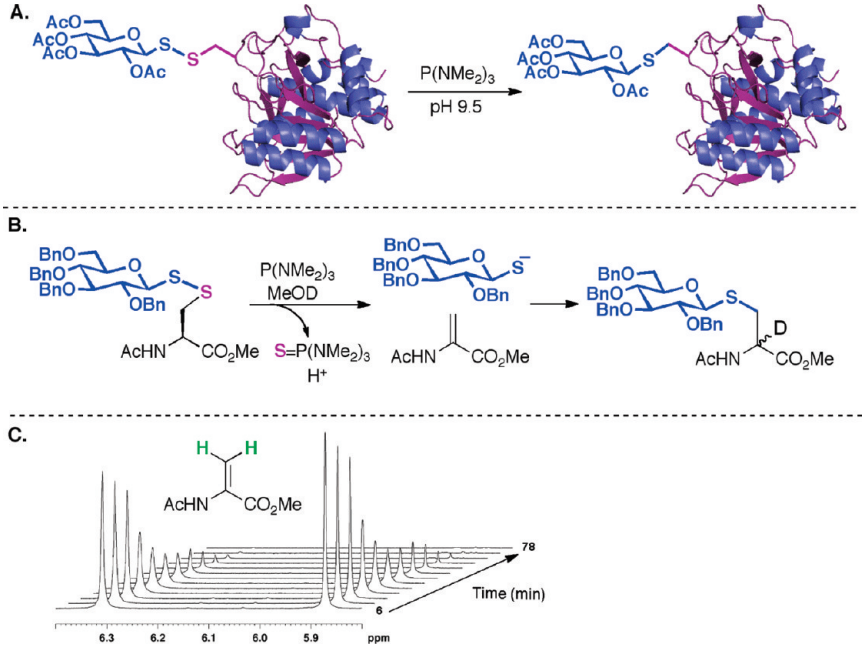
Trivially, mixing a thiol and a cysteine-tagged protein can generate a mixed disulfide after air oxidation. However, under such conditions unwanted symmetrical disulfides can form. In contrast, reagents that bear methanethiosulfonate (MTS),[24,25] phenylthiosulfonate (PTS),[26] or phenylselenenylsulfide (SeS),[27] react rapidly and specifically with cysteine to provide mixed disulfides (Scheme 2A). These transformations are nearly instantaneous, allow kinetic control of product formation, and do not require large excesses of reagents. We have taken advantage of these highly efficient reactions in a number of glycoprotein syntheses. For instance, glycodendriproteins[28] synthesized through this chemistry neutralized pathogens by degrading their adhesion proteins.[29] This chemistry was also used to synthesize a delicate trisulfide-linked glycoprotein.[30] Finally, the selenenylsulfide method allowed the conjugation of a heptasaccharide, one of the largest synthetic glycans to be attached to a protein.[27]

The ease and utility of disulfide formation inspired a method for the direct thionation of alcohols using Lawesson's reagent (Scheme 2B). While there are many methods for the synthesis of thioglycosides and glycosyl thiols, most require multistep syntheses with a heavy reliance on protecting groups.[31] The conversion of reducing sugars to glycosyl thiols using Lawesson's reagent requires no protecting group intervention.[32] In principle, this method allows the thionation of sugars isolated from natural sources and direct attachment to a protein.[32] Lawesson's reagent also provides

SCHEME 2. (A) Site-Specific Formation of Disulfide-Linked Glycoproteins, (B) Thionation of Reducing Sugars and Allylic Alcohols Using Lawesson's Reagent Enabling Protein Glycosylation and Prenylation, Respectively, and (C) Representative Thiols Synthesized Using Lawesson's Reagent



SCHEME 3. (A) Thioether-Linked Glycoprotein Synthesis by Disulfide Contraction, (B) A Model Desulfurization of a Cysteine Disulfide, and (C) ^{1}H NMR Evidence for the Dehydroalanine Intermediate

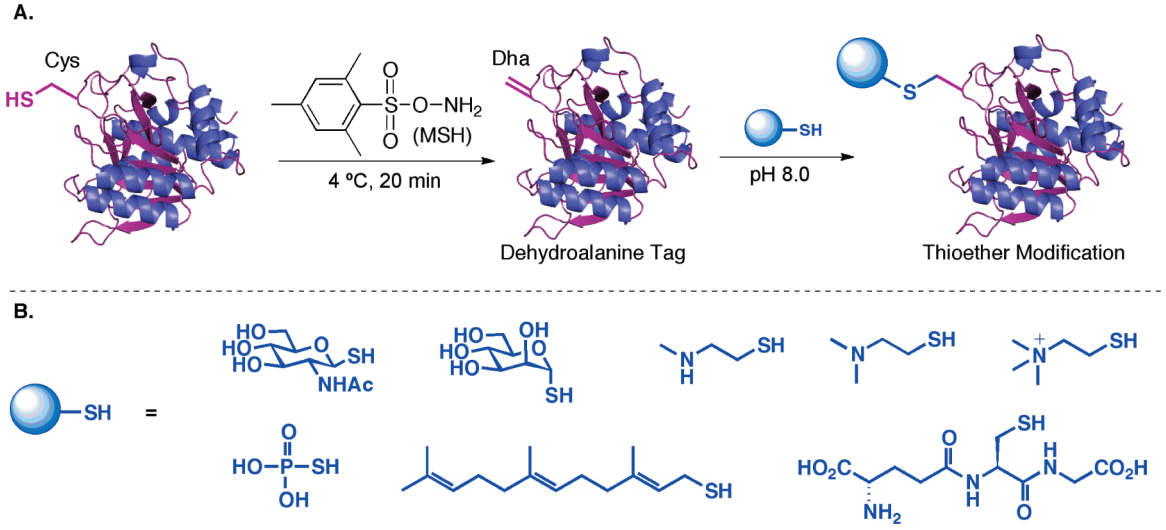


allylic thiols from the corresponding alcohols, enabling access to well-defined prenylated proteins (Scheme 2C).[33]

Converting Disulfides to Thioethers

Disulfides are easily installed and are often stable *in vivo*, but disulfides may be labile under reducing conditions. When the tethered modification is critical for the specific function of the protein, a stable linkage is essential. In some cases, one can take advantage of disulfide lability. For example, electron-rich phosphines such as hexamethylphosphorus triamide (HMPT) react with disulfide glycoconjugates to give a thioether derivative (Scheme 3A).[34] This transformation,

SCHEME 4. Conversion of Cysteine to Dehydroalanine Using *O*-Mesitylenesulfonylhydroxylamine (MSH)[a]



[a]Thiol addition to dehydroalanine enables synthesis of post-translational modifications and their mimics.

inspired by Harpp and Gleason,[35] let us combine the convenience of disulfide modification with the benefit of thioether stability.[34] This transformation, at least in some substrates, proceeds via attack of the aminophosphine at the cysteine sulfur with subsequent elimination of the thiophosphonium ion to generate dehydroalanine.[34] Addition of the liberated thiol to dehydroalanine provides the thioether (Scheme 3B,C).

This desulfurization of disulfides allowed the first formal glycosylation of cysteine on a protein.[34] Underlying mechanistic observations next led us to explore the direct conversion of cysteine to dehydroalanine. We anticipated many opportunities in protein modification with a general method for dehyroalanine incorporation.

Thioethers from Dehydroalanine

Inspired by Koshland's seminal work on dehydroalanine-containing proteins[13,15] and our disulfide contraction studies discussed above,[34] we set out to develop a general method for the conversion of cysteine to dehydroalanine. We considered that the efficient *syn*-elimination of sulfilimines[36,37] was good reason to believe that the sulfenamide of cysteine might undergo an analogous elimination. Indeed, when treated with the electrophilic nitrogen source *O*-mesitylsulfonylhydroxylamine (MSH), cysteine eliminated rapidly to dehydroalanine.[38] Application of this novel reaction to a single-cysteine-tagged protein enabled efficient incorporation of dehydroalanine (Scheme 4A).[38]

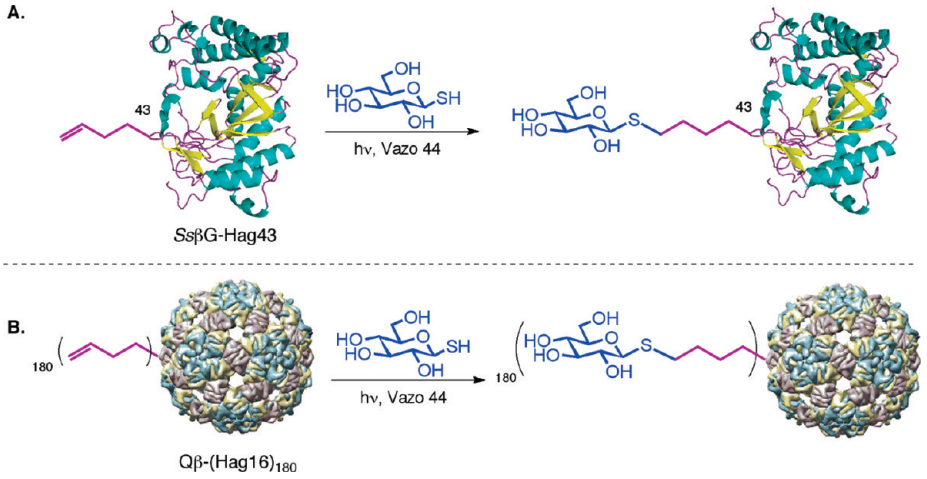
The use of dehydroalanine as a reactive tag enables access to a wide range of post-translational modifications (Scheme 4B):[38] Addition of glycosyl thiols leads to *S*-linked glycoconjugates. These structures are naturally occurring[39−41] and also serve as analogs to the more common serine-linked glycoproteins.[21,22] Conjugate addition of thiophosphate yields phosphocysteine, a natural modification.[42] We propose that phosphocysteine can also serve as an functional mimic of phosphoserine that should aid in the chemical biology of phosphorylation.[43] The naturally occurring farnesyl cysteine[44] is also accessible through dehydroalanine, as well as analogs of methylated lysine.[45] These lysine analogs are useful in the study of chromatin biology, which is regulated, in part, by lysine methylation of histone proteins.[46−48]

While methylated lysine analogs can be accessed by direct alkylation of cysteine,[46] direct glycosylation or phosphorylation of cysteine has not been reported for protein substrates. Dehydroalanine provides direct access to these structures. Addition to dehydroalanine most likely results in an epimeric mixture, but this mixture does not preclude use of these modified proteins in biology. For instance, the acquired capability to install GlcNAc cysteine residues in proteins via dehydroalanine enabled an investigation of endoglycosidase A (Endo A) in the chemoenzymatic synthesis of glycoproteins.[49] Complete conversions were observed in the Endo A-catalyzed transfer of sugar oxazoline donors to proteins displaying synthetically installed GlcNAc−cysteine, suggesting that both epimers were accepted by the enzyme.

Protein Modification at Unnatural Tags

So far our survey has focused on the unique reactivity of cysteine, a natural amino acid. There are, however, many

SCHEME 5. A Single Homoallylglycine Tag Was Reacted with Glycosyl Thiols under Free Radical Conditions in *Sulfolobus solfataricus* β-Glycosidase (A) and the Monomer of the Virus-Like Particle Qβ (B)



methods for the incorporation of unnatural amino acid tags[9−11] and their bio-orthogonal modification.[1,7,8] We now shift our focus to unnatural amino acids in the tag-and-modify strategy.

Radical Addition of Thiols to Homoallylglycine

The addition of thiols to dehydroalanine is a direct route to several post-translational modifications (Scheme 4). However, when only efficient conjugation is required, rather than post-translational mimicry, other methods allow access to *S*-linked modifications. The addition of thiyl radicals to alkene-tagged amino acids such as homoallylglycine (Hag) is one method. Hag can be incorporated into proteins as a methionine surrogate when using methionine auxotrophic bacteria in methionine-depleted media.[50,51] In this way, the Met triplet codon can be reassigned to Hag.

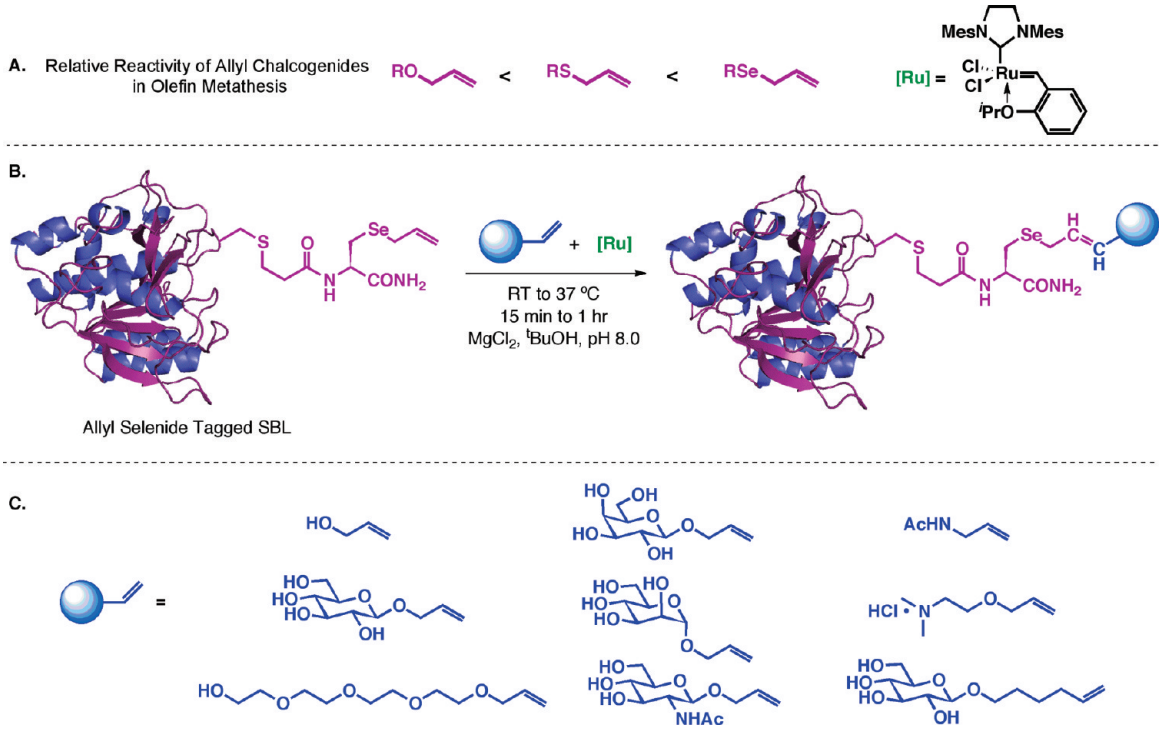
As a first model for Hag modification, a single methionine mutant of a β-glycosidase from *Sulfolobus solfataricus* (*Ss*βG) was expressed in methionine auxtrophic *Escherichia coli* supplemented with Hag. Hydrothiolation was then investigated with glycosyl thiols. This method of glycoprotein synthesis was designed to complement the related reaction of cysteine with alkenes.[52] Dual use of the water-soluble radical initiator Vazo 44 and UV light (365 nm) led to the most efficient conjugation (Scheme 5A).[53] We then applied this transformation to a virus-like particle (Qβ) that displays 180 homoallylglycine tags at predetermined sites (Scheme 5B). Hydrothiolation with a glycosyl thiol provided a multivalent glycoprotein in which all Hag residues were modified. These homogenously modified particles are of potential interest in synthetic vaccine development where statistical labeling of protein carriers with antigens is almost invariably used for immunological tests.[54] While we have contributed methods for analysis of statistically modified proteins,[55] we anticipate that that the use of well-defined vaccine constructs will allow a clear correlation between antigen copy number, linker effects, carrier structure, and the immunological outcome of vaccination.[56]

Olefin Metathesis on Proteins

Another transformation that takes advantage of an alkene tag is olefin metathesis.[57] The functional group tolerance of olefin metathesis and the prospect for carbon−carbon bond formation motivated us to adapt this reaction to protein modification. At the start of this project, however, homogeneous aqueous cross metathesis was limited to simple alkenes such as allyl alcohol.[58] Progress in aqueous olefin metathesis[59] was required for application to biomolecules.

Investigating olefin metathesis on a range of amino acids in aqueous media, one residue was particularly reactive: *S*-allyl cysteine.[60] Extending the alkene by even one methylene unit to the butenyl sulfide resulted in no cross-metathesis under the same conditions. Apparently, allyl sulfides were exceptional substrates for olefin metathesis. At first this observation was counterintuitive since prior observations demonstrated a "general incompatibility of the ruthenium-based metathesis catalysts with substrates containing sulfur-(II) donor sites."[61] Allyl sulfides, it turns out, are an important exception to this rule. Moreover, allyl sulfides are not simply compatible with ruthenium metathesis catalysts; allyl sulfides enhance the rate of metathesis. We proposed a "sulfur-relay" mechanism[60] to account for this rate enhancement whereby sulfur coordinates to ruthenium, positioning the

SCHEME 6. (A) Relative Reactivity of Allyl Chalcogenides in Olefin Metathesis, (B) Olefin Cross-Metathesis on an Allyl Selenide Tagged Protein, and (C) Cross-Metathesis Partners That React Efficiently with Allyl Selenide Tagged Proteins



allyl group proximal to the alkylidene and provoking rapid metathesis. For longer chain alkenyl sulfides, such as butenyl or pentyl sulfides, nonproductive chelation arrests catalysis.[60]

Next we benefited from the fortuitous intersection of our program in cysteine modification and our foray in aqueous olefin metathesis. *S*-Allyl cysteine could be incorporated into proteins by the addition of allyl thiol to dehydroalanine or by allylation of cysteine itself, enabling the first examples of olefin metathesis on a protein substrate.[60,62] Despite this breakthrough, only a single example in our initial studies proceeded to completion: cross-metathesis with allyl alcohol.

To address this shortcoming, the influence of sterics was investigated. Unhindered allyl sulfides allowed a broader substrate scope for protein modification.[63] It also became clear that allyl chalcogenides generally enhance the rate of olefin metathesis, with allyl selenides as the most reactive (Scheme 6A).[63−65] By installation of an unhindered allyl selenide on a protein surface, metathesis proceeded to completion for several substrates (Scheme 6B,C). The reactivity of allyl chalcogenides in olefin metathesis has motivated our current efforts to incorporate these tags biosynthetically.

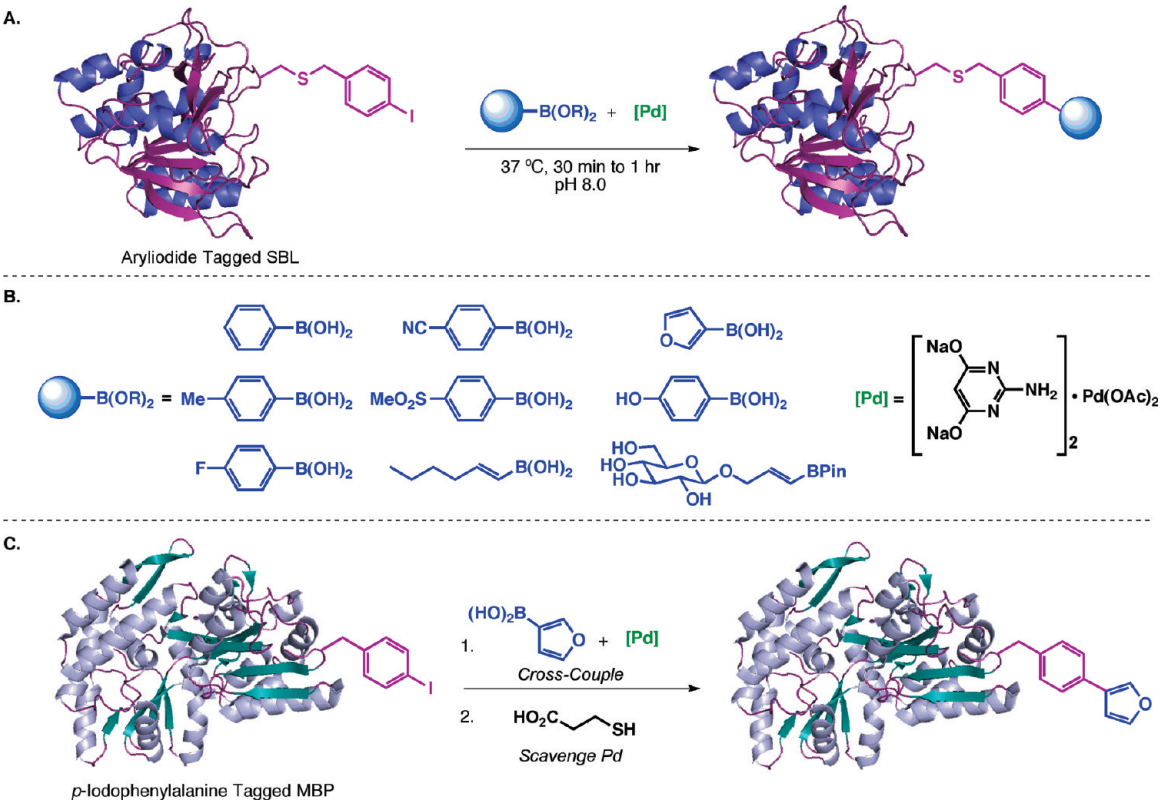
While our primary interest in olefin metathesis is its application to biomolecules,[65] the principles uncovered during these investigations are useful in general synthetic endeavors.[64] We next turn to another widespread transformation that enables carbon−carbon bond formation and discuss its implementation in the tag-and-modify strategy.

Cross-Coupling on Proteins

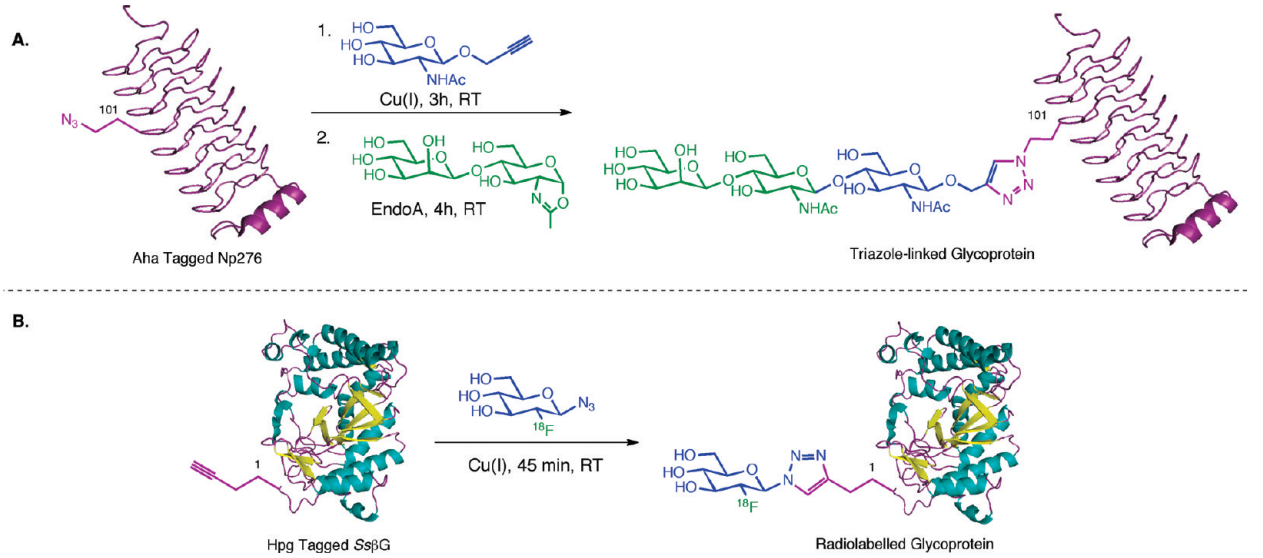
Palladium-catalyzed cross-coupling is ubiquitous in organic synthesis.[66] In particular, the Suzuki−Miyaura coupling[67] is well suited for biomolecules due to its high functional group tolerance and compatibility with water.[68] Yet, despite the ability to genetically incorporate cross-coupling partners, the challenge of cross-coupling on proteins had been met with little success since first proposed in 2000.[69] Sonogashira, Heck, and Suzuki−Miyaura reactions on proteins invariably gave low conversion.[70−72] From these reports, the need for active catalysts was apparent.

While addressing this issue, we discovered a novel catalyst for efficient Suzuki−Miyaura coupling in water (Scheme 7).[73] To validate this catalyst in protein modification, an aryl iodide was installed chemically by alkylating a single cysteine protein mutant (Scheme 7A). Rapid cross-coupling was observed on this protein with a variety of substrates (Scheme 7B). No organic solvent was used in these reactions, and full conversion was observed in all cases. The

SCHEME 7. (A) Suzuki–Miyaura Coupling on a Subtilisin Mutant (SBL) Containing a Chemically Installed Aryl Idodide, (B) Cross-Coupling Partners Used in Model Couplings on SBL, (C) The First Suzuki–Miyaura Cross-Coupling at a Genetically Encoded *p*-Iodophenylalanine in Maltose Binding Protein (MBP)

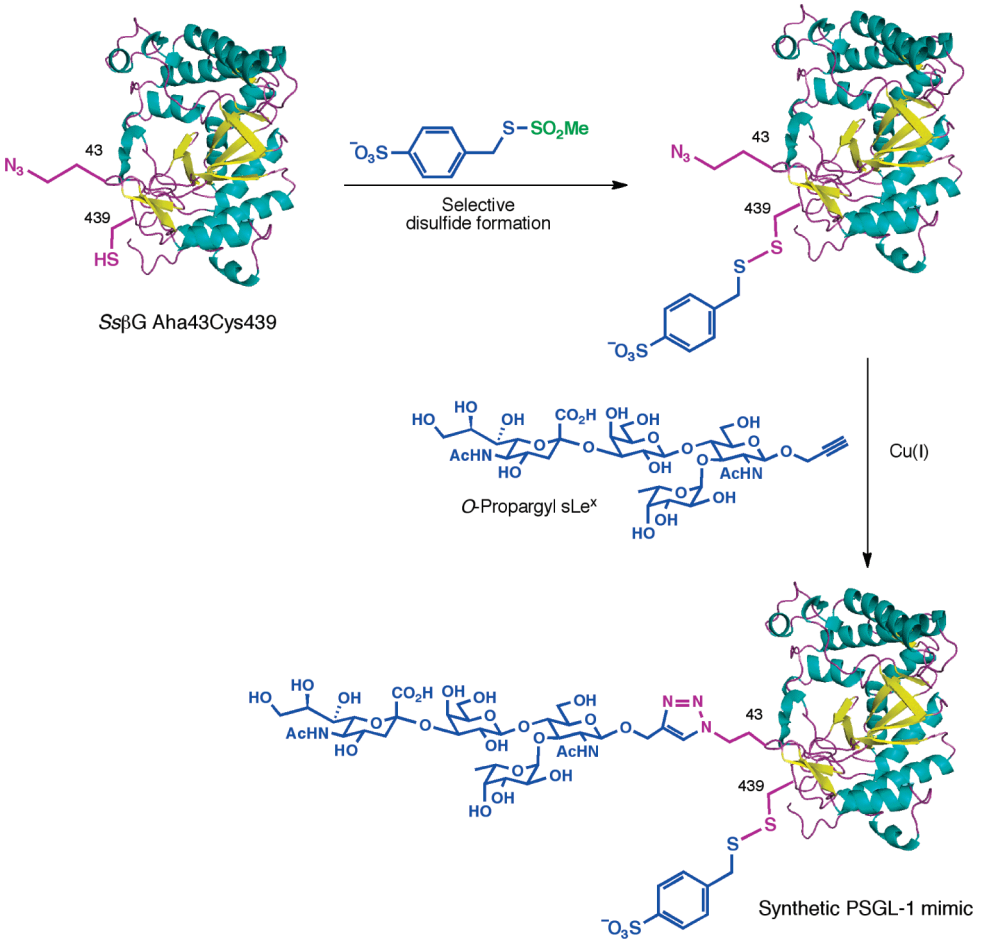


SCHEME 8. (A) The Aha Tag in Glycoprotein Synthesis on a Model Protein (Np276 from *Nostoc punctiforme*) and (B) The Hpg Tag in the Synthesis of an ^{18}F-Labeled Glycoprotein



Suzuki–Miyaura reaction is particularly useful for installing hydrophobic modifications since water solubility of these substrates is conferred by formation of the borate salt. An additional prospect for Suzuki–Miyaura coupling is fluorogenic protein labeling since biaryl side chains display unique spectral signatures.

SCHEME 9. A Dual Tag-and-Modify Process Enables Multisite Protein Modification and the Synthesis of a Functional PSGL-1 Mimic



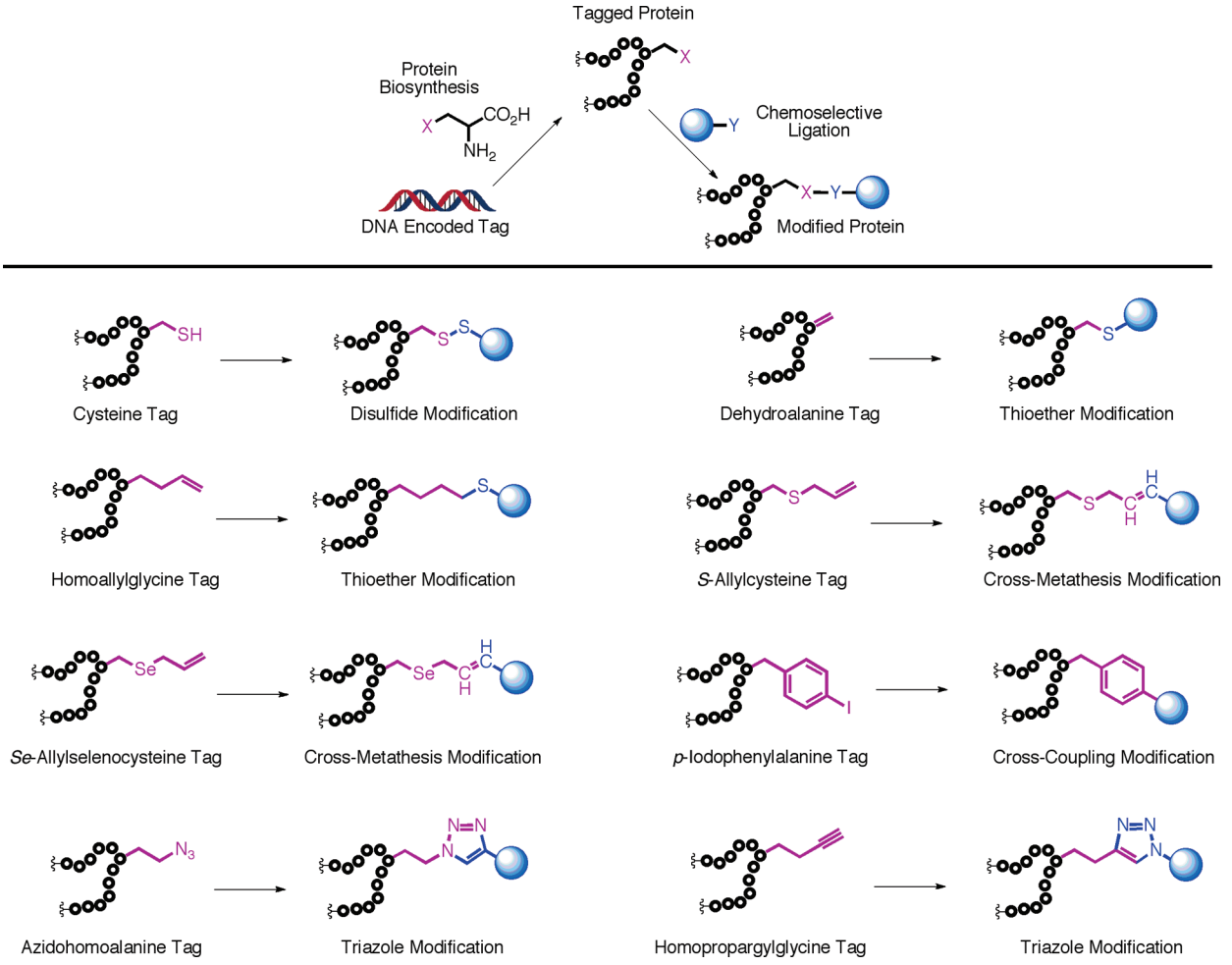
While this was the first cross-coupling to proceed to completion on a protein, the aryl iodide was installed chemically. Genetically incorporated aryl iodides, in contrast, create opportunities for *in vivo* protein labeling. Accordingly, we incorporated *p*-iodophenylalanine into maltose binding protein (MBP) in response to the amber stop codon.[74] The use of our pyrimidine–palladium catalyst enabled the first Suzuki–Miyaura coupling at a genetically incorporated aryl iodide (Scheme 7C).[75] Additionally, a method to scavenge Pd was described in this report (Scheme 7C). This scavenging may be important for metal-mediated modifications where residual metals must be removed.

Protein Modification at Azide and Alkyne Tags

Like homoallylglycine, azidohomoalanine (Aha) can be incorporated into proteins as a methionine surrogate; reassigning the Met triplet codon to an azide tag.[76] The azide is a versatile functional group for bio-orthogonal chemistry[77] and reacts efficiently with alkynes in the copper-catalyzed[78,79] or strain-promoted [3 + 2] cycloaddition.[80,81] We have included the azide in our standard panel of tags and have capitalized on efficient cycloaddition at Aha. In one example, *N*-acetylglucosamine (GlcNAc)-modified proteins were prepared by this method and used to probe the substrate tolerance of endoglycosidase A (Scheme 8A).[49] This investigation revealed that the triazole was the linker of choice for hindered sites, while a thioether-linked GlcNAc gave highest conversion for exposed sites.[49] These subtle linker effects will help guide future chemoenzymatic strategies in the synthesis of glycoproteins.

A hallmark of the azide–alkyne cycloaddition is the high reaction rate. To coordinate positron emission tomography with protein modification, we have taken advantage of these fast ligations to synthesize ^{18}F-labeled glycoproteins. Using a homopropargylglycine (Hpg) tag incorporated as a methionine surrogate (reassigning the Met codon to an alkyne tag),[51] we were able to link [^{18}F]-2-fluorodeoxyglucose to a model protein (Scheme 8B).[82] This experiment was completed in just over 2 h, demonstrating that radiolabeled glycoproteins can be synthesized in a time frame compatible with the short half-life of ^{18}F.

SCHEME 10. The "Tag-and-Modify" Approach to Protein Modification



Installing Two Distinct Modifications

The modifications discussed above were all at a single tag. Nature, in contrast, frequently installs multiple and distinct modifications on proteins.[4] To access these proteins or suitable mimics, orthogonal tags are required. As a demonstration of dual differential modification, we have synthesized a functional mimic of P-selectin-glycoprotein-ligand-1 (PSGL-1).[83] PSGL-1 naturally displays two different modifications: an *O*-linked sialyl Lewis^x (sLe^x) glycan and a sulfated tyrosine that together mediate binding to P-selectin, a process observed during the inflammatory response. To access a PSGL-1 mimic, the β-glycosidase from *Sulfolobus solfataricus* (*Ss*βG) was first engineered to display cysteine and azidohomoalanine[76] tags at approximately the same spatial relationship as the modifications in the natural PSGL-1. At the cysteine tag, a mimic of sulfated tyrosine was installed by disulfide formation. At the azide tag, sialyl Lewis^x was installed by the copper-catalyzed azide–alkyne cycloaddition (Scheme 9).[78,79] The resulting synthetic mimic of PSGL-1 showed binding to human P-selectin *in vitro* and *in vivo* at sites associated with inflammation. This binding demonstrated that functional mimicry is possible even without using the PSGL-1 protein sequence. The innate glycosidase activity of the *Ss*βG scaffold further serves as a reporter mechanism. Thus, when treated with glycosidase substrate X-Gal, the PSGL-1 mimic generates an amplified signal and allows sensitive visualization of inflammation. Importantly, because of the choice of a generic protein scaffold (the TIM barrel protein *Ss*βG), this approach, in principle, could be translated to other TIM barrel proteins to exploit their inherent functions.

The synthesis of the PSGL-1 mimic is a rare example of sequential protein modification at multiple sites. This strategy is made possible by access to multiple tags with mutually compatible reactivity. While the triazole and disulfide linkages sufficed in this context, many proteins bear multiple post-translational modifications on residues that may not be effectively mimicked through these linkages. For the study

of multiple and different post-translational modifications on a single protein, additional orthogonal tags and reactions are needed.[8] This need for expanded orthogonal chemistry is shaping our current research efforts.

Conclusion and Outlook

Over the last decade, we have explored chemistry useful in the selective modification of proteins (Scheme 10). Each method in this Account has revealed unique insights into selectivity, mechanism, and scope of the chemical modification of proteins. This panel of techniques is useful in the analysis of a wide array of post-translational modifications and shows promise in diagnostic and therapeutic arenas that require synthetic biologics.

The future of chemical protein modification is rife with opportunity. Even the most fundamental cellular processes are regulated by proteins bearing multiple modifications. A challenge awaits in both chemistry and biology to access these structures and deduce their precise role. Moreover, while functional mimics of post-translational modifications can be accessed through bio-orthogonal tags, accessing the native modification through selective chemistry is an ongoing challenge. A goal in this endeavor is to use chemistry that rivals the precision of enzymes. As these transformations develop, this chemistry should complement and expand standard genetic and biochemical techniques in protein synthesis.[84] As we pursue these goals, we will maintain a general sensitivity to both chemical synthesis and biology. Ultimately, we hope the methods and materials that are the fruits of these efforts will enable the design of protein function, shed light on the intricate biology of post-translational modifications, and provide therapeutics for outstanding problems in medicine.

BIOGRAPHICAL INFORMATION

Justin M. Chalker was born in Kansas City, Kansas, in 1983. He graduated from the University of Pittsburgh with a B.S. in Chemistry and a B.A. in the History and Philosophy of Science in 2006. Justin carried out undergraduate research with Prof. Theodore Cohen where he investigated the Zn-ene cyclization and its use in total synthesis. As a Rhodes Scholar and National Science Foundation Graduate Research Fellow, Justin is completing his Ph.D. at the University of Oxford under the direction of Prof. Benjamin G. Davis where he is exploring novel aqueous chemistry for selective protein modification.

Gonçalo J. L. Bernardes was born in Torres Vedras, Portugal, in 1980. He received his M.Chem. degree from the University of Lisbon in 2004. He then moved to the University of Oxford where he completed his D.Phil. degree in 2008 under the supervision of Prof. Benjamin G. Davis working on the synthesis of homogenously modified proteins. As a Marie-Curie Fellow, he worked with Prof. Peter H. Seeberger for 1 year, after which time he returned to Portugal to work as a senior scientist at Alfama Inc. Since the fall of 2010, Gonçalo is an EMBO and Novartis Fellow at the Institute of Pharmaceutical Sciences of ETH Zürich where he is developing novel antibody−drug conjugates in the group of Prof. Dario Neri.

Benjamin G. Davis earned his B.A. (1993) and D.Phil. (1996) from the University of Oxford. He is currently a full Professor at the University of Oxford and a Fellow at Pembroke College. His group's research centers on chemical biology with an emphasis on carbohydrates and proteins. Prof. Davis' accomplishments have been recognized by a number of awards: the Royal Society of Chemistry Meldola (1999), Corday-Morgan (2005), and Norman Heatley (2009) Medals; the Wain Medal (2008); and the Royal Society Mullard (2005) and Wolfson Research Merit (2009) Awards. Prof. Davis was also the first U.K. recipient of the American Chemical Society's Horace S. Isbell Award in 2008.

Justin M. Chalker thanks the Rhodes Trust and the National Science Foundation for funding. Gonçalo J. L. Bernardes also thanks his sources of financial support: Fundação Calouste Gulbenkian, FCT Portugal, and the International AIDS Vaccine Innitiative. Benjamin G. Davis thanks his talented and dedicated co-workers who carried out the research highlighted in this Account and his mentor and collaborator Prof. J. Bryan Jones. The following agencies are gratefully acknowledged for funding this research: the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, the International AIDS Vaccine Initiative, UCB-Celltech, the Medical Research Council, Cancer Reasearch UK, Genecor, and The Leverhulme Trust.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: ben.davis@chem.ox.ac.uk. Telephone: +44 (0) 1865 275652. Fax: +44 (0) 1865 275674.

REFERENCES

1 Sletten, E. M.; Bertozzi, C. R. Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
2 Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. Generation of new enzymes via covalent modification of existing proteins. *Chem. Rev.* **2001**, *101*, 3081–3111.
3 Doores, K. J.; Gamblin, D. P.; Davis, B. G. Exploring and exploiting the therapeutic potential of glycoconjugates. *Chem. Eur. J.* **2006**, *12*, 656–665.
4 Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem., Int. Ed.* **2005**, *44*, 7342–7372.
5 Gamblin, D. P.; van Kasteren, S. I.; Chalker, J. M.; Davis, B. G. Chemical approaches to mapping the function of post-translational modifications. *FEBS J.* **2008**, *275*, 1949–1959.
6 Hackenberger, C. P. R.; Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 10030–10074.
7 de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chem.* **2009**, *20*, 1281–1295.
8 Lim, R. K. V.; Lin, Q. Bioorthogonal chemistry: Recent progress and future directions. *Chem. Commun.* **2010**, *46*, 1589–1600.
9 Xie, J.; Schultz, P. G. A chemical toolkit for proteins - an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775–782.

10 Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr. Opin. Chem. Biol.* **2010**, *14*, 774–780.

11 Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159–167.

12 Davis, B. G. Sugars and proteins: New strategies in synthetic biology. *Pure Appl. Chem.* **2009**, *81*, 285–298.

13 Strumeyer, D. H.; White, W. N.; Koshland, D. E. Role of serine in chymotrypsin action. Conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 931–935.

14 Neet, K. E.; Koshland, D. E. The conversion of serine at the active site of subtilisin to cysteine: A "chemical mutation". *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *56*, 1606–1611.

15 Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E. The formation of anhydrochymotrypsin by removing the elements of water from the serine at the active site. *J. Am. Chem. Soc.* **1966**, *88*, 3851–3859.

16 Polgar, L.; Bender, M. L. A new enzyme containing a synthetically formed active site. Thiol-subtilisin. *J. Am. Chem. Soc.* **1966**, *88*, 3153–3154.

17 Clark, P. I.; Lowe, G. Conversion of the active-site cysteine residue of papain into a dehydroserine, a serine and a glycine residue. *Eur. J. Biochem.* **1978**, *84*, 293–299.

18 Berglund, P.; DeSantis, G.; Stabile, M. R.; Shang, X.; Gold, M.; Bott, R. R.; Graycar, T. P.; Hing Lau, T.; Mitchinson, C.; Jones, J. B. Chemical modification of cysteine mutants of subtilisin Bacillus lentus can create better catalyts than the wild-type enzyme. *J. Am. Chem. Soc.* **1997**, *119*, 5265–5266.

19 Wu, Z.-P.; Hilvert, D. Conversion of a protease into an acyl transferase: selenosubtilisin. *J. Am. Chem. Soc.* **1989**, *111*, 4513–4514.

20 Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical modification of proteins at cysteine: Opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630–640.

21 *Essentials of Glycobiology*, 2nd ed.; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.

22 Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. Glycoprotein synthesis: An update. *Chem. Rev.* **2009**, *109*, 131–163.

23 Bernardes, G. J. L.; Castagner, B.; Seeberger, P. H. Combined approaches to the synthesis and study of glycoproteins. *ACS Chem. Biol.* **2009**, *4*, 703–713.

24 Davis, B. G.; Lloyd, R. C.; Jones, J. B. Controlled site-selective glycosylation of proteins by a combined site-directed mutagenesis and chemical modification approach. *J. Org. Chem.* **1998**, *63*, 9614–9615.

25 Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. Glycomethanethiosulfonates: Powerful reagents for protein glycosylation. *Tetrahedron: Asymmetry* **2000**, *11*, 245–262.

26 Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glycosyl phenylthiosulfonates (glyco-PTS): novel reagents for glycoprotein synthesis. *Org. Biomol. Chem.* **2003**, *1*, 3642–3644.

27 Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glyco-SeS: Selenenylsulfide-mediated protein glycoconjugation-a new strategy in post-translational modification. *Angew. Chem. Int. Ed.* **2004**, *43*, 828–833.

28 Davis, B. G. The controlled glycosylation of a protein with a bivalent glycan: Towards a new class of glycoconjugates, glycodendriproteins. *Chem. Commun.* **2001**, 351–352.

29 Rendle, P. M.; Seger, A.; Rodrigues, J.; Oldham, N. J.; Bott, R. R.; Jones, J. B.; Cowan, M. M.; Davis, B. G. Glycodendriproteins: A synthetic glycoprotein mimic enzyme with branched sugar-display potently inhibits bacterial aggregation. *J. Am. Chem. Soc.* **2004**, *126*, 4750–4751.

30 Bernardes, G. J. L.; Marston, J. P.; Batsanov, A. S.; Howard, J. A. K.; Davis, B. G. A trisulfide-linked glycoprotein. *Chem. Commun.* **2007**, 3145–3147.

31 Pachamuthu, K.; Schmidt, R. R. Synthetic routes to thiooligosaccharides and thioglycopeptides. *Chem. Rev.* **2006**, *106*, 160–187.

32 Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G. The direct formation of glycosyl thiols from reducing sugars allows one-pot protein glycoconjugation. *Angew. Chem. Int. Ed.* **2006**, *45*, 4007–4011.

33 Gamblin, D. P.; van Kasteren, S.; Bernardes, G. J. L.; Chalker, J. M.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Chemical site-selective prenylation of proteins. *Mol. BioSyst.* **2008**, *4*, 558–561.

34 Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. From disulfide- to thioether-linked glycoproteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244–2247.

35 Harpp, D. N.; Gleason, J. G.; Snyder, J. P. Organic sulfur chemistry. I. The disulfide-phosphine reaction. Desulfurization with tris(diethylamino)phosphine. *J. Am. Chem. Soc.* **1968**, *90*, 4181–4182.

36 Franek, W.; Claus, P. K. Alkyl- und arylsulfensäureamide durch cycloeliminierung von alkenen aus sulfimiden, 3: Bildung primärer und sekundärer aliphatischer sulfenamide des typs $RSNH_2$ bzw. $RSNHR^1$. *Monatsh. Chem.* **1990**, *121*, 539–547.

37 Matsuo, J.; Kozai, T.; Ishibashi, H. Mild preparation of alkenes from phenyl sulfides: One-pot elimination of phenylthio group via sulfilimine at ambient temperature. *Org. Lett.* **2006**, *8*, 6095–6098.

38 Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc.* **2008**, *130*, 5052–5053.

39 Weiss, J. B.; Lote, C. J.; Bobinski, H. New low molecular weight glycopeptide containing triglucosylcyseteine in human erythrocyte membrane. *Nat. New Biol.* **1971**, *234*, 25–26.

40 Lote, C. J.; Weiss, J. B. Identification of digalactosylcysteine in a glycopeptide isolated from urine by a new preparative technique. *FEBS Lett.* **1971**, *16*, 81–85.

41 Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nat. Chem. Biol.* **2011**, *7*, 78–80.

42 Stadtman, T. C. Emerging awareness of the critical roles of S-phospho-cysteine and selenophosphate in biological systems. *BioFactors* **1994**, *4*, 181–185.

43 Tarrant, M. K.; Cole, P. A. The chemical biology of protein phosphorylation. *Annu. Rev. Biochem.* **2009**, *78*, 797–825.

44 Zhang, F. L.; Casey, P. J. Protein prenylation: Molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **1996**, *65*, 241–269.

45 Martin, C.; Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838–849.

46 Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003–1012.

47 Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 6399–6401.

48 Chatterjee, C.; Muir, T. W. Chemical approaches for studying histone modifications. *J. Biol. Chem.* **2010**, *285*, 11045–11050.

49 Fernández-González, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G. Site-selective chemoenzymatic construction of synthetic glycoproteins using endoglycosidases. *Chem. Sci.* **2010**, *1*, 709–715.

50 van Hest, J. C. M.; Tirrell, D. A. Efficient introduction of alkene functionality into proteins in vivo. *FEBS Lett.* **1998**, *428*, 68–70.

51 van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **2000**, *122*, 1282–1288.

52 Dondoni, A.; Massi, A.; Nanni, P.; Roda, A. A new ligation strategy for peptide and protein glycosylation: photoinduced thiol-ene coupling. *Chem. Eur. J.* **2009**, *15*, 11444–11449.

53 Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. Thiyl glycosylation of olefinic proteins: S-linked glycoconjugate synthesis. *Angew. Chem. Int. Ed.* **2009**, *48*, 7798–7802.

54 *Carbohydrate-Based Vaccines*; Roy, R., Ed.; ACS Symposium Series 989; American Chemical Society: Washington, DC, 2008.

55 Patel, M. K.; Vijayakrishnan, B.; Koeppe, J. R.; Chalker, J. M.; Doores, K. J.; Davis, B. G. Analysis of the dispersity in carbohydrate loading of synthetic glycoproteins using MALDI-TOF mass spectrometry. *Chem. Commun.* **2010**, *46*, 9119–9121.

56 Grayson, E. J.; Bernardes, G. J. L.; Chalker, J. M.; Boutureira, O.; Koeppe, J. R.; Davis, B. G. A coordinated synthesis and conjugation strategy for the preparation of homogeneous glycoconjugate vaccine candidates. *Angew. Chem. Int. Ed.* **2011**, *50*, 4127–4132.

57 Grubbs, R. H. *Handbook of Metathesis*; Wiley-VCH: Weinheim, Germany, 2003.

58 Hong, S. H.; Grubbs, R. H. Highly active water-soluble olefin metathesis catalyst. *J. Am. Chem. Soc.* **2006**, *128*, 3508–3509.

59 Burtscher, D.; Grela, K. Aqueous olefin metathesis. *Angew. Chem. Int. Ed.* **2009**, *48*, 442–454.

60 Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. Allyl sulfides are privileged substrates in aqueous cross-metathesis: application to site-selective protein modification. *J. Am. Chem. Soc.* **2008**, *130*, 9642–9643.

61 Fürstner, A.; Seidel, G.; Kindler, N. Macrocycles by ring-closing-metathesis, XI: Synthesis of (R)-(+)-lasiodiplodin, zeranol and truncated salicylihalamides. *Tetrahedron* **1999**, *55*, 8215–8230.

62 Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. Enabling olefin metathesis on proteins: Chemical methods for installation of S-allyl cysteine. *Chem. Commun.* **2009**, 3714–3716.

63 Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin cross-metathesis on proteins: Investigation of allylic chalcogen effects and guiding principles in metathesis partner selection. *J. Am. Chem. Soc.* **2010**, *132*, 16805–16811.

64 Lin, Y. A.; Davis, B. G. The allylic chalcogen effect in olefin metathesis. *Beilstein J. Org. Chem.* **2010**, *6*, 1219–1228.

65 Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin metathesis for site-selective protein modification. *ChemBioChem* **2009**, *10*, 959–969.

66 *Handbook of Organopalladium Chemistry for Organic Synthesis*; Negishi, E.-i., Ed.; Wiley Interscience: New York, 2002.

67 Miyaura, N.; Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **1995**, *95*, 2457–2483.

68 Polshettiwar, V.; Decottignies, A.; Len, C.; Fihri, A. Suzuki-Miyaura cross-coupling reactions in aqueous media: Green and sustainable syntheses of biaryls. *ChemSusChem* **2010**, *3*, 502–522.

69 Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS Lett.* **2000**, *467*, 37–40.

70 Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *ChemBioChem* **2006**, *7*, 134–139.

71 Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. Site-specific functionalization of proteins by organopalladium reactions. *ChemBioChem* **2007**, *8*, 232–238.

72 Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220–8223.

73 Chalker, J. M.; Wood, C. S. C.; Davis, B. G. A convenient catalyst for aqueous and protein Suzuki−Miyaura cross-coupling. *J. Am. Chem. Soc.* **2009**, *131*, 16346–16347.

74 Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotechnol.* **2004**, *22*, 1297–1301.

75 Spicer, C. D.; Davis, B. G. Palladium-mediated site-selective Suzuki-Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun.* **2011**, *47*, 1698–1700.

76 Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U. S.A.* **2002**, *99*, 19–24.

77 Debets, M. F.; van der Doelen, C. W. J.; Rutjes, F. P. J. T.; van Delft, F. L. Azide: A unique dipole for metal-free bioorthogonal ligations. *ChemBioChem* **2010**, *11*, 1168–1184.

78 Tornoe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–3064.

79 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.

80 Wittig, G.; Krebs, A. Zur existenz niedergliedriger cycloaline, I. *Chem. Ber.* **1961**, *94*, 3260–3275.

81 Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

82 Boutureira, O.; D'Hooge, F.; Fernández-González, M.; Bernardes, G. J. L.; Sánchez-Navarro, M.; Koeppe, J. R.; Davis, B. G. Fluoroglycoproteins: Ready chemical site-selective incorporation of fluorosugars into proteins. *Chem. Commun.* **2010**, *46*, 8142–8144.

83 van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature* **2007**, *446*, 1105–1109.

84 Chalker, J. M.; Davis, B. G. Chemical mutagenesis: Selective post-expression interconversion of protein amino acid residues. *Curr. Opin. Chem. Biol.* **2010**, *14*, 781–789.