the FEBS Journal

REVIEW ARTICLE

Chemical approaches to mapping the function of post-translational modifications

David P. Gamblin, Sander I. van Kasteren, Justin M. Chalker and Benjamin G. Davis

Chemistry Research Laboratory, Department of Chemistry, University of Oxford, UK

Keywords
chemoselective ligation; post-translational modification; protein glycosylation; protein modification; synthetic proteins

Correspondence
B. G. Davis, Chemistry Research Laboratory, 12 Mansfield Road, Oxford OX1 3TA, UK
Fax: 44 (0) 1865 285 002
Tel: 44 (0) 1865 275652
E-mail: ben.davis@chem.ox.ac.uk
Website: http://www.chem.ox.ac.uk/researchguide/bgdavis.html

Note
Taken in part from Young Investigator Award lecture delivered to the MPSA 2006 meeting in Lille

(Received 18 July 2007, revised 10 February 2008, accepted 21 February 2008)

doi:10.1111/j.1742-4658.2008.06347.x

Strategies for the chemical construction of synthetic proteins with precisely positioned post-translational modifications or their mimics offer a powerful method for dissecting the complexity of functional protein alteration and the associated complexity of proteomes.

Introduction

Post-translational modifications (PTMs) of proteins modulate protein activity and greatly expand the diversity and complexity of their biological function. The ubiquity of PTMs is reflected in their widespread roles in signaling, protein folding, localization, enzyme activation, and protein stability [1–3]. Indeed, the prevalence of such modifications in higher organisms, such as humans, is a leading candidate for the origin of such complex biological functions [4], which may arise from a comparatively restricted genetic code [5–7]. As a consequence of the lack of direct genetic control of their biosynthesis, natural PTMs vary in site and level of incorporation, leading to mixtures of modified proteins that may differ in function. In order to fully dissect the biological role of PTMs and determine precise structure–activity relationships, access to pure protein derivatives is essential. One approach is to exploit the fine control that may be offered by chemistry [4]. A combination of chemical, enzymatic and biological augmentation strategies can provide a modification process that occurs with the chemoselectivity and regioselectivity that is often lacking in the natural production of post-translationally modified proteins [8]. This allows the construction not only of post-translationally

Abbreviations
EPL, expressed protein ligation; glycoMTS, glycosyl methanethiosulfonates; glycoSeS, selenenylsulfide-mediated glycosylation; MTS, methanethiosulfonates; NCL, native chemical ligation; PTM, post-translational modification; SBL, subtilisin *Bacillus lentus*.

modified proteins but also of their mimics [4,9,10]. The chemical motif introduced should thus be sufficiently similar to the natural modification to mimic its function; varying this chemical appendage presents the opportunity for imparting different or enhanced biological activity.

Among PTMs, protein glycosylation is the most prevalent and diverse [11,12]. The glycans on proteins play key roles in expression and folding [13], thermal and proteolytic stability [14], and cellular differentiation [15]. Carbohydrate-bearing proteins also serve as cell surface markers in communication events such as microbial invasion [16], inflammation [17], and immune response [11,12]. The study of these events is taxing, as the biosynthesis of glycoproteins is not template driven. This results in the formation of so-called 'glycoforms' [11,12], proteins with the same peptide backbone that differ in the nature and site of glycan incorporation. Ready access to homogeneous glycoforms is hampered by inadequate separation technology that has afforded homogeneous glycoproteins only in rare instances [18]. The limited availability of singular glycoforms has prompted a concerted effort to develop new methods for their synthesis [8].

Biological methods to obtain glycoproteins

The natural expression of glycoproteins is highly dependent on the host cell glycosylation machinery. However, the re-engineering of the glycosylation pathway in the yeast *Pichia pastoris* has resulted in near-homogeneous expression [19–23], although, at present, this method lacks flexibility and non-natural variants are not tolerated. The examples of pure glycans displayed on recombinant proteins are therefore limited, thus far, to only a few structures such as the biantennary structure $GlcNAc_2Man_5GlcNAc_2$ [20] and its extended variants $Gal_2GlcNAc_2Man_3GlcNAc_2$ [19] and $Sia_2Gal_2GlcNAc_2Man_3GlcNAc_3$ [21].

An alternative approach exploits 'misacylated' tRNAs in codon suppression read-through techniques to produce homogeneous glycoproteins [24]. *In vivo* evolution of a tRNA synthetase–tRNA pair from *Methanococcus jannaschii* capable of accepting and loading glycosylated amino acids has allowed the introduction of *O*-β-D-GlcNAc-L-Ser [25] and *O*-α-D-GalNAc-L-Thr [26] into proteins with efficiencies of 96% and ~ 40% respectively.

In addition to expression-based approaches, biocatalytic methods can allow the so-called remodeling of modifications such as glycosylation. Endoglycosidases and glycosyltransferases have been used to modify existing glycoforms, e.g. in the creation of a single unnatural glycoform of enzyme RNaseB [27] catalyzed by the glycoprotein endoglycosidase enzyme endo A using novel synthetic oxazoline oligosaccharide reagents [28,29].

The above solely biological methods offer great potential. However, despite the impressive results listed above, these strategies may be limited by the often stringent specificity of natural catalytic machinery in a way that can limit their versatility and general application to modified protein (glycoprotein) synthesis.

Chemical strategies in glycoprotein synthesis

The chemical attachment of glycans offers an alternative, pragmatic route to homogeneous glycoproteins. Chemical methods can be divided into two complementary strategies [4] (Fig. 1): linear assembly, such as the introduction of a well-defined modified peptide (glycopeptide) into a larger peptide backbone; and convergent assembly, such as chemoselective ligation of a modification (glycoside) to a side chain in an intact protein scaffold. These terms reflect not only the linearity or convergence of the chemical steps that may lead to a given synthetic protein, but also the structural strategy that links the (linear) segments of the protein backbone or (convergently) attachs components/modifications to this backbone (typically to residue side chains) with little or no alteration of the backbone itself.

In linear assembly, small modified peptides (glycopeptides and glycoamino acids) can be ligated to other peptide fragments. Linear assembly methods include the use of native chemical ligation (NCL) [30], which has been applied to form, for example, unmodified protein barnase [31] and a poly(ethylene glycol)-modified variant of erythropoeitin (EPO) [32]. More recently, the use of expressed protein ligation (EPL) has provided access to larger peptide fragments. Macmillan *et al.* have used EPL to construct three well-defined model GlyCAM-1 glycoproteins [33], the first reported modular total synthesis of a biologically relevant glycoprotein. The immediate compatibility of NCL and EPL methods has led to their widespread adoption. Other methods, however, also provide emerging alternatives, such as traceless Staudinger peptide [34] ligation and protease-mediated peptide ligation [35,36].

Not withstanding these clear demonstrations of the utility of linear ligation assembly, a convergent chemoselective approach can offer the key advantages of more ready and flexible modification of a well-defined protein structure. While also developing novel methods

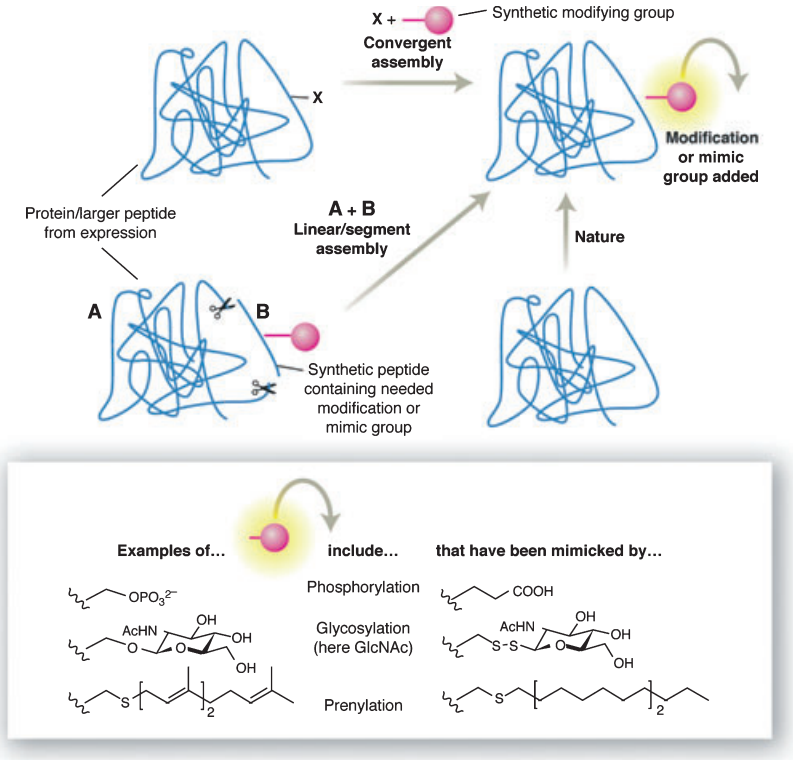

Fig. 1. Two complementary chemical strategies for mimicking PTM. Taken from [4].

for linear assembly [36], it is this convergent strategy that we have typically adopted in our own efforts in the synthesis and study of precisely modified proteins. The central strategic concept behind this convergent chemical protein modification (glycosylation) is one of 'tag and modify' (Fig. 2): the introduction of a tag into the protein backbone followed by chemoselective modification of that tag. This allows for greater flexibility in choice of protein, carbohydrate and modification (glycosylation) site.

With the relatively low abundance and unique reactivity profile of cysteine, S-linked chemical modifications are attractive targets for selective, well-defined PTM mimicry. In protein glycosylation, surface-exposed cysteine residues can be alkylated [37–39] or converted to the corresponding disulfide [40]. Furthermore, when it is used in combination with site-directed mutagenesis [41,42], glycans of choice can be introduced at any predetermined site. First-generation disulfide-forming reagents such as glycosyl methanethiosulfonates (glycoMTS) or phenylthiosulfonates provided reliable access to homogeneous glycoproteins with high efficiency [41,43]. These allowed the first examples of the systematic modulation of enzyme activity [amidase and esterase activity of the serine protease subtilisin *Bacillus lentus* (SBL)] and demonstrated not only precise glycosylation but also the dependence of activity on the exact site and identity of the disulfide-linked glycan [44].

Interestingly, judicious site selection for incorporation of a desired PTM revealed the dramatic effects of 'polar patch' modifications [45,46]. Precisely introduced charged modifications converted the protease SBL into an improved biocatalyst in peptide ligation. Particularly striking was the broad substrate tolerance that could be engineered (e.g. towards non-natural amino acids) by appropriate incorporation of the polar domain [47]. In an example that combines the exploration of two modes of modification, 'polar patch'-modified enzymes have also been applied to the catalysis of glycan-modified glycopeptide ligation [36].

Our early success using glycoMTS-mediated protein glycosylation along with a rich history of modifications using MTS reagents [48] highlighted the method as a general tool in protein modification, and we have since used this chemistry in a variety of site-selective 'tag and modify' reactions, reliably incorporating desired functionality or PTM. For instance, a library of 'catalytic antagonists' was engineered for affinity proteolysis by incorporation of a variety of ligands onto protease SBL, including examples of natural PTMs such as biotinylation and D-mannosylation (Fig. 3) [49]. The pendant ligands allowed SBL to selectively bind a protein target or partner and, by virtue of proximity,

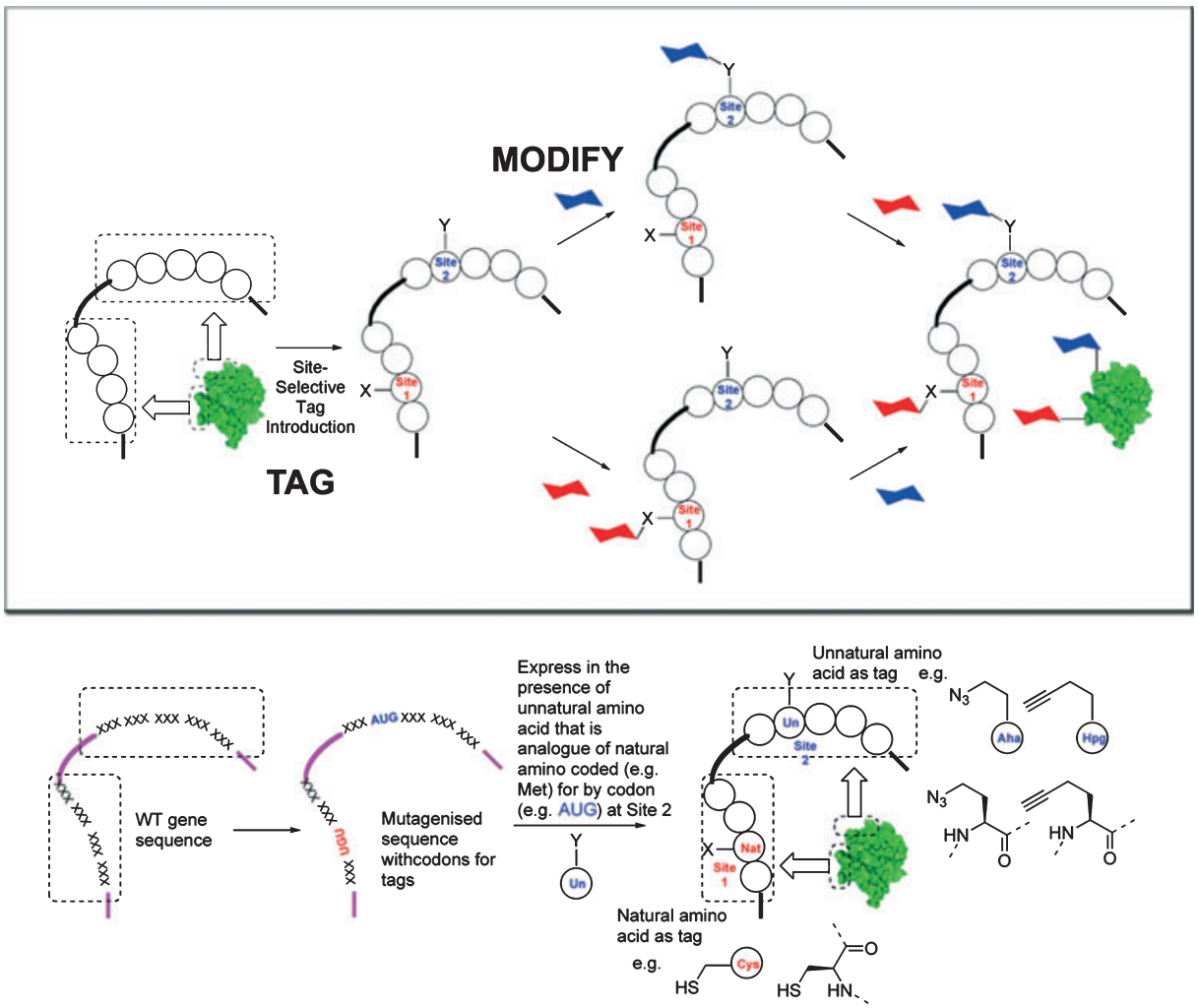

Fig. 2. The 'tag and modify' strategy behind convergent modification, illustrated here for dual tag and dual modify. Taken from [10].

catalyze enhanced hydrolytic degradation of the target protein.

More recently, the glycoMTS method has allowed the synthesis of the first examples of a homogeneous protein bearing symmetrically branched multivalent glycans [50,51]. This new class of glycoconjugate, the 'glycodendriprotein', exists in two-arm, three-arm or four-arm variants tipped with sugars. These are designed to mimic the branching levels in complex N-glycans, which come in bi-antennary, tri-antennary and tetra-antennary form. For example, the synthesized divalent, trivalent and tetravalent D-galactosyl-tipped glycodendriproteins effectively mimicked glycoproteins with branched sugar displays, as indicated by a high level of competitive inhibition of the coaggregation between the pathogen *Actinomyces naeslundii* and its copathogen *Streptococcus oralis*. This inhibition, when coupled with targeted pathogen degradation, offers therapeutic potential for the treatment of opportunistic pathogens [50,51].

This 'tag and modify' two-step approach has proved a widely successful strategy for site-selective glycosylation, used by several groups. For example, Flitsch *et al.* have employed glycosyliodoacetimides to site-selectively modify erythropoietin [52]. A similar strategy has been reported by Withers *et al.* where glycosyliodoacetimides were used in conjunction with site-selective modification of the protein endoxylanase from *Bacillus circulans* (Bcx) [53]. A protected thiol-containing sugar was conjugated and then chemically exposed before enzymatic extension. Boons *et al.* have used aerial oxidation and disulfide exchange to form homogeneous disulfide-linked glycoproteins via a cysteine mutation in the Fc region of IgG_1 [42,54].

More recently, second-generation thiol-selective protein glycosylation reagents that rely upon selenenyl-

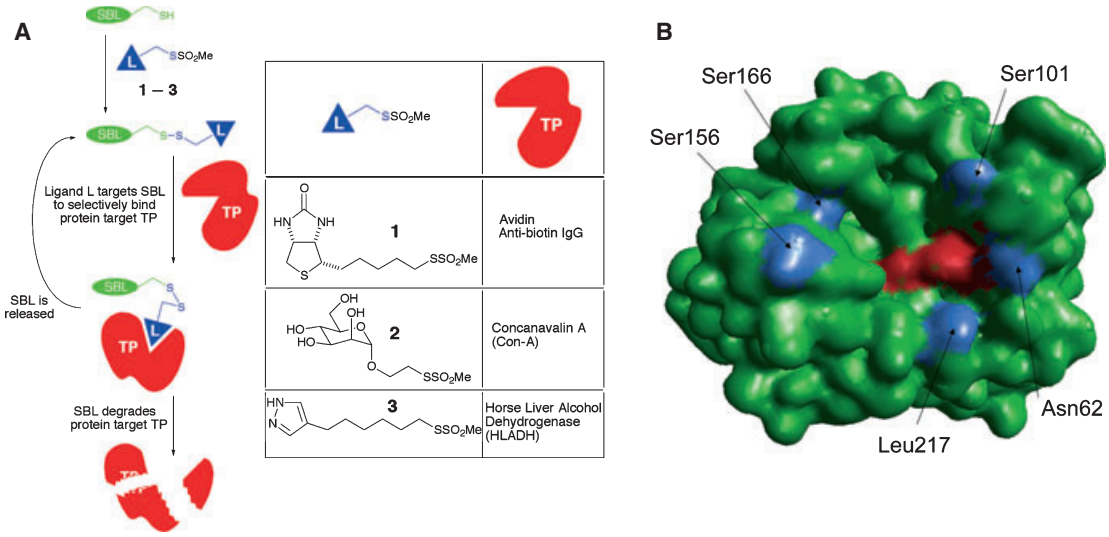

Fig. 3. (A) The use of a thiol 'tag and modify' strategy allowed site-selective attachment of natural PTMs such as biotin (1) and D-mannose (2) that, in turn, acted as 'homing' ligands for affinity proteolysis of target PTM-binding proteins. (B) A ring of modification sites (blue) around the active site (red) of the modified protease was explored. Taken from [49].

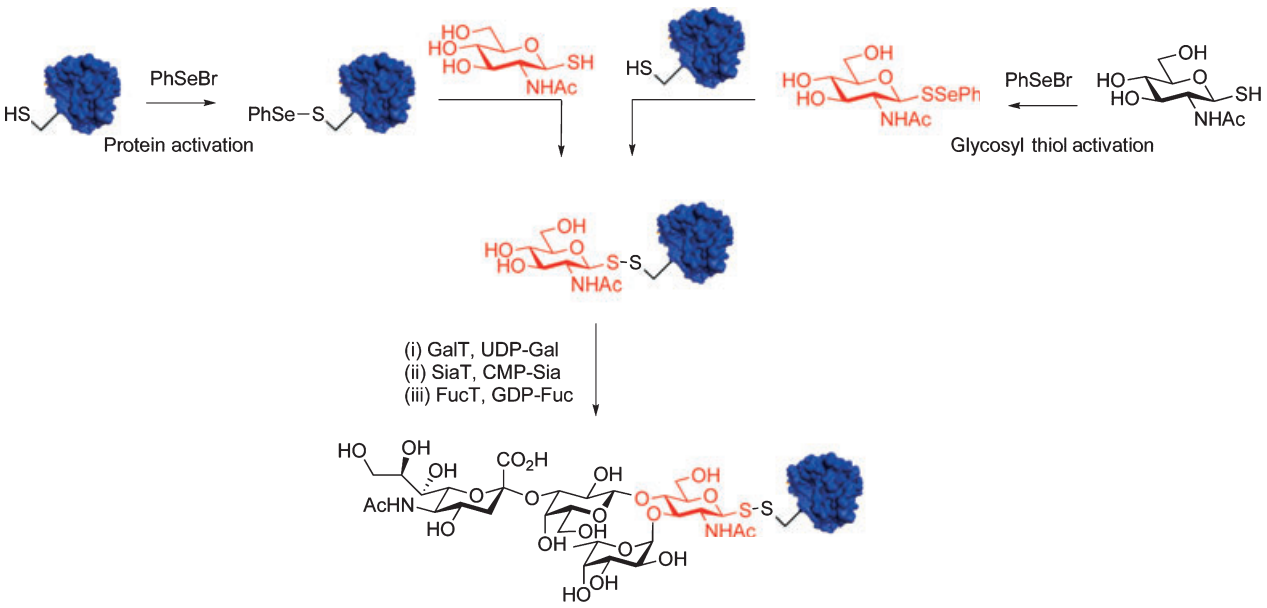

Fig. 4. Two complementary routes in glyco-SeS: protein activation and glycosyl thiol activation. The disulfide-linked glycoproteins were then readily processed in on-protein transformations catalyzed by glycosyltransferases, leading to, for example, a sialyl Lewis^X-tetrasaccharide glycan.

sulfide-mediated glycosylation (glycoSeS) have greatly improved the efficiency of 'tag and modify' methods [55]. In this approach, cysteine-containing proteins and glycosyl thiols combine through phenyl selenenylsulfide intermediates (Fig. 4). Preactivation of either the cysteine mutant protein or thiosugar is possible following exposure to PhSeBr.

GlycoSeS was initially demonstrated on simple cysteine-containing peptides, and then shown to be successful on a variety of different proteins, highlighting its versatility for glycosylation in a variety of protein environments. This high-yielding procedure also provided the first example of multisite-selective glycosylation with the same glycan and the coupling of a

heptasaccharide. Importantly, the reaction proceeds to completion using, in some cases, as little as one equivalent of glycosylating reagent. This is a great improvement on the sometimes greater than 1000 molar equivalents used in standard protein modification chemistry [8]. Furthermore, the disulfide-linked glycoprotein was readily processed by glycosyltransferases, as demonstrated by the enzymatic β-1,4-galactosylation of an *N*-acetylglucosaminyl-modified SBL protein. Recently, we have managed to further extend this disaccharide using additional glycosyltransferases to create, for example, sialyl Lewis^X-tetrasaccharide on the surface of the protein. Quantitative conversions can be obtained for the chemical glycosylation and each of these subsequent enzymatic glycosylations, leading ultimately to one pure glycoform being detected after chemical modification and each of three successive enzymatic extensions. This maintenance of purity compares favorably with enzymatic extensions performed on other natural and unnaturally linked glycoproteins [35,56]. We have also demonstrated enzymatic extensions on complex-type and branched oligosaccharides in synthetic glycoproteins.

Many of the above methods depend on a ready source of glycosyl thiol. To aid their preparation from natural sources, we have recently developed a novel direct thionation reaction for both protected and unprotected reducing sugars [57]. This allows the direct synthesis of glycosyl thiols from naturally sourced, unprotected glycans, which can then can be attached using glycoSeS to proteins in a one-pot protein glycosylation method [55]. Thus, natural sugars can be stripped from a natural protein and reinstalled site-selectively into an alternative protein scaffold of choice.

To further explore the potential of selenenylsulfide-mediated ligation in creating post-translationally modified proteins, we have mimicked protein prenylation (Fig. 5). The attachment of prenyl moieties to protein scaffolds is required for the correct function of the modified protein [58], either as a mediator of membrane association or as a determinant for specific protein–protein interactions [59,60]. Furthermore, such prenylated proteins have been shown to play crucial roles in many cellular processes, such as signal transduction [61], intracellular trafficking [62,63], and cytoskeletal structure alterations [64]. In order to fully probe and access well-defined prenylated proteins, we have recently developed a novel thionation reaction for the direct conversion of prenyl alcohols to the corresponding thiol, thereby allowing direct compatibility with selenenylsulfide protein conjugation (D. P. Gamblin, S. I. van Kasteren, G. J. L. Bernardes, N. J. Oldham, A. J. Fairbanks & B. G. Davis, manuscript in preparation). These preliminary results not only represent the first examples of site-selective protein lipidation, but also demonstrate the dramatic effect of prenylation upon the physical properties of the protein.

The construction of disulfide-linked post-translationally modified protein mimics has also been used to explore dynamic regulatory PTMs such as tyrosine phosphorylation [65,66] and glutathionation [67,68]. In all cases, the post-translationally modified protein mimics displayed native biological responses in, for example, antibody screening, highlighting the use of chemistry to further adapt and enhance protein function.

Dual differential modification

In nature, modified proteins such as glycoproteins often carry more than one distinct glycan on their surface. In order to access dual, differentially modified

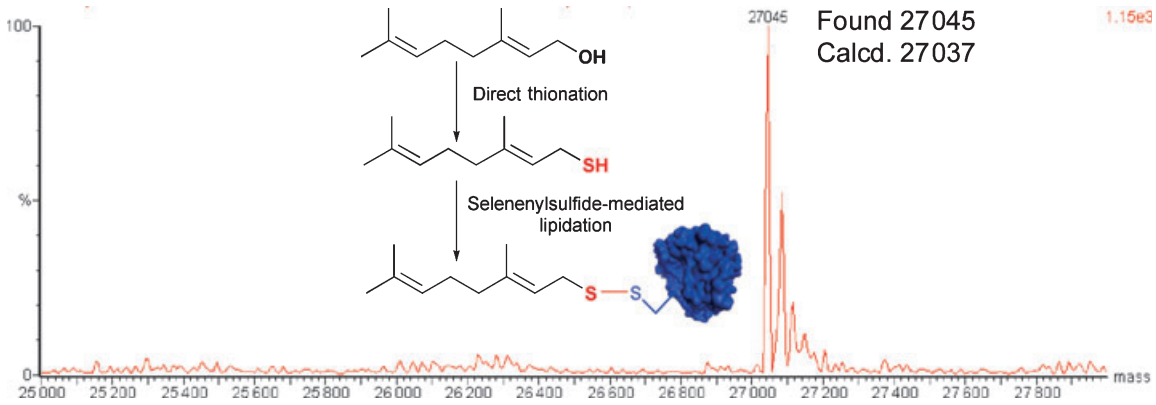

Fig. 5. A novel thionation reaction allows for the first examples of site-selective chemical protein prenylation.

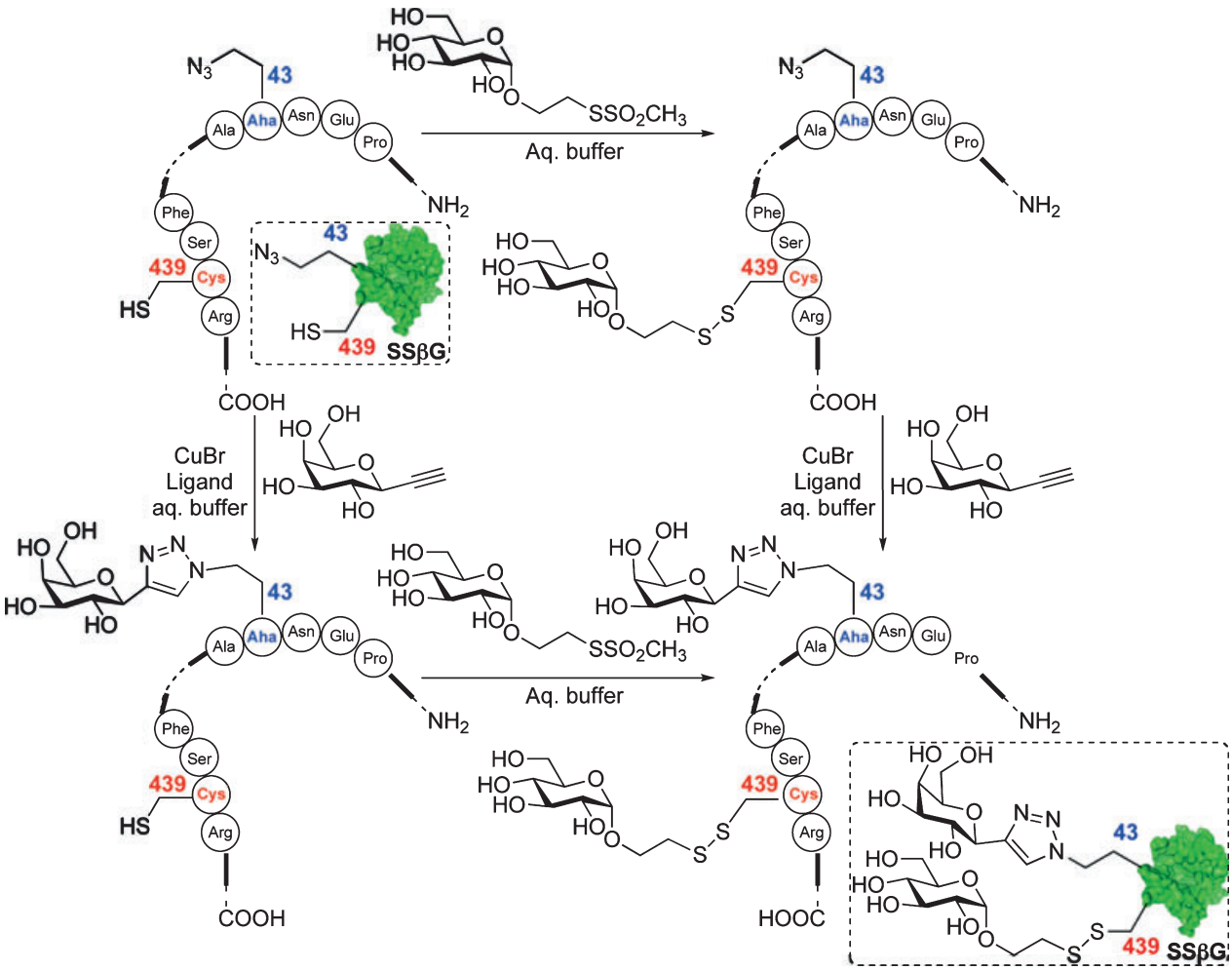

Fig. 6. The use of orthogonal chemoselective strategies allows for multisite-selective differential protein glycosylation. Taken from [10].

proteins, orthogonal methodologies are required. A strategy based on a combination of site-directed mutagenesis, unnatural amino acid incorporation, a copper(I)-catalyzed Huisgen cycloaddition [69,70] and MTS reagents has successfully been used in the first syntheses of doubly modified glycoproteins (Figs 2 and 6) [10].

The chemical protein tags were introduced through site-directed mutagenesis and incorporation of either azido- or alkyne-containing residues through methionine replacement in an auxotrophic *Escherichia coli* strain [71,72]. Treatment of these unnatural residues with either propargylic or azido glycosides, respectively, provided triazole-linked glycoproteins. This double modification strategy was used to mimic a putative glycoprotein domain of human Tamm–Horsfall protein, which carries two glycans, and the introduction of two glycans onto a galactosidase (lacZ) reporter protein. In all cases, the proteins maintained native function as well as being endowed with additional lectin-binding properties. The two methods of modification, although employing different chemistries, may be used in a complementary manner. They are also mutually compatible (orthogonal), allowing the chemistry to be performed in either order. The disulfide formation method is more rapid than the cycloaddition method, but under optimized conditions, both allow complete conversion in a matter of hours.

As a demonstration of the biological relevance, this methodology was used to model the P-selectin-binding domain of the mucin-like glycoprotein PSGL-1 [73,74]. This ligand is involved in the initial homing of leukocytes to sites of inflammation [73,74]. The binding of PSGL-1 to P-selectin is largely due to two PTMs, namely an *O*-glycan that contains tetrasaccharide sialyl-Lewis^X, and a sulfated tyrosine [74]. By careful selection of the amino acid residue accessibility and

inter-residue distance on the lacZ-reporter protein, the PSGL-1 binding domain was imitated after modification with a copper(I)-catalyzed Huisgen cycloaddition-reactive sialyl Lewis^X sugar and an MTS sulfonate as a mimic of the tyrosine sulfate. Binding of this PSGL-1 mimic to human P-selectin was shown by ELISA. This PSGL mimic also retained its inherent galactosidase activity. This dual-function, synthetic protein is therefore an effective P-selectin ligand, while simultaneously serving as a lacZ-like reporter. This mimic, named PSGL-lacZ, was subsequently used for the monitoring of acute and chronic inflammation in mammalian brain tissue both *in vitro* and *in vivo*, including in the detection of cerebral malaria.

Retooling of this reporter system also allowed systematic investigation of the role of GlcNAc-ylation as a potentially important and emerging protein PTM process [75]. Using a synthetic glycoprotein reporter GlcNAc–lacZ, specific binding was detected with the mouse innate immunity protein DC-SIGN-R2. This synthetic protein probe also selectively bound to the nuclei of a neuron subpopulation, with no binding to the nuclei of glial cells. This result suggests that neurons display selective GlcNAc-binding proteins, an intriguing result in the light of previous work on the proposed role of GlcNAc regarding both the nuclear localization of Alzheimer's-associated protein Tau [76] and nuclear shuttling in *Aplysia* neurons [77]. This work also illustrates that synthetic protein probes can be highly effective in a manner that is complementary to other protein probes such as monoclonal antibodies.

Future directions

Over the last few years, chemical protein glycosylation has become a powerful tool for accessing and studying the roles of single glycoforms [8]. In order to mimic nature's full arsenal of PTMs, the development of additional mutually orthogonal strategies is needed. This may require targeting traditionally ignored residues and application of transformations common to organic synthesis but not yet amenable to protein modification. As new methodologies emerge, the study of other PTMs and that of regulatory PTMs will hopefully provide a powerful tool for shedding light on certain key processes *in vivo* and perhaps on one of the origins of biological complexity itself.

References

1 Walsh CT (2006) *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. Roberts and Co., Englewood, CO.
2 Walsh CT, Garneau-Tsodikova S & Gatto GJ Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chemie Int Edn* **44**, 7342–7372.
3 Wold F (1981) In vivo chemical modification of proteins (post-translational modification). *Annu Rev Biochem* **50**, 783–814.
4 Davis BG (2004) Mimicking posttranslational modifications of proteins. *Science* **303**, 480–482.
5 Mirsky AE & Ris H (1951) The desoxyribonucleic acid content of animal cells and its evolutionary significance. *J Gen Physiol* **34**, 451–462.
6 Thomas CA (1971) Genetic organization of chromosomes. *Annu Rev Genet* **5**, 237.
7 Petrov DA, Sangster TA, Johnston JS, Hartl DL & Shaw KL (2000) Evidence for DNA loss as a determinant of genome size. *Science* **287**, 1060–1062.
8 Davis BG (2002) Synthesis of glycoproteins. *Chem Rev* **102**, 579–601.
9 Simon MD, Chu F, Racki LR, de la Cruz CC, Burlingame AL, Panning B, Narlikar GJ & Shokat KM (2007) The site-specific installation of methyllysine analogs into recombinant histones. *Cell* **128**, 1003–1012.
10 van Kasteren SI, Kramer HB, Jensen HH, Campbell SJ, Kirkpatrick J, Oldham NJ, Anthony DC & Davis BG (2007) Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature* **446**, 1105–1109.
11 Dwek RA (1996) Glycobiology: toward understanding the function of sugars. *Chem Rev* **96**, 683–720.
12 Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97–130.
13 Parodi AJ (2000) Protein glucosylation and its role in protein folding. *Annu Rev Biochem* **69**, 69–93.
14 Opdenakker G, Rudd PM, Ponting CP & Dwek RA (1993) Concepts and principles of glycobiology. *FASEB J* **7**, 1330–1337.
15 Lau KS, Partridge EA, Grigorian A, Silvescu CI, Reinhold VN, Demetriou M & Dennis JW (2007) Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* **129**, 123–134.
16 Comstock LE & Kasper DL (2006) Bacterial glycans: key mediators of diverse host immune responses. *Cell* **126**, 847–850.
17 Lasky LA (1995) Selectin–carbohydrate interactions and the initiation of the inflammatory response. *Annu Rev Biochem* **64**, 113–139.
18 Rudd PM, Joao HC, Coghill E, Fiten P, Saunders MR, Opdenakker G & Dwek RA (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* **33**, 17–22.
19 Bobrowicz P, Davidson RC, Li H, Potgieter TI, Nett JH, Hamilton SR, Stadheim TA, Miele RG, Bobrowicz

B, Mitchell T *et al*. (2004) Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* **14**, 757–766.
20 Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA, Wischnewski H *et al*. (2003) Production of complex human glycoproteins in yeast. *Science* **301**, 1244–1246.
21 Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, Bobrowicz P, Stadheim TA, Li H, Choi B-K *et al*. (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**, 1441–1443.
22 Li H, Sethuraman N, Stadheim TA, Zha D, Prinz B, Ballew N, Bobrowicz P, Choi B-K, Cook WJ, Cukan M *et al*. (2006) Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nat Biotechnol* **24**, 210–215.
23 Choi B-K, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH, Li H, Miele RG, Nett JH, Wildt S & Gerngross TU (2003) Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Proc Natl Acad Sci USA* **100**, 5022–5027.
24 Xie J & Schultz PG (2006) A chemical toolkit for proteins – an expanded genetic code. *Nat Rev Mol Cell Biol* **7**, 775–782.
25 Zhang Z, Gildersleeve J, Yang Y-Y, Xu R, Loo JA, Uryu S, Wong C-H & Schultz PG (2004) A new strategy for the synthesis of glycoproteins. *Science* **303**, 371–373.
26 Xu R, Hanson SR, Zhang Z, Yang Y-Y, Schultz PG & Wong C-H (2004) Site-specific incorporation of the mucin-type N-acetylgalactosamine-a-O-threonine into protein in *Escherichia coli*. *J Am Chem Soc* **126**, 15654–15655.
27 Li B, Song H, Hauser S & Wang L-X (2006) A Highly efficient chemoenzymatic approach toward glycoprotein synthesis. *Organic Lett* **8**, 3081–3084.
28 Li B, Zeng Y, Hauser S, Song H & Wang L-X (2005) Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. *J Am Chem Soc* **127**, 9692–9693.
29 Zeng Y, Wang J, Li B, Hauser S, Li H & Wang L-X (2006) Glycopeptide synthesis through endo-glycosidase-catalyzed oligosaccharide transfer of sugar oxazolines: probing substrate structural requirement. *Chem Eur J* **12**, 3355–3364.
30 Dawson PE, Muir TW, Clark-Lewis I & Kent SBH (1994) Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779.
31 Yan LZ & Dawson PE (2001) Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J Am Chem Soc* **123**, 526–533.
32 Kochendoerfer GG, Chen S-Y, Mao F, Cressman S, Traviglia S, Shao H, Hunter CL, Low DW, Cagle EN, Carnevali M *et al*. (2003) Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* **299**, 884–887.
33 Macmillan D & Bertozzi CR (2004) Modular assembly of glycoproteins: towards the synthesis of GlyCAM-1 by using expressed protein ligation. *Angew Chem Int Edn* **43**, 1355–1359.
34 Liu L, Hong Z-Y & Wong C-H (2006) Convergent glycopeptide synthesis by traceless Staudinger ligation and enzymatic coupling. *ChemBioChem* **7**, 429–432.
35 Witte K, Sears P & Wong C-H (1997) Enzymic glycoprotein synthesis: preparation of ribonuclease glycoforms via enzymic glycopeptide condensation and glycosylation. *J Am Chem Soc* **119**, 2114–2118.
36 Doores KJ & Davis BG (2005) 'Polar patch' proteases as glycopeptiligases. *Chem Commun*, 168–170.
37 Davis NJ & Flitsch SL (1991) A novel method for the specific glycosylation of proteins. *Tetrahedron Lett* **32**, 6793–6796.
38 Ito Y, Hagihara S, Matsuo I & Totani K (2005) Structural approaches to the study of oligosaccharides in glycoprotein quality control. *Current Opin Struct Biol* **15**, 481–489.
39 Wong SY, Guile GR, Dwek RA & Arsequell G (1994) Synthetic glycosylation of proteins using N-(beta-saccharide) iodoacetamides: applications in site-specific glycosylation and solid-phase enzymic oligosaccharide synthesis. *Biochem J* **300** (Pt 3), 843–850.
40 Macindoe WM, van Oijen AH & Boons G-J (1998) A unique and highly facile method for synthesizing disulfide linked neoglycoconjugates: a new approach for remodeling of peptides and proteins. *Chem Commun*, 847–848.
41 Davis BG, Lloyd RC & Jones JB (1998) Controlled site-selective glycosylation of proteins by a combined site-directed mutagenesis and chemical modification approach. *J Organic Chem* **63**, 9614–9615.
42 Watt GM, Lund J, Levens M, Kolli VSK, Jefferis R & Boons G-J (2003) Site-specific glycosylation of an aglycosylated human IgG1-Fc antibody protein generates neoglycoproteins with enhanced function. *Chem Biol* **10**, 807–814.
43 Gamblin DP, Garnier P, Ward SJ, Oldham NJ, Fairbanks AJ & Davis BG (2003) Glycosyl phenylthiosulfonates (Glyco-PTS): novel reagents for glycoprotein synthesis. *Organic Biomol Chem* **1**, 3642–3644.
44 Davis BG, Maughan MAT, Green MP, Ullman A & Jones JB (2000) Glycomethanethiosulfonates: powerful reagents for protein glycosylation. *Tetrahedron Asymmetry* **11**, 245–262.
45 Davis BG, Khumtaveeporn K, Bott RR & Jones JB (1999) Altering the specificity of subtilisin *Bacillus lentus*

through the introduction of positive charge at single amino acid sites. *Bioorganic Med Chem* **7**, 2303–2311.
46 Davis BG, Shang X, DeSantis G, Bott RR & Jones JB (1999) The controlled introduction of multiple negative charge at single amino acid sites in subtilisin *Bacillus lentus*. *Bioorganic Med Chem* **7**, 2293–2301.
47 Matsumoto K, Davis BG & Jones JB (2002) Chemically modified 'polar patch' mutants of subtilisin in peptide synthesis with remarkably broad substrate acceptance: designing combinatorial biocatalysts. *Chem Eur J* **8**, 4129–4137.
48 Kenyon GL & Bruice TW (1977) Novel sulfhydryl reagents. *Methods Enzymol* **47**, 407–430.
49 Davis BG, Sala RF, Hodgson DRW, Ullman A, Khumtaveeporn K, Estell DA, Sanford K, Bott RR & Jones JB (2003) Selective protein degradation by ligand-targeted enzymes: towards the creation of catalytic antagonists. *ChemBioChem* **4**, 533–537.
50 Davis BG (2001) The controlled glycosylation of a protein with a bivalent glycan: towards a new class of glycoconjugates, glycodendriproteins. *Chem Commun*, 351–352.
51 Rendle PM, Seger A, Rodrigues J, Oldham NJ, Bott RR, Jones JB, Cowan MM & Davis BG (2004) Glycodendriproteins: a synthetic glycoprotein mimic enzyme with branched sugar-display potently inhibits bacterial aggregation. *J Am Chem Soc* **126**, 4750–4751.
52 Macmillan D, Bill RM, Sage KA, Fern D & Flitsch SL (2001) Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem Biol* **8**, 133–145.
53 Mullegger J, Chen HM, Warren RAJ & Withers SG (2006) Glycosylation of a neoglycoprotein by using glycosynthase and thioglycoligase approaches: the generation of a thioglycoprotein. *Angew Chem Int Edn* **45**, 2585–2588.
54 Watt GM & Boons G-J (2004) A convergent strategy for the preparation of N-glycan core di-, tri-, and pentasaccharide thioaldoses for the site-specific glycosylation of peptides and proteins bearing free cysteines. *Carbohydr Res* **339**, 181–193.
55 Gamblin DP, Garnier P, van Kasteren S, Oldham NJ, Fairbanks AJ & Davis BG (2004) Glyco-SeS: selenenylsulfide-mediated protein glycoconjugation – a new strategy in post-translational modification. *Angew Chem Int Edn* **43**, 828–833.
56 Liu H, Wang L, Brock A, Wong C-H & Schultz PG (2003) A method for the generation of glycoprotein mimetics. *J Am Chem Soc* **125**, 1702–1703.
57 Bernardes GJL, Gamblin DP & Davis BG (2006) The direct formation of glycosyl thiols from reducing sugars allows one-pot protein glycoconjugation. *Angew Chem Int Edn* **45**, 4007–4011.
58 Cox AD & Der CJ (1992) Protein prenylation: more than just glue? *Curr Opin Cell Biol* **4**, 1008–1016.
59 Magee AI & Seabra MC (2003) Are prenyl groups on proteins sticky fingers or greasy handles? *Biochem J* **376**, e3–e4.
60 Ramamurthy V, Roberts M, van den Akker F, Niemi G, Reh TA & Hurley JB (2003) AIPL1, a protein implicated in Leber's congenital amaurosis, interacts with and aids in processing of farnesylated proteins. *Proc Natl Acad Sci USA* **100**, 12630–12635.
61 Sinensky M (2000) Functional aspects of polyisoprenoid protein substituents: roles in protein–protein interaction and trafficking. *Biochim Biophys Acta* **1529**, 203–209.
62 Gomes AQ, Ali BR, Ramalho JS, Godfrey RF, Barral DC, Hume AN & Seabra MC (2003) Membrane targeting of Rab GTPases is influenced by the prenylation motif. *Mol Biol Cell* **14**, 1882–1899.
63 Wherlock M, Gampel A, Futter C & Mellor H (2004) Farnesyltransferase inhibitors disrupt EGF receptor traffic through modulation of the RhoB GTPase. *J Cell Sci* **117**, 3221–3231.
64 Pittler SJ, Fliesler SJ, Fisher PL, Keller PK & Rapp LM (1995) In vivo requirement of protein prenylation for maintenance of retinal cytoarchitecture and photoreceptor structure. *J Cell Biol* **130**, 431–439.
65 Cozzone AJ (2005) Role of protein phosphorylation on serine/threonine and tyrosine in the virulence of bacterial pathogens. *J Mol Microbiol Biotechnol* **9**, 198–213.
66 Oestman A, Hellberg C & Boehmer FD (2006) Protein-tyrosine phosphatases and cancer. *Nat Rev Cancer* **6**, 307–320.
67 Ghezzi P (2005) Regulation of protein function by glutathionylation. *Free Rad Res* **39**, 573–580.
68 O'Brian CA & Chu F (2005) Post-translational disulfide modifications in cell signaling-role of inter-protein, intra-protein, S-glutathionyl, and S-cysteaminyl disulfide modifications in signal transmission. *Free Rad Res* **39**, 471–480.
69 Rostovtsev VV, Green LG, Fokin VV & Sharpless KB (2002) A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew Chem Int Edn* **41**, 2596–2599.
70 Tornøe CW, Christensen C & Meldal M (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J Org Chem* **67**, 3057–3064.
71 Kiick KL, Saxon E, Tirrell DA & Bertozzi CR (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc Natl Acad Sci USA* **99**, 19–24.
72 Van Hest JCM, Kiick KL & Tirrell DA (2000) Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J Am Chem Soc* **122**, 1282–1288.
73 Kansas GS (1996) Selectins and their ligands: current concepts and controversies. *Blood* **88**, 3259–3287.
74 Somers WS, Tang J, Shaw GD & Camphausen RT (2000) Insights into the molecular basis of leukocyte

tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell* **103**, 467–479.

75 Vosseller K, Sakabe K, Wells L & Hart Gerald W (2002) Diverse regulation of protein function by O-GlcNAc: a nuclear and cytoplasmic carbohydrate post-translational modification. *Curr Opin Chem Biol* **6**, 851–857.

76 Lefebvre T, Guinez C, Dehennaut V, Beseme-Dekeyser O, Morelle W & Michalski J-C (2005) Does O-GlcNAc play a role in neurodegenerative diseases? *Expert Rev Proteomics* **2**, 265–275.

77 Elliot SP, Schmied R, Gabel CA & Ambron RT (1993) An 83 kDa O-GlcNAc-glycoprotein is found in the axoplasm and nucleus of *Aplysia* neurons. *J Neurosci* **13**, 2424–2429.