Getting a chemical handle on protein post-translational modification†

William P. Heal and Edward W. Tate*

Received 1st September 2009, Accepted 31st October 2009
First published as an Advance Article on the web 3rd December 2009
DOI: 10.1039/b917894e

This article highlights the emerging field of chemical proteomics, a powerful technology for the study of post- and co-translational modification of proteins. Genome mapping and the study of protein post-translational modifications have revealed the astounding chemical complexity present in the proteome of even the simplest organisms. The identification and characterisation of the modifications present on specific proteins in such complex mixtures has become a central challenge for post-genomic functional studies in cell and systems biology. In the chemical proteomic approach to this problem, protein-modifying enzymes and bioorthogonal chemoselective elaboration are exploited to deliver chemical tags to specific modified residues, enabling new advances in our understanding of protein modification.

Introduction

Gene sequencing and the elucidation of a growing variety of pathways for co- and post-translational modification (PTM) have revealed a staggering combinatorial chemical complexity in the proteome of all organisms. Classically, protein populations have been mapped using global methods such as 2D gel electrophoresis. More recent high-throughput techniques can enable identification and quantification of thousands of proteins in a single experiment.[1–3] Protein labelling techniques have greatly enhanced our ability to add handles such as fluorophores, isotope labels and affinity tags to enhance mass spectrometry-based proteomics, but the study of PTM specifically remains a significant challenge. Recently, techniques have emerged from the fields of chemical biology and chemical proteomics that allow the introduction of small, unobtrusive, site specific tags, offering a much more selective means of probing specific PTMs.[4–8]

Chemical proteomics

As illustrated in Fig. 1, tagging of a target protein with a small-molecule at the site of modification is common to all chemical proteomic approaches to PTM. This is typically accomplished by the introduction of a chemically-tagged analogue of the natural substrate for the enzyme (*e.g.* a transferase) that catalyses PTM of the target protein(s). The tag is designed to be as close to

Department of Chemistry and Chemical Biology Centre, South Kensington Campus, Imperial College, London, UK SW7 2AZ. E-mail: e.tate@imperial.ac.uk; Fax: +44 (0)20 75941139; Tel: +44 (0)20 75943752

† This paper is part of an *Organic & Biomolecular Chemistry* web theme issue on protein labelling and chemical proteomics. Guest editors: Ed Tate and Matthew Bogyo.

William P. Heal

William Heal received a BSc with honours in chemistry in 1998, and an MSc in chemical process R&D the following year from the University of Liverpool (UK). He stayed in Liverpool to complete a PhD in 2003, studying asymmetric synthesis under the supervision of Professor Stan Roberts, followed by postdoctoral research in medicinal chemistry (prion disease) at the University of Sheffield (UK). He is presently at Imperial College London (UK), undertaking postdoctoral research in the Tate group in the field of chemical biology. He is particularly interested in the design and synthesis of chemical tools for use in studying biological systems.

Edward W. Tate

Dr Tate received his PhD in the group of Prof. Steve Ley FRS at the University of Cambridge, and then worked with Prof. Sam Zard at Ecole Polytechnique (Paris) as an 1851 Research Fellow. Following postdoctoral research in molecular microbiology at the Pasteur Institute in Paris and in chemical biology at Imperial College London, he was awarded a BBSRC David Phillips Research Fellowship in 2006, and in 2008 he was appointed to a Lectureship in Chemical Biology in the Department of Chemistry, Imperial College London. His research group currently comprises over 20 researchers engaged in multiple aspects of the design and application of novel chemical approaches to understanding living systems, with an emphasis on the roles of protein modification in disease.

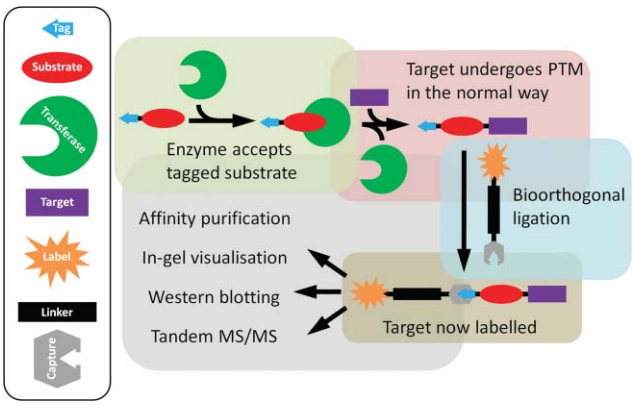
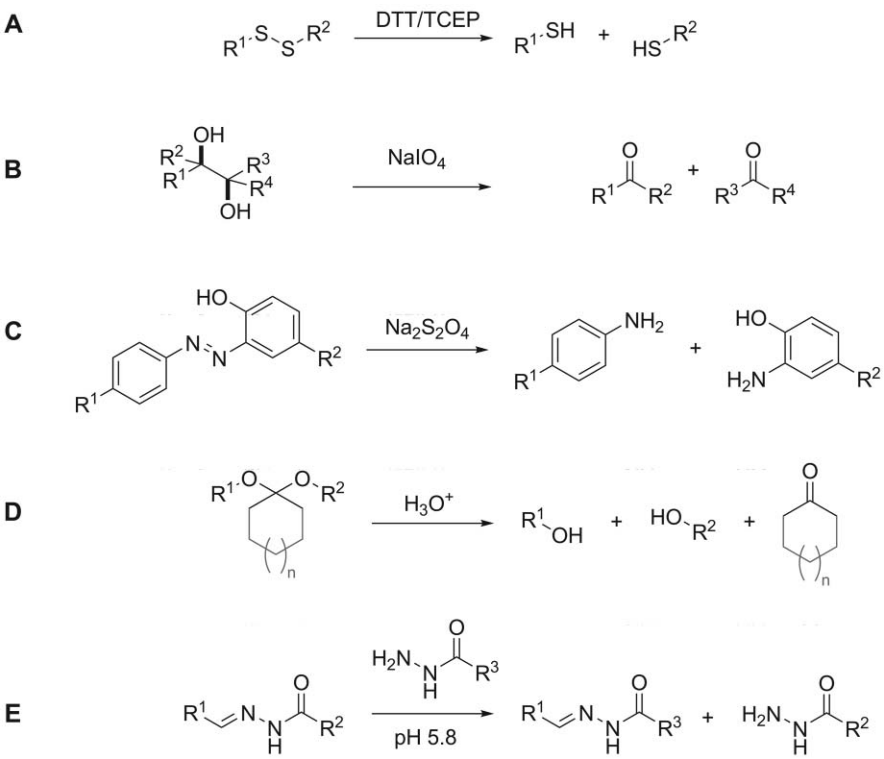


Fig. 1 Cartoon representation of the technique of 'tagging *via* substrate'. The transferase-mediated step can be carried out *in vivo* or *in vitro*, as can the bioorthogonal ligation. Down-stream options include affinity purification and in-gel fluorescence to detect and identify modified proteins.

biologically imperceptible as possible, such that tagging can be carried out *in vivo* using the native cellular machinery. However, *in vitro* tagging is also possible using purified recombinant enzymes or chemoselective reactions. Once the target proteins have been tagged, a highly selective bioorthogonal ligation is performed between the tag and another chemical entity (or *capture reagent*) bearing one or more secondary labels. In the large majority of studies reported to date this step has been based on 'click-chemistry', a set of reactions chosen for their extremely high chemoselectivity in a biological environment (often termed *bioorthogonality*), efficiency and mild, aqueous compatible reaction conditions.[9–14] The design of capture reagents can exploit the full range of modern synthetic organic chemistry resulting in a plethora of powerful chemical tools. These include such features as affinity groups (biotin, histidine tags, FLAG peptides, *etc.* allowing enrichment), fluorophores (for applications from direct in-gel visualisation to live cell imaging), oligonucleotide tags, radiolabels and/or stable isotope labels.

These components are typically separated from the clickable moiety by a flexible hydrophilic linker (such as PEG), improving aqueous solubility and facilitating interaction with a binding partner (*e.g.* biotin/avidin). The use of such affinity interactions for target enrichment can be confounded by nonspecific binding to both the binding partner and the solid support. The introduction of cleavable linkers has greatly expedited selectivity, allowing liberation of targets from the background (Scheme 1). These groups can be sundered by mild chemistry in the case of acetals,[15] 1,2-diols,[16] disulfides,[17] acylhydrazones[18] and diazo groups,[19] or by proteolysis (*e.g.* a TEV cleavage site[20]). Due to the range of chemistries available, mutually compatible methods can be found so that multiple features can be combined in a single reagent, specifically tailored to the system of interest.

A key advantage of this 'tagging-*via*-substrate' approach over genetic approaches, such as fusion to a reporter protein (*e.g.* EGFP), is the reduced chance of interference with biological function. Fusion proteins are large, typically >10 kDa, and there is no temporal control over the labelling event; in contrast, chemical tags are typically <50 Da, and the subsequent labelling



Scheme 1 Chemically-cleavable moieties that have proven value in the linker region of capture reagents. A, Dithiol cleavage, carried out under mildly reducing conditions. B, Diol cleavage, carried out in the presence of periodate. C, Diazo linker cleavage, by addition of sodium dithionite. D, Acid cleavable acetal. E, acylhydrazone exchange.

step can be performed at any stage of analysis. An important variant, termed *orthogonal chemical proteomics*, allows the study of a specific enzyme's activity discreet from any isozymes. This may be accomplished by modifying the enzyme to tolerate a tagged analogue of the natural substrate that would otherwise not be recognised. This approach is well-suited to the study of protein kinases, for example, where it is becoming an established technique.[21,22]

Currently, a number of research groups are engaged in developing chemical proteomics techniques, expanding the range of PTMs open to study. In addition, the toxicity and biocompatibility (solubility, off-target effects, *etc.*) of the reagents are under constant improvement, such that *in vivo* experiments inside living cells, and even whole organisms, are now possible.[12,23–26] This article will discuss significant recent developments in chemical proteomics, and their application to the study of four of the key PTMs: glycosylation, prenylation, acylation, and phosphorylation.[27]

Glycosylation

Protein glycosylation is one of the most abundant and diverse PTMs in mammalian cells, and plays an important role in many cellular processes, particularly cell–cell recognition, adhesion and signalling. Yet the considerable complexity of glycosylation and the difficulty of differentiating individual glycosides from complex mixtures of other polysaccharides has made this a challenging area of study using conventional techniques (mass spectrometry, antibody-based purification *etc.*). This fact, in combination with access to tagged sugars, meant that the study of glycosylation was one of the earliest developments in chemical proteomics of PTM,[28] and subsequent work has greatly increased our understanding of certain types of *N*- and *O*-linked glycosylation (Fig. 2). The first method (Fig. 2A), whereby azide tagged monosaccarides were fed to living cells, was pioneered by Bertozzi and co-workers.[28–30] This makes use of the cell's native machinery to incorporate tagged sugars into proteins *in vivo*. Subsequent capture *via* bioorthogonal ligation allows inclusion of an affinity group, making selective enrichment of the modified proteins of interest. This versatile technology has also been applied to the labelling of cell surface proteins, allowing surface re-functionalisation, and even to the tagging of glycosylated proteins in whole animals. The metabolic incorporation of *N*-azidoacetylmannosamine (ManNAz), *N*-azidoacetylgalactosamine (GalNAz), *N*-azidoacetylglucosamine (GlcNAz) and 6-azidofucose (6AzFuc) have all been studied, using capture reagents incorporating features such as FLAG peptides, polyhistidine tags and biotin for detection/enrichment.[31–38]

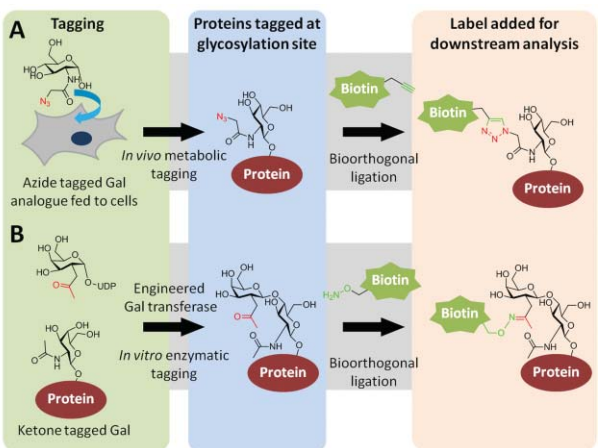


Fig. 2 Tagging-*via*-substrate applied to glycomics: **A**, Tagged glycoside (*e.g.* GalNAc, fucose and ManAc) fed to cells. **B**, Enzymatic addition (post-lysis) of tagged Gal analogues at *O*-GlcNAc sites by use of an engineered Gal transferase.

The second method (Fig. 2B), developed by Hsieh-Wilson and colleagues, focuses on β-*N*-acetylglucosamine (*O*-GlcNAc) modification of serine and threonine residues. This PTM is present in all higher eukaryotic organisms, and has been shown to be inducible and highly dynamic. It shares certain features with phosphorylation, and so-called ying/yang sites have been found to be alternately *O*-GlcNAc-ylated or phosphorylated as a regulatory switch. However, the nature of the modification (*i.e.* dynamic and often present at low cellular abundance) also represents a significant challenge for its detection and study. An engineered β-1,4-galactosyltransferase enabled post-lysis site-specific incorporation of a ketone tag that was ligated readily and selectively with aminooxy biotin derivatives, allowing downstream investigation.[39,40] This technology has since been extended and commercialised, and has found application in the identification of potential biomarkers and imaging.[41–45]

Protein lipidation

Lipid-based PTM, whilst a key feature of the regulation of cellular processes such as trafficking and signalling, poses a challenge for proteomics. This is chiefly due to the difficulty of detecting the lipid moiety, and a lack of effective antibody-based recognition. However, these obstacles have been overcome by the post-translational inclusion, *via* synthetic organic chemistry, of novel chemical tags.

Acylation

Protein acylation with long chain fatty acids is found in all eukaryotic cells, and is central to membrane localistion and other hydrophobic interactions (formation of lipid rafts or membrane micro domains, protein–protein interactions, signalling pathways, *etc.*). The addition of C14 and C16 unbranched saturated alkyl chains (myristate and palmitate respectively) are the most widely studied. However, the paucity of methods for biochemical analysis have rendered this a challenge.

Myristoylation. *N*-Myristoylation, the covalent attachment of myristate to the *N*-terminal glycine of a target protein, is catalysed by the enzyme Myristoyl-CoA:protein *N*-myristoyl transferase (NMT).[46–48] NMT has been characterised in a wide range of organisms from yeast to humans, and is a prominent drug target in several maladies and infections including the protozoan parasites[49–51] and cancer.[52,53] Furthermore, a severe heritable condition termed 'Noonan-like syndrome with loose anagen hair' was recently traced to a Ser2Gly mutation of the protein SHOC2, the first instance of a genetic disease resulting from an acquired *N*-terminal lipidation event.[54]

The ability to investigate the population of myristoylated proteins in a given system is highly significant in both answering basic questions in functional biology and the validation of NMT as a drug target. Despite being an essential PTM, myristic acid is

a relatively rare entity *in vivo* and in humans, for example, <3% of the proteome is myristoylated. This necessitates sensitive and selective methods for study. Classically, radiolabelled analogues of myristate were used, but low specific activity and the hazards involved in handling such reagents, along with a lack of technique for affinity enrichment, mean that there is a demand for alternative approaches.

In the early 1990s Gordon *et al.* demonstrated that *Saccharomyces cerevisiae* NMT (*Sc*NMT) tolerates fatty acylcoenzyme A (CoA) analogues that closely mimic myristoyl-CoA, particularly regarding chain length and flexibility.[55,56] Building on this, our group[57,58] and others[59,60] have developed a method for the introduction of small unobtrusive 'tags', *via* synthetic analogues of myristic acid (Fig. 3). Tag groups, such as azide or alkyne, that are tolerated by the transferase NMT are selectively and site-specifically added to the *N*-terminus of target proteins, allowing subsequent 'capture' by bioorthogonal ligation. Like the work on glycosylation described above, this ligation chemistry exploits 'click' reactions. Reactions of this type are rapid, highly chemoselective, efficient and, importantly, may be performed under mild conditions.[61] They are compatible with biological media such as buffered protein solutions, cell lysates, and even, in some cases, tissue culture. The tags are reacted with capture reagents bearing the reciprocal component of the ligation reaction, and additional functionality for down-stream processing. Capture reagents bearing affinity tags (*e.g.* biotin for detection and affinity purification) and fluorophores (allowing ready detection) have been investigated by several groups. The versatility of this approach as a general labelling tool has recently been demonstrated on purified recombinant proteins, *in vitro* and *in vivo*, in an *E. coli* co-expression system.[57,58] The latter makes use of *E. coli*, which like all prokaryotes has no endogenous NMT, to synthesise the transferase (*Candida albicans* NMT, *Ca*NMT) and a target protein (*Plasmodium falciparum* ARF1, *Pf*ARF1) by use of a dual plasmid approach. After growing the bacteria supplemented with tagged myristic acid and induced with IPTG, lysis and subsequent bioorthogonal ligation

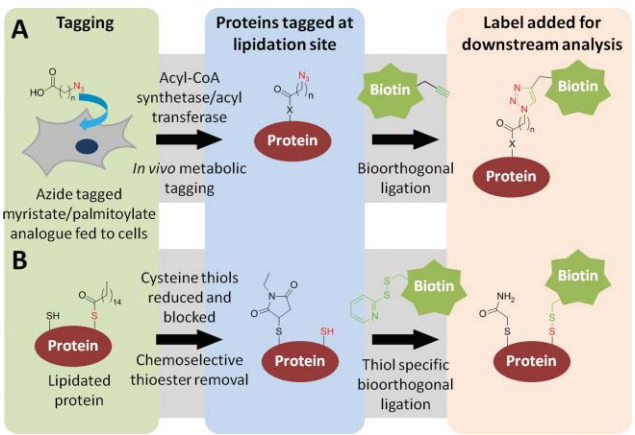


Fig. 3 Examples of tagging proteins at sites of fatty acid acylation. **A**, Tagged lipid is fed to cells, the cellular machinery processes the analogue as for the natural fatty acid, and the target protein is tagged in a site specific manner. Subsequent elaboration by bioorthogonal chemical ligation allows attachment of labels. **B**, All available cysteine thiols are reduced and blocked, before chemoselective cleavage of thioester bonds. The newly revealed thiols, previously sites of palmitoylation, are labelled selectively.

resulted in biotinylated *Pf*ARF1, which was identified by Western blotting. The site-specific nature of this tagging-*via*-myristate was demonstrated by use, *in vitro* and *in vivo*, of the G2A mutant of *Pf*ARF1, which was not myristoylated, and thus did not undergo subsequent biotinylation. In this work, both [3 + 2] copper-catalysed cycloaddition and the Staudinger–Bertozzi ligation were employed.

The same technology has also been applied to the study of myristoylation in mammalian cells. For example, Hang *et al.* were able to detect the well-characterised fatty-acylated protein Lck, a Src-family protein kinase essential for T cell activation, by similar metabolic labelling of Jurkat cells, lysis and subsequent capture by Staudinger ligation using phosphine–biotin.[60] Berthiaume *et al.* have studied the myristoylation in Jurkat T cell apoptosis, demonstrating that at least 15 proteins undergo this post-translational lipidation.[62] This work made use of the Staudinger ligation to attach biotin or the FLAG peptide. Across all these studies, metabolic chemical labelling has been shown to result in a $>10^7$-fold increase in signal compared to radiolabelling and with detection (in-gel or by Western blotting) in seconds rather than days or weeks.[57,58,60,62,63]

Palmitoylation. Palmitoylation, the addition of the unsaturated C16 palmitoyl group to the cysteine thiols of a target protein, is more prevalent than myristoylation. It promotes hydrophobic interactions in a similar way to the other co- and post-translational lipidations, but differs in its reversibility.[64] This commutability was exploited by Davis *et al.* to study palmitoylation in the yeast *Saccharomyces cerevisiae*.[65] Purification of palmitoylated proteins was accomplished by capping free cysteine thiols with *N*-ethyl maleimide, removal of palmitoyl groups by treatment with hydroxylamine, followed by biotinylation of the newly exposed thiols. In this manner, 35 new palmitoylation targets were identified after binding to streptavidin and MudPIT (multi-dimensional protein identification technology)[66] analysis. Interestingly, the biotinylating reagent (*N*-(6-(biotinamido)hexyl)-3′-(2′-pyridyldithio)-propionamide (biotin-HPDP) was used to link the affinity group *via* a disulfide bond, allowing later cleavage. This landmark study of palmitoylation was followed by a comprehensive assessment of synaptic palmitoylation using the same technique.[67]

In a similar manner, Zacharias and co-workers (who have recently reviewed palmitoyl acyltransferases, known substrates and methods for their assay[68]) were able to identify CKAP4/p63 (a putative tumour suppressor) as a major substrate of palmitoyl acyltransferase DHHC2.[69] They used *S*-methyl methanethiosulfonate (MMTS) as the thiol capping reagent, followed by hydroxylamine treatment to remove palmitoyl groups, and lastly, isotope coded affinity tag (ICAT, for tandem MS experiments) labelling of the newly revealed cysteine thiols.

However, this acyl-biotinyl exchange (ABE) methodology suffers from a relatively high false positive rate due to the multiple chemical manipulations required. A complementary approach was recently reported by Cravatt and Martin, who used the tagging *via* substrate approach to identify 125 predicted palmitoylated proteins in Jurkat T cells.[70] Like the myristoylation studies (*v.s.*), an azido–biotin species was used to affinity tag proteins bearing an alkynyl palmitate. Once isolated from the membrane fractions of lysed cells, palmitoylated proteins were identified by MudPIT.

Prenylation

Prenylation involves the covalent attachment of a polyisoprene lipid (*e.g.* farnesyl, geranylgeranyl) to a cysteine residue, typically in the *C*-terminal part of the target protein. Analogous to acylation, prenylation is central to cellular processes involving protein–protein interactions, trafficking and membrane localisation.[71] Most notably, incorrect prenylation of the Ras and Rab super families of proteins has been implicated in the etiology of infectious diseases, heritable conditions and cancer,[72–76] and the transferases responsible are therefore under investigation as potential drug targets. The geranylgeranylation of Rab GTPases is exclusively performed by Rab geranylgeranyl transferase (RGGTase), and has been reviewed extensively by Seabra *et al.*[75,77] and Waldman *et al.*,[78] who also provided the first crystal structure of RGGTase, and of RGGTase in complex with an inhibitor.[79,80] However, for similar reasons to those discussed for acylation, the study of protein prenylation has proved challenging, compounded by difficulties encountered in the mass analysis of very hydrophobic prenylated peptides.[3] Thus, less selective or sensitive methods such as hydrophobic affinity chromatography or radioisotope labelling have been relied upon.

As early as 1997, Edelstein and Distefano showed that protein farnesyl transferase (PFTase) tolerates some chemical modification of the prenyl pyrophosphate unit[81] and later work showed that a considerably wider range of functionalities, including ketones, ethers, esters and amides, could be accepted in the prenyl substrates, which could thus be added to substrate peptides and intact proteins.[82] The inclusion of bioorthogonal tags and subsequent chemical elaboration, in the same manner as outlined above for acylation, was therefore a natural development. Indeed, site specific labelling of recombinant proteins has been reported using azidofarnesyl pyrophosphates,[37,83] including the enrichment of the resultant prenylated entities from cells in culture. Although not as developed as chemical glycomics, the significance of prenylation in various diseases is driving many research efforts in this field.[84,85] In a very recent study of Hutchinson-Gilford progeria syndrome, azidogeranylgeranyl alcohol, a reagent recently commercialised for non-radioactive metabolic labelling, was used with subsequent ligation to a TAMRA-containing molecule to study the prenylation of progerin in fibroblasts.[86]

Rab prenylation is carried out by a complex of RGGTase and a Rab escort protein (REP). Mutation of REP-1 is known to cause the X-linked retinal degenerative disease Choroideremia. The Rab proteins in the choroid and retinal pigment epithelium of the eye are incorrectly prenylated and thus not transported to the correct membranes, resulting in blindness.[87,88] Work is underway in our labs to use a combination of enzyme-ligand engineering and bioorthogonal ligation chemistry to attach labels to Rab proteins, to identify those that are inefficiently prenylated, with the aim of improving our understanding of the biology of the disease and identifying novel targets for therapeutic intervention. In a somewhat different approach, Alexandrov *et al.* were able to apply a simple synthetic biotin geranylpyrophosphate that resembles geranylgeranylpyrophosphate in terms of length, by structure-guided engineering of protein prenyltransferases.[89] The researchers used these tools to study the effects of protein prenyltransferase inhibitors by following the mis-prenylation of RabGTPases throughout the proteome.

Phosphorylation

Although one of the most widely studied protein PTMs, phosphorylation still poses many challenges in the field of proteomics.[90] Although many methods have been developed for the isolation of phosphorylated proteins, mainly involving types of metal affinity chromatography, they are frequently blighted by poor specificity, particularly when acidic peptides/proteins are present, and tend to favour polyphosphorylated entities.[3] Despite the more recent development of monoclonal anti-phosphotyrosine antibodies,[91,92] which allow for the specific enrichment of phosphotyrosine-containing proteins, non-sequence specific antibodies towards PTMs have remained elusive. For example, a recent attempt employing phage display to identify sequence-independent antibodies specific for sulfyltyrosine-modified proteins managed to find only one hit from a screening of nearly 8000 selected clones.[93]

As an alternative and complementary approach, the site-specific modification of residues bearing phosphoseryl or phosphothreonyl modifications can be accomplished by exploiting the tendency of the phosphate moiety to undergo β-elimination under strongly basic conditions.[90,94–96] This chemistry results in a reactive dehydroalanine (*i.e.* a Michael acceptor), which is susceptible to nucleophilic attack (Fig. 4A). Ethanedithiol was used to establish a nucleophile at the site of phosphorylation (necessitating prior capping of the cysteine thiols), before treatment with an electrophilic biotinylated entity. Issues with the efficiency of the chemistry (*O*-glycans can also eliminate under these conditions) and sample recovery notwithstanding, this approach allowed a limited analysis of the phosphoproteome from *Arabidopsis thaliana*, with some 30 or so proteins being characterised.[94]

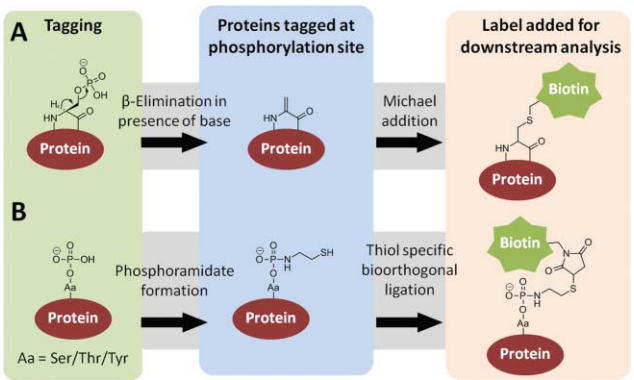


Fig. 4 Approaches to protein tagging at the site of phosphorylation. **A**, After oxidation of free thiols (to prevent unwanted nucleophilic addition), base-mediated β-elimination of the phosphates from serine results in the formation of a Michael acceptor. A thiol-containing source of biotin is then added for site specific derivatisation. **B**, Alternatively, all phosphates undergo phosphoramidate formation, followed by thiol-specific labelling.

In a related method for mapping the phosphorylome, the proteins are reversibly immobilised on a solid support or dendrimer *via* a short sequence of reactions.[97–99] After removal of the non-phosphorylated entities, the bound proteins are released for analysis. One of the advantages of this approach over that described previously is that under the conditions of phosphoramidate chemistry used, the phosphate remains attached to the protein. This allows the site of phosphorylation to be determined during the subsequent analysis. This advantage was

demonstrated elegantly in the discovery of a large number of phosphorylation sites on the protein dFOXO, from the cytosolic fraction of *Drosophila melanogaster* Kc167 cells.[97] Additionally, selective enrichment of thiophosphorylated polypeptides has been accomplished by chemoselective alkylation.[100] Exploiting the enhanced nucleophilicity of the thiophosphate sulfur at low pH (over other nucleophilic amino acid features), selective attachment of a biotinylated (or solid supported) alkylating reagent was possible (Fig. 4B). Interestingly, these researchers made use of a photocleavable linker to facilitate separation of the thiophosphorylated entities.

Pioneered by Shokat *et al.*,[21] a chemical genetics approach to the study of protein phosphorylation is among the exciting recent developments in the area. Essentially, an unnatural base-modified ATP analogue is used, which can only be tolerated by a kinase possessing a mutationally enlarged ATP binding site (the so-called 'bump-hole' approach). In this way, only the substrates of the (modified) kinase of interest are derivatised with the bulkier ATP analogue. The inclusion of a chemical tag allows, *via* transfer of the analogue, site-specific tagging of the kinase substrates, and subsequent elucidation. In a recent example, the synthesis of several enlarged/tagged ATP analogues is described, along with a screen of these with protein kinase cdk2, and its known substrate $p27^{kip1}$.[101] Subsequent work by Green and Pflum has shown that, in selected cases, it is possible to transfer a labelled γ-phosphate moiety from an ATP analogue, to generate a biotinylated or dansylated kinase substrate. This approach has so far only been shown to work for kinases in which the γ-phosphate is solvent exposed in the ATP binding site, which is a relative rarity amongst kinases characterised to date.[102,103] Nevertheless, this very direct approach is clearly an attractive prospect for screening kinase activity in cell-free systems, and merits further exploration in the future.

Other PTMs

PTM is an incredibly rich area for the development and application of new techniques in chemical biology. Many dozens of PTMs are known with more examples identified on a monthly basis, offering a wealth of opportunities to apply chemical approaches to further our understanding of these diverse and fascinating phenomena. For example, the addition of small polypeptides such as ubiquitin or SUMO (small ubiquitin-related modifier); whilst the destructive role of ubiquitination has been understood for some time (marking proteins for recognition by the proteosome prior to degradation), it is now also known that the modification has a role in regulating gene expression, signal transduction and intracellular transport.[104] Less is known about sumoylation, although it has been implicated in several regulatory functions including the regulation of chromatin substructure, transcription factor activity, subcellular compartmentalisation, protein complex assembly and DNA binding.[105–107] Zhao and co-workers reported the use of affinity chromatography-coupled HPLC/MS/MS analysis for identification of sumoylated entities from HEK-293 cells. The approach made use of HA-tagged SUMOs (HA tag sequence: YPYDVPDYA), which allowed immunopurification prior to protein identification. By this approach, 21 candidate sumoylated targets were identified from whole cell lysates.[108] A related approach was previously applied with great success to the identification of ubiquitinated proteins in a strain of the yeast *Saccharomyces cerevisiae* that expresses a His_6-tagged ubiquitin analogue.[109] In another recent development, Lin *et al.* were able to use an alkyne-tagged nicotinamide adenine dinucleotide (NAD) to investigate the protein ADP-ribosyltransferase-catalysed transfer of ADP-ribose from NAD to target proteins, promising the future identification of the many potential targets of poly- and mono-ADP-ribosylation, both of which are important PTMs in regulating gene expression.[110]

Conclusions

Whilst the information provided by genome sequencing has revolutionised research in biology, RNA splicing and the myriad pathways of protein maturation/modification mean that the gene sequence only scratches the surface of the complexity of the proteome. Genetic information is relatively static, whilst the proteome undergoes a staggering range of modifications, many highly dynamic, in a constant state of flux as part of signalling, intracellular transport, gene regulation, shock response, disease progression, *etc*. Furthermore, epigenetic modifications represent an additional level of genomic complexity that is not fixed at the DNA sequence level, and is partly mediated and encoded by protein PTMs. It is the vibrant field at the chemistry/biology interface that is proving the most fertile ground for cultivation of new technologies for studying PTM. As the power of emerging techniques is realised research efforts in the field have grown, widening the scope both in terms of the biological questions to which chemical proteomics can be applied, and in terms of the diversity of the chemical tool box available. The recent development of commercial kits for metabolic bioorthogonal tagging (Invitrogen[111]) bodes well for the future of chemical proteomics. The approach will be at its most effective when paired with complementary techniques such as chemical genetics or RNA interference that manipulate specific aspects of the system in a temporally-defined manner. Under these conditions, chemical proteomic analysis can provide a unique window onto changes in post-translational modification that serve to mediate downstream processes.

In the near future, it may be anticipated that this powerful method of chemical tagging will be extended to the study of PTMs as biomarkers of disease,[112,113] and permit the effects of transferase inhibitors to be elucidated, with collateral applications in enzyme assays and target validation studies. The paramount importance of PTMs such as methylation and acetylation, and of DNA modification in transcriptional regulation and epigenetic control presents a strikingly significant opportunity for the development and application of chemical tagging methodologies that has yet to be explored in depth. It is also worth noting that the elegant chemistry first exploited for bioorthogonal ligation in chemical proteomics has found many applications outside of the remit of this review. For example, DNA modification[114–117] and activity-based protein profiling (ABPP)[118–122] both utilise such reactions. Indeed, the activities of several enzymes discussed in this manuscript have been studied using the complementary approach of ABPP,[123–125] and in concert these methods have the potential to provide a comprehensive picture of the regulation of PTM. Tagging-*via*-substrate is also able to provide a general method for the site-specific transferase-mediated labelling of recombinant proteins,

both *in vitro* and *in vivo*, and in this biotechnological guise it has already resulted in many varied applications.[126]

Perhaps the main challenge for researchers in this field is the high level of expertise required in biochemistry, cell and molecular biology, as well as in synthetic chemistry. Nevertheless, in the modern research climate it is necessary and desirable for synthetic chemists to apply their skills to problems of importance to the wider scientific community. Chemistry has a fundamental and leading role to play in the future of biology and medicine, and chemical proteomics is emerging as an integral part of this multidisciplinary landscape.

Acknowledgements

The authors gratefully acknowledge financial support from Cancer Research UK (grant C29637/A9913) and the Biotechnology and Biological Sciences Research Council, UK (grant BB/D02014X/1).

Notes and references

1 N. G. Ahn, J. B. Shabb, W. M. Old and K. A. Resing, *ACS Chem. Biol.*, 2007, **2**, 39–52.
2 T. Kislinger and A. Emili, *Expert Rev. Proteomics*, 2005, **2**, 27–39.
3 J. Reinders, U. Lewandrowski, J. Moebius, Y. Wagner and A. Sickmann, *Proteomics*, 2004, **4**, 3686–3703.
4 A. B. Berger, P. M. Vitorino and M. Bogyo, *Am. J. PharmacoGenomics*, 2004, **4**, 371–381.
5 H. Daub, *Biochim. Biophys. Acta, Proteins Proteomics*, 2005, **1754**, 183–190.
6 D. A. Jeffery and M. Bogyo, *Curr. Opin. Biotechnol.*, 2003, **14**, 87–95.
7 D. S. Sem, *Expert Rev. Proteomics*, 2004, **1**, 165–178.
8 S. A. Sieber and B. F. Cravatt, *Chem. Commun.*, 2006, 2311–2319.
9 N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi, *ACS Chem. Biol.*, 2006, **1**, 644–648.
10 I. S. Carrico, B. L. Carlson and C. R. Bertozzi, *Nat. Chem. Biol.*, 2007, **3**, 321–322.
11 M. Köhn and R. Breinbauer, *Angew. Chem., Int. Ed.*, 2004, **43**, 3106–3116.
12 J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13–21.
13 S. Sen Gupta, J. Kuzelka, P. Singh, W. G. Lewis, M. Manchester and M. G. Finn, *Bioconjugate Chem.*, 2005, **16**, 1572–1579.
14 P. F. van Swieten, M. A. Leeuwenburgh, B. M. Kessler and H. S. Overkleeft, *Org. Biomol. Chem.*, 2005, **3**, 20–27.
15 K. Srinivasachar and D. M. Neville, Jr., *Biochemistry*, 1989, **28**, 2501–2509.
16 O. Melnyk, J. S. Fruchart, C. Grandjean and H. Gras-Masse, *J. Org. Chem.*, 2001, **66**, 4153–4160.
17 C. A. Gartner, J. E. Elias, C. E. Bakalarski and S. P. Gygi, *J. Proteome Res.*, 2007, **6**, 1482–1491.
18 K. D. Park, R. Liu and H. Kohn, *Chem. Biol.*, 2009, **16**, 763–772.
19 S. H. Verhelst, M. Fonovic and M. Bogyo, *Angew. Chem., Int. Ed.*, 2007, **46**, 1284–1286.
20 A. E. Speers and B. F. Cravatt, *J. Am. Chem. Soc.*, 2005, **127**, 10018–10019.
21 P. J. Alaimo, M. A. Shogren-Knaak and K. M. Shokat, *Curr. Opin. Chem. Biol.*, 2001, **5**, 360–367.
22 L. M. Elphick, S. E. Lee, V. Gouverneur and D. J. Mann, *ACS Chem. Biol.*, 2007, **2**, 299–314.
23 J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 16793–16797.
24 K. P. Chiang, S. Niessen, A. Saghatelian and B. F. Cravatt, *Chem. Biol.*, 2006, **13**, 1041–1050.
25 D. H. Dube, J. A. Prescher, C. N. Quang and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 4819–4824.
26 J. A. Prescher, D. H. Dube and C. R. Bertozzi, *Nature*, 2004, **430**, 873–877.
27 E. W. Tate, *J. Chem. Biol.*, 2008, **1**, 17–26.
28 C. R. Bertozzi and L. L. Kiessling, *Science*, 2001, **291**, 2357–2364.
29 E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007–2010.
30 E. Saxon and C. R. Bertozzi, *Annu. Rev. Cell Dev. Biol.*, 2001, **17**, 1–23.
31 A. P. Bussink, P. F. van Swieten, K. Ghauharali, S. Scheij, M. van Eijk, T. Wennekes, G. A. van der Marel, R. G. Boot, J. M. Aerts and H. S. Overkleeft, *J. Lipid Res.*, 2007, **48**, 1417–1421.
32 H. C. Hang, C. Yu, D. L. Kato and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 14846–14851.
33 T. L. Hsu, S. R. Hanson, K. Kishikawa, S. K. Wang, M. Sawa and C. H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2614–2619.
34 S. T. Laughlin and C. R. Bertozzi, *Nat. Protoc.*, 2007, **2**, 2930–2944.
35 S. J. Luchansky, S. Argade, B. K. Hayes and C. R. Bertozzi, *Biochemistry*, 2004, **43**, 12358–12366.
36 A. Nandi, R. Sprung, D. K. Barma, Y. Zhao, S. C. Kim, J. R. Falck and Y. Zhao, *Anal. Chem.*, 2006, **78**, 452–458.
37 R. Sprung, A. Nandi, Y. Chen, S. C. Kim, D. Barma, J. R. Falck and Y. Zhao, *J. Proteome Res.*, 2005, **4**, 950–957.
38 D. J. Vocadlo, H. C. Hang, E. J. Kim, J. A. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9116–9121.
39 N. Khidekel, S. Arndt, N. Lamarre-Vincent, A. Lippert, K. G. Poulin-Kerstien, B. Ramakrishnan, P. K. Qasba and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2003, **125**, 16162–16163.
40 N. Khidekel, S. B. Ficarro, E. C. Peters and L. C. Hsieh-Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13132–13137.
41 Z. Wang, K. Park, F. Comer, L. C. Hsieh-Wilson, C. D. Saudek and G. W. Hart, *Diabetes*, 2009, **58**, 309–317.
42 H. E. Murrey, S. B. Ficarro, C. Krishnamurthy, S. E. Domino, E. C. Peters and L. C. Hsieh-Wilson, *Biochemistry*, 2009, **48**, 7261–7270.
43 J. E. Rexach, P. M. Clark and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2008, **4**, 97–106.
44 P. M. Clark, J. F. Dweck, D. E. Mason, C. R. Hart, S. B. Buck, E. C. Peters, B. J. Agnew and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2008, **130**, 11576–11577.
45 N. Khidekel, S. B. Ficarro, P. M. Clark, M. C. Bryan, D. L. Swaney, J. E. Rexach, Y. E. Sun, J. J. Coon, E. C. Peters and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2007, **3**, 339–348.
46 C. A. Chen and D. R. Manning, *Oncogene*, 2001, **20**, 1643–1652.
47 M. Pechlivanis and J. Kuhlmann, *Biochim. Biophys. Acta, Proteins Proteomics*, 2006, **1764**, 1914–1931.
48 M. D. Resh in *Subcellular Biochemistry*, 2004, vol. 37, pp. 217232.
49 N. H. Georgopapadakou, *Expert Opin. Invest. Drugs*, 2002, **11**, 1117–1125.
50 H. P. Price, M. R. Menon, C. Panethymitaki, D. Goulding, P. G. McKean and D. F. Smith, *J. Biol. Chem.*, 2003, **278**, 7206–7214.
51 N. Takamune, H. Hamada, S. Misumi and S. Shoji, *FEBS Lett.*, 2002, **527**, 138–142.
52 C. E. Ducker, J. J. Upson, K. J. French and C. D. Smith, *Mol. Cancer Res.*, 2005, **3**, 463–476.
53 A. Shrivastav, M. K. Pasha, P. Selvakumar, S. Gowda, D. J. Olson, A. R. Ross, J. R. Dimmock and R. K. Sharma, *Cancer Res.*, 2003, **63**, 7975–7978.
54 V. Cordeddu, E. Di Schiavi, L. A. Pennacchio, A. Ma'ayan, A. Sarkozy, V. Fodale, S. Cecchetti, A. Cardinale, J. Martin, W. Schackwitz, A. Lipzen, G. Zampino, L. Mazzanti, M. C. Digilio, S. Martinelli, E. Flex, F. Lepri, D. Bartholdi, K. Kutsche, G. B. Ferrero, C. Anichini, A. Selicorni, C. Rossi, R. Tenconi, M. Zenker, D. Merlo, B. Dallapiccola, R. Iyengar, P. Bazzicalupo, B. D. Gelb and M. Tartaglia, *Nat. Genet.*, 2009, **41**, 1022–1026.
55 R. O. Heuckeroth, E. Jackson-Machelski, S. P. Adams, N. S. Kishore, M. Huhn, A. Katoh, T. Lu, G. W. Gokel and J. I. Gordon, *J. Lipid Res.*, 1990, **31**, 1121–1129.
56 N. S. Kishore, T. B. Lu, L. J. Knoll, A. Katoh, D. A. Rudnick, P. P. Mehta, B. Devadas, M. Huhn, J. L. Atwood and S. P. Adams, *J. Biol. Chem.*, 1991, **266**, 8835–8855.
57 W. P. Heal, S. R. Wickramasinghe, P. W. Bowyer, A. A. Holder, D. F. Smith, R. J. Leatherbarrow and E. W. Tate, *Chem. Commun.*, 2008, 480–482.
58 W. P. Heal, S. R. Wickramasinghe, R. J. Leatherbarrow and E. W. Tate, *Org. Biomol. Chem.*, 2008, **6**, 2308–2315.
59 G. Charron, M. M. Zhang, J. S. Yount, J. Wilson, A. S. Raghavan, E. Shamir and H. C. Hang, *J. Am. Chem. Soc.*, 2009, **131**, 4967–4975.
60 H. C. Hang, E. J. Geutjes, G. Grotenbreg, A. M. Pollington, M. J. Bijlmakers and H. L. Ploegh, *J. Am. Chem. Soc.*, 2007, **129**, 2744–2745.

61 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249–1262.
62 D. D. Martin, G. L. Vilas, J. A. Prescher, G. Rajaiah, J. R. Falck, C. R. Bertozzi and L. G. Berthiaume, *FASEB J.*, 2008, **22**, 797–806.
63 M. A. Kostiuk, M. M. Corvi, B. O. Keller, G. Plummer, J. A. Prescher, M. J. Hangauer, C. R. Bertozzi, G. Rajaiah, J. R. Falck and L. G. Berthiaume, *FASEB J.*, 2008, **22**, 721–732.
64 J. Charollais and F. G. Van Der Goot, *Mol. Membr. Biol.*, 2009, **26**, 55–66.
65 A. F. Roth, J. Wan, A. O. Bailey, B. Sun, J. A. Kuchar, W. N. Green, B. S. Phinney, J. R. Yates, III and N. G. Davis, *Cell*, 2006, **125**, 1003–1013.
66 J. K. Eng, A. L. McCormack and J. R. Yates, III, *J. Am. Soc. Mass Spectrom.*, 1994, **5**, 976–989.
67 R. Kang, J. Wan, P. Arstikaitis, H. Takahashi, K. Huang, A. O. Bailey, J. X. Thompson, A. F. Roth, R. C. Drisdel, R. Mastro, W. N. Green, J. R. Yates, III, N. G. Davis and A. El-Husseini, *Nature*, 2008, **456**, 904–909.
68 S. L. Planey and D. A. Zacharias, *Mol. Membr. Biol.*, 2009, **26**, 14–31.
69 J. Zhang, S. L. Planey, C. Ceballos, S. M. Stevens, Jr., S. K. Keay and D. A. Zacharias, *Mol. Cell. Proteomics*, 2008, **7**, 1378–1388.
70 B. R. Martin and B. F. Cravatt, *Nat. Methods*, 2009, **6**, 135–138.
71 F. L. Zhang and P. J. Casey, *Annu. Rev. Biochem.*, 1996, **65**, 241–269.
72 C. Flotho, C. Kratz and C. M. Niemeyer, *Curr. Drug Targets*, 2007, **8**, 715–725.
73 P. A. Konstantinopoulos, M. V. Karamouzis and A. G. Papavassiliou, *Nat. Rev. Drug Discovery*, 2007, **6**, 541–555.
74 K. T. Lane and L. S. Beese, *J. Lipid Res.*, 2006, **47**, 681–699.
75 K. F. Leung, R. Baron and M. C. Seabra, *J. Lipid Res.*, 2006, **47**, 467–475.
76 S. M. Sebti, *Cancer Cell*, 2005, **7**, 297–300.
77 K. F. Leung, R. Baron, B. R. Ali, A. I. Magee and M. C. Seabra, *J. Biol. Chem.*, 2007, **282**, 1487–1497.
78 L. Brunsveld, J. Kuhlmann, K. Alexandrov, A. Wittinghofer, R. S. Goody and H. Waldmann, *Angew. Chem., Int. Ed.*, 2006, **45**, 6622–6646.
79 O. Pylypenko, A. Rak, R. Reents, A. Niculae, V. Sidorovitch, M.-D. Cioaca, E. Bessolitsyna, N. H. Thoma, H. Waldmann, I. Schlichting, R. S. Goody and K. Alexandrov, *Mol. Cell*, 2003, **11**, 483–494.
80 A. Rak, O. Pylypenko, T. Durek, A. Watzke, S. Kushnir, L. Brunsveld, H. Waldmann, R. S. Goody and K. Alexandrov, *Science*, 2003, **302**, 646–650.
81 R. L. Edelstein and M. D. Distefano, *Biochem. Biophys. Res. Commun.*, 1997, **235**, 377–382.
82 B. P. Duckworth, Z. Zhang, A. Hosokawa and M. D. Distefano, *ChemBioChem*, 2007, **8**, 98–105.
83 Y. Kho, S. C. Kim, C. Jiang, D. Barma, S. W. Kwon, J. Cheng, J. Jaunbergs, C. Weinbaum, F. Tamanoi, J. Falck and Y. Zhao, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12479–12484.
84 A. Khwaja, J. O'Connolly and B. M. Hendry, *Lancet*, 2000, **355**, 741–744.
85 M. C. Seabra, E. H. Mules and A. N. Hume, *Trends Mol. Med.*, 2002, **8**, 23–30.
86 B. S. Davies, S. H. Yang, E. Farber, R. Lee, S. B. Buck, D. A. Andres, H. P. Spielmann, B. J. Agnew, F. Tamanoi, L. G. Fong and S. G. Young, *J. Lipid Res.*, 2008, **50**, 126–134.
87 M. C. Seabra, *Ophthalmic Genetics*, 1996, **17**, 43–46.
88 M. C. Seabra, M. S. Brown and J. L. Goldstein, *Science*, 1993, **259**, 377–381.
89 U. T. T. Nguyen, Z. Guo, C. Delon, Y. Wu, C. Deraeve, B. Franzel, R. S. Bon, W. Blankenfeldt, R. S. Goody, H. Waldmann, D. Wolters and K. Alexandrov, *Nat. Chem. Biol.*, 2009, **5**, 227–235.
90 M. Hjerrild and S. Gammeltoft, *FEBS Lett.*, 2006, **580**, 4764–4770.
91 S. J. Ding, W. J. Qian and R. D. Smith, *Expert Rev. Proteomics*, 2007, **4**, 13–23.
92 F. Liang, S. Kumar and Z. Y. Zhang, *Mol. BioSyst.*, 2007, **3**, 308–316.
93 J. W. Kehoe, N. Velappan, M. Walbolt, J. Rasmussen, D. King, J. Lou, K. Knopp, P. Pavlik, J. D. Marks, C. R. Bertozzi and A. R. Bradbury, *Mol. Cell. Proteomics*, 2006, **5**, 2350–2363.
94 S. J. Kwon, E. Y. Choi, J. B. Seo and O. K. Park, *Mol. Cells*, 2007, **24**, 268–275.
95 A. Leitner and W. Lindner, *Proteomics*, 2006, **6**, 5418–5434.
96 Y. Oda, T. Nagasu and B. T. Chait, *Nat. Biotechnol.*, 2001, **19**, 379–382.
97 B. Bodenmiller, L. N. Mueller, P. G. Pedrioli, D. Pflieger, M. A. Junger, J. K. Eng, R. Aebersold and W. A. Tao, *Mol. BioSyst.*, 2007, **3**, 275–286.
98 M. Warthaka, P. Karwowska-Desaulniers and M. K. Pflum, *ACS Chem. Biol.*, 2006, **1**, 697–701.
99 H. Zhou, J. D. Watts and R. Aebersold, *Nat. Biotechnol.*, 2001, **19**, 375–378.
100 S. W. Kwon, S. C. Kim, J. Jaunbergs, J. R. Falck and Y. Zhao, *Mol. Cell. Proteomics*, 2003, **2**, 242–247.
101 S. E. Lee, L. M. Elphick, A. A. Anderson, L. Bonnac, E. S. Child, D. J. Mann and V. Gouverneur, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3804–3807.
102 K. D. Green and M. K. Pflum, *J. Am. Chem. Soc.*, 2007, **129**, 10–11.
103 K. D. Green and M. K. Pflum, *ChemBioChem*, 2009, **10**, 234–237.
104 F. J. Herrera and S. J. Triezenberg, *Curr. Biol.*, 2004, **14**, R622–624.
105 F. Melchior, *Annu. Rev. Cell Dev. Biol.*, 2000, **16**, 591–626.
106 F. Melchior, M. Schergaut and A. Pichler, *Trends Biochem. Sci.*, 2003, **28**, 612–618.
107 A. Verger, J. Perdomo and M. Crossley, *EMBO Rep.*, 2003, **4**, 137–142.
108 Y. Zhao, S. W. Kwon, A. Anselmo, K. Kaur and M. A. White, *J. Biol. Chem.*, 2004, **279**, 20999–21002.
109 J. Peng, D. Schwartz, J. E. Elias, C. C. Thoreen, D. Cheng, G. Marsischky, J. Roelofs, D. Finley and S. P. Gygi, *Nat. Biotechnol.*, 2003, **21**, 921–926.
110 J. Du, H. Jiang and H. Lin, *Biochemistry*, 2009, **48**, 2878–2890.
111 http://probes.invitrogen.com/media/pis/mp10186.pdf.
112 M. R. Bond and J. J. Kohler, *Curr. Opin. Chem. Biol.*, 2007, **11**, 52–58.
113 N. Ferri, R. Paoletti and A. Corsini, *Biomarkers*, 2005, **10**, 219–237.
114 G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen and T. Carell, *J. Am. Chem. Soc.*, 2006, **128**, 1398–1399.
115 H. C. Kolb and K. B. Sharpless, *Drug Discovery Today*, 2003, **8**, 1128–1137.
116 R. Kumar, A. El-Sagheer, J. Tumpane, P. Lincoln, L. M. Wilhelmsson and T. Brown, *J. Am. Chem. Soc.*, 2007, **129**, 6859–6864.
117 A. D. Moorhouse, A. M. Santos, M. Gunaratnam, M. Moore, S. Neidle and J. E. Moses, *J. Am. Chem. Soc.*, 2006, **128**, 15972–15973.
118 A. M. Sadaghiani, S. H. Verhelst and M. Bogyo, *Curr. Opin. Chem. Biol.*, 2007, **11**, 20–28.
119 A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535–546.
120 A. E. Speers and B. F. Cravatt, *ChemBioChem*, 2004, **5**, 41–47.
121 E. Weerapana, A. E. Speers and B. F. Cravatt, *Nat. Protoc.*, 2007, **2**, 1414–1425.
122 W. P. Heal, S. R. Wickramasinghe and E. W. Tate, *Curr. Drug Discovery Technol.*, 2008, **5**, 200–212.
123 S. Kumar, B. Zhou, F. Liang, W. Q. Wang, Z. Huang and Z. Y. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 7943–7948.
124 M. C. Yee, S. C. Fas, M. M. Stohlmeyer, T. J. Wandless and K. A. Cimprich, *J. Biol. Chem.*, 2005, **280**, 29053–29059.
125 K. R. Love, R. K. Pandya, E. Spooner and H. L. Ploegh, *ACS Chem. Biol.*, 2009, **4**, 275–287.
126 M. Sunbul and J. Yin, *Org. Biomol. Chem.*, 2009, **7**, 3361–3371.

