View Online / Journal Homepage / Table of Contents for this issue

Organic & Biomolecular Chemistry

This article is part of the

OBC 10th anniversary themed issue

All articles in this issue will be gathered together online at

www.rsc.org/OBC10

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

Organic &
Biomolecular
Chemistry

Dynamic Article Links
View Online

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 5684

www.rsc.org/obc

PERSPECTIVE

Chemical and semisynthesis of posttranslationally modified proteins

Peter Siman and Ashraf Brik*

Received 19th January 2012, Accepted 26th March 2012
DOI: 10.1039/c2ob25149c

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

Posttranslational modifications of proteins play crucial roles in health and disease by affecting numerous aspects of protein structure, function, stability and sub cellular localization. Yet understanding the effects of these modifications on several of these processes at the molecular level has been hindered by the lack of homogeneously modified proteins obtained *via* traditional biochemical and molecular biology approaches. Moreover, the preparation of such bioconjugates at a workable level is highly demanding. Recent advances in protein chemistry applying chemical and semisynthetic approaches are becoming increasingly beneficial to overcome these challenges. These methods allow site-specific modifications of a desired protein and afford the product in large quantities for biochemical and structural analyses. In this review, we survey these efforts and their importance in dissecting the role of several posttranslational modifications in various proteins. Several examples are presented where glycosylated, phosphorylated, ubiquitinated, lipidated, acetylated and methylated proteins were prepared.

Department of Chemistry, Ben-Gurion University of the Negev, Beer Sheva, 84105, Israel. E-mail: abrik@bgu.ac.il

Introduction

The journey of cellular protein synthesis starts by transcription of the specific DNA to the mRNA followed by ribosomal translation to the corresponding polypeptide. This machinery produces a limited number of proteins *i.e.* ~5000 in yeast and up to

Peter Siman

Peter Siman received his BSc in Chemistry in 2007 and his MSc in 2009 from The Hebrew University of Jerusalem, where he worked under the guidance of Prof. Daniel Cohn on biomedical polymers and designing new tissue adhesives. Peter is currently a PhD student in Prof. Ashraf Brik's laboratory in the Ben-Gurion University of the Negev. His research interests include the development of novel tools for the synthesis and manipulation of highly aggregative proteins. Peter is the recipient of the Israel Ministry of Science and Technology Scholarship.

Ashraf Brik

Ashraf Brik received his BSc from Ben-Gurion University (BGU) in Israel, his MSc from the Technion-Isreal Institute of Technology, and his PhD in Chemistry from the Technion under the supervision of Prof. Ehud Keinan jointly with Prof. Philip E. Dawson from The Scripps Research Institute (TSRI). He received a postdoctoral fellowship from the Israel Science Foundation and spent 3 years as a postdoctoral fellow with Prof. Chi-Huey Wong at TSRI. He was promoted to a senior research associate and worked with Prof. Wong for an additional two years. In 2007, he returned to BGU as an Assistant Professor in the Chemistry Department. In 2011, he was promoted to Associate Professor. His current research involves the development of novel chemistries to prepare posttranslationally modified proteins for biological studies. He is the recipient of the Ma'of Fellowship and 2011 Israel Chemical Society prize for Outstanding Young Chemist.

This journal is © The Royal Society of Chemistry 2012

25 000 in humans, with a size ranging from ~70 to 5000 amino acids (~7 kDa–1.7 mDa). However, the complexity and diversity of cellular functions exceeds the number of proteins within a particular cell. Nature has compensated for the low number of genes through mRNA splicing at the transcriptional level and covalent modifications of proteins called posttranslational modifications (PTMs) which together increase the proteome diversity by more than 10–100 fold. PTMs occur after the translational step and are catalyzed by a number of enzymes that are highly specific for each modification. In general, PTMs largely affect protein structure, function, stability and subcellular localization.[1]

PTMs occur on the protein side-chains or its N-terminus, which are electron rich, through the addition of groups that are usually electron deficient. The nucleophilic residues, which are prone to these modifications are the thiolate anion of Cys, hydroxyl group of Ser and Thr, phenolic hydroxyl of Tyr, amine of Lys, His, Arg, and the carboxylate side chain of Asp and Glu (Scheme 1). Two distinct PTMs are Met oxidation to the sulfoxide form and glycosylation of Asn/Gln side-chain amide, which is considered a very weak nucleophile.[1]

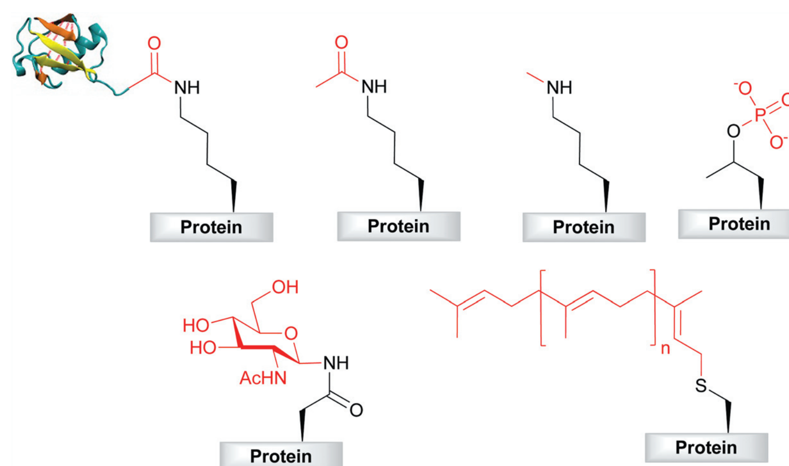
Phosphorylation is an example for a ubiquitous PTM, appearing in thousands of proteins and is considered one of the most studied PTMs. This modification affects the nature and behavior of a protein by modifying the polar hydroxyl side chain of Thr, Ser or Tyr, with a relatively bulky group of phosphate.[2] As a result, the radius and the surrounding of the modified residue are changed, which could affect the overall charge of the protein. Such modifications could, for instance, induce structural changes in the protein, promote protein–protein interactions through electrostatic interaction with the positively charged side-chains and switch on and off the catalytic activity of enzymes. While some PTMs can affect the functionality of modified protein, others interfere with protein subcellular localization and its fate. For instance, monoubiquitination can result in sending a membrane protein to recycling compartments in the trans Golgi network and to lysosomes, while the addition of a polyubiquitin chain targets proteins for proteasomal degradation.[3]

Unarguably, PTMs play major roles in the regulations of numerous of biological processes and are increasingly linked to different diseases. Hence, dissecting the roles of these modifications on the protein structure and function as well as in health and disease has been of great interest to chemists and biologists. A well-known example of the importance of understanding the role of PTMs and its impact on global health is related to kinases that catalyze phosphorylation. Over the past decade, several drugs that target these enzymes have been approved for clinical use in cancer and a number of inhibitors are currently in clinical trials for combating other diseases such as rheumatoid arthritis.[4]

In order to explore the role of PTMs at the molecular level, there is a great demand for methods that allow site-specific modifications to enable the production of posttranslationally modified proteins with high homogeneity and workable quantities. Such requirements cannot be achieved by traditional biological approaches due to the heterogeneity of the products and the low amounts that these strategies could afford. Moreover, these methods require the isolation of the specific enzyme, which is not often available for *in vitro* experiments. These challenges become even more demanding when one is aiming to study the interplay between different PTMs. Solutions to such obstacles are mainly emerging from two directions: the use of chemical and semisynthetic strategies and by applying advanced molecular biology approaches. The latter approach uses a unique codon and the corresponding tRNA/aminoacyl-tRNA synthetase pair to genetically encode unnatural amino acids, *e.g.* phosphorylated residue, in *Escherichia coli*, yeast, and mammalian cells. Alternatively, these approaches were also used to incorporate a certain chemical group that allows a site-specific modification to install PTM or its analogue. Such strategies will not be covered here as several excellent reviews were dedicated to this topic.[5] Recent advances in organic chemistry and chemical protein synthesis have made it possible to prepare posttranslationally modified proteins in excellent homogeneity and sufficient amounts for various studies. In this review, we survey these efforts in preparing these precious bioconjugates, highlighting examples from different areas of PTMs.

Chemical synthesis of proteins

Chemical synthesis of proteins is emerging as a very powerful approach for manufacturing of posttranslationally modified proteins.[6] The great advantage of such a method stems from the

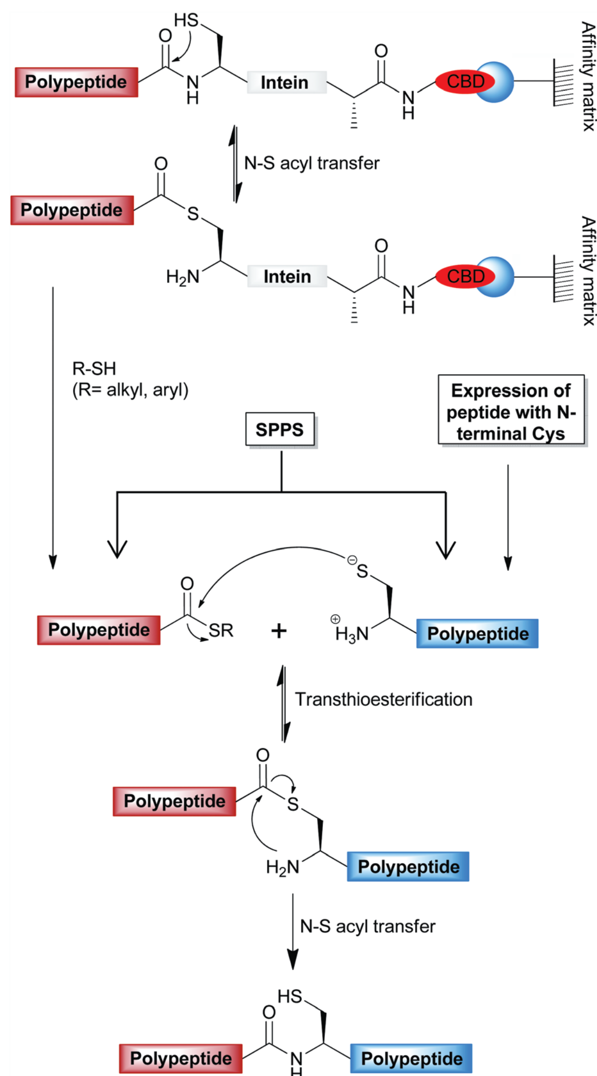


Scheme 1 Examples of different PTMs.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

ability to introduce essentially any modification at any position in the protein of interest. This is possible thanks to the chemoselective strategies that allow the construction of large proteins from smaller fragments obtained *via* solid phase peptide synthesis (SPPS). Unarguably, native chemical ligation (NCL) introduced by Kent and coworkers[7] is currently the method of choice to fulfil many of the requirements for the synthesis of small to middle-sized proteins (~200 residues).[8] This chemoselective method relies on an N-terminal Cys residue of a peptide fragment and the thioester functionality at the C-terminus of the other fragment to allow transthioesterification. The formed thioester intermediate undergoes spontaneous self-rearrangement, *via* S–N acyl transfer, resulting in the formation of native amide bond (Scheme 2).

The requirement for N-terminal Cys and its presence at a specific position in the sequence became a frequent limitation in applying NCL for protein synthesis. Hence, several directions were adopted to overcome these restrictions and enable the synthesis of proteins lacking the Cys residue(s).[9] The introduction of the desulfurization concept coupled with peptide ligation by Yan and Dawson expanded significantly the repertoire of ligation junctions and is becoming increasingly applicable in protein synthesis.[10] Today and thanks to the contributions of several research labs, including ours, ligation at Phe,[11] Val,[12] Leu,[13] Thr,[14] Lys,[15] Pro[16] and more recently Gln[17] junctions are now possible. Moreover, several groups have contributed to the development of different desulfurization conditions such as nickel borate/H_2,[10] Pd/Al_2O_3/H_2 [18] free metal conditions[12a,19] as well as performing ligation and desulfurization *in situ*.[20] Inspired by the NCL principles, S–N acyl transfer assisted by proximity affect[21] and desulfurization, a method called sugar-assisted ligation for glycopeptide and protein synthesis was also developed.[22] Together, all these methods have contributed in several ways to prepare proteins with PTMs and examples from different areas will be presented in the following sections.



Scheme 2 Principles of chemical and semisynthesis of proteins applying NCL and EPL.

Semisynthesis of proteins

The term "semisynthesis" refers to a protein preparation where at least one of the building blocks is obtained *via* expression while the other fragment(s) is chemically prepared. Expressed protein ligation (EPL), is a powerful method for protein semisynthesis and combines the principles of NCL, *in vitro* expression of proteins and protein splicing.[23] This method uses chemical and biological tools to manufacture the peptide precursors wherein NCL play a key role to bridge the peptide fragments (Scheme 2). For example, the thioester peptide could be obtained *via* intein-based expression[24] while the Cys fragment is chemically prepared to allow the introduction of any modification in such a fragment. Alternatively, a peptide with N-terminal Cys can be expressed while the thioester fragment is chemically synthesized to introduce the desired modification(s). Several methods were developed to efficiently generate a polypeptide with N-terminal Cys, including expression of a protein fused to cleavable tags such as TEV,[25] SUMO,[26] or a Met residue.[27] The removal of the tag/extension with TEV protease, SUMO protease, or endogenous MAP (Met amino peptidase), respectively, exposes the N-terminal Cys for the subsequent NCL step. One of the advantages of semisynthesis over chemical protein synthesis is the ability to prepare large proteins with a specific modification. Recently, the semisynthesis of the Eph receptor tyrosine kinase (96 KD protein), which required the expression of a thioester fragment comprising 517 residues was achieved, highlighting the power of the intein technology in the semisynthesis of large proteins.[28]

Preparation of naturally occurring posttranslationally modified proteins

Using the current methods of chemical and semisynthesis, the preparation of various proteins containing naturally occurring PTMs were achieved. These advances have allowed important studies on the role of PTMs on protein's structure, enzyme functionality and the effect on protein–protein, DNA–protein interactions. Notably, such a level of investigation was not possible using traditional methods. To give the reader an overview of the field rather than a comprehensive survey of a specific PTM, we selected examples that represent the use of these methods in

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

 This journal is © The Royal Society of Chemistry 2012

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

the preparations of a variety of posttranslationally modified proteins. This review also does not cover preparation of proteins with analogues of PTMs, a field that emerged as a result of the challenges in achieving the synthesis of a modified protein in its native form and to introduce stable derivatives of PTMs for different goals. To obtain PTM surrogates, methods like "click" chemistry, Michael addition of thiol to double bond or disulfide bond formation were applied to generate protein analogues bearing glycan, ubiquitin and lipids as well others PTMs.[29]

I. Glycosylated proteins

Glycosylation is one of the most used and complex PTM, which affects several physiochemical and functional properties of the modified protein. This modification could, for example, regulate protein turnover and immune responses, stabilize proteins against proteolysis, enhance protein solubility, fine-tune the charge and isoelectric point of proteins, and mediate interactions with pathogens.[30] Glycosylation could lead to a modified protein bearing as simple as a monosaccharide moiety or a very complex polysaccharide structure containing up to 13 different types of monosaccharide. Moreover, glycosylation could involve as much as eight different amino acids resulting in a wide variety of glycoproteins. Notably, at least 16 enzymes are known to be involved in such processes at various intracellular sites, which are encountered by several enzymes *i.e.* glycosidases that are known to cleave the *N*- or *O*-glycosidic bonds and play important biological roles and are crucial in the development of disease state.[31]

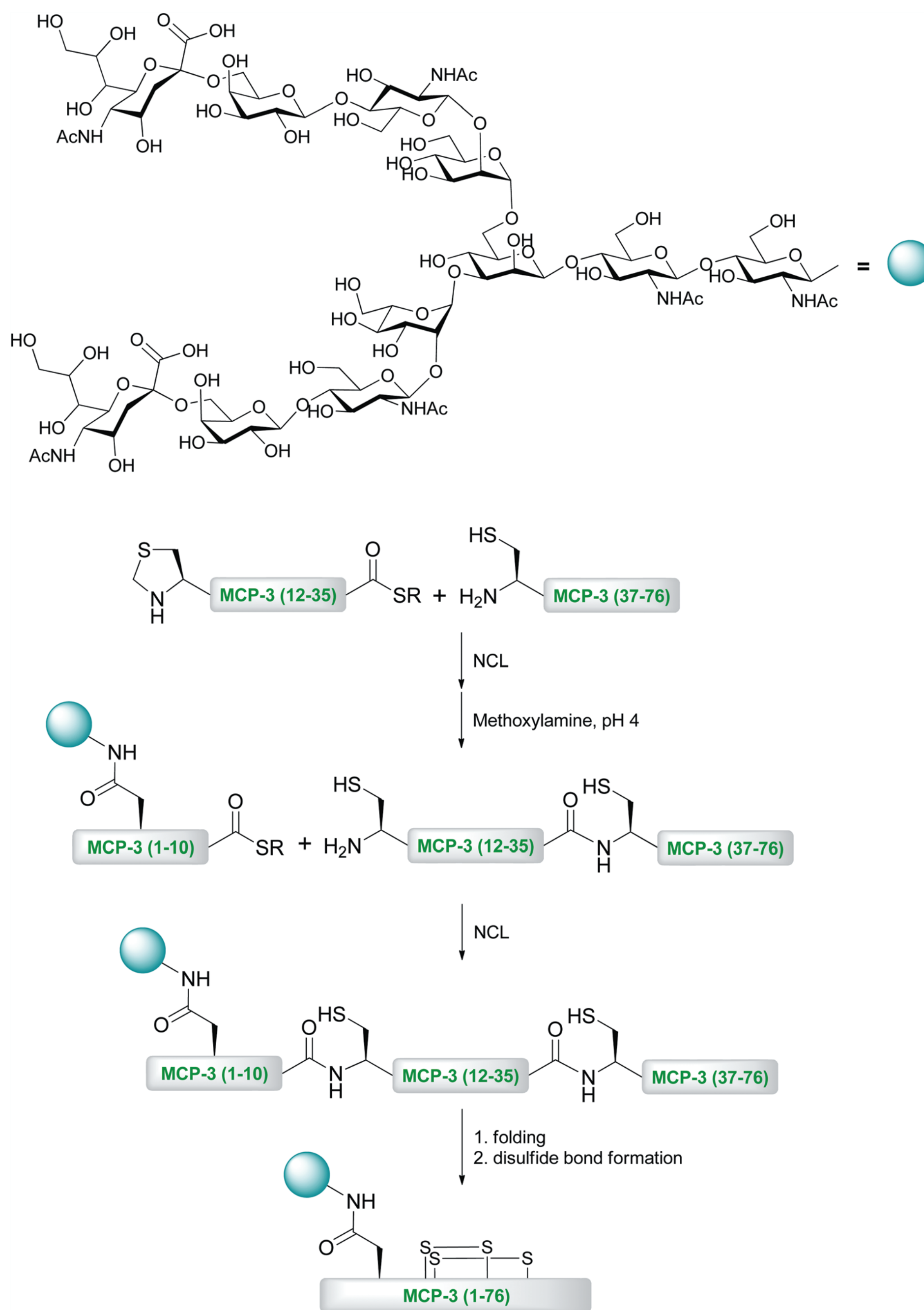
Since glycosylation involves enzymatic steps in complex biosynthetic pathways and is not controlled by a template, unpredictable glycoprotein products are often formed. This mixture possesses different glycoforms having the same polypeptide sequence. Hence, this issue challenges scientists, in particular when attempting to study the biological functions of glycoproteins and dissecting the role of glycosylation at the molecular level. Moreover, this complexity makes it very difficult to generate protein-based therapeutics, which in some cases are administered as a mixture of glycoforms *e.g.* erythropoietin. Therefore, there is a great demand for methods that facilitate the preparations of homogeneous materials to assist researchers in their efforts to understand the exact function/role of these modifications. Scientists have tackled these problems from different directions in which some of these merely based on chemical synthesis[32] while others combine biochemical approaches.[33] The readers are referred to excellent reviews on preparing the glycopeptides building blocks required for the synthesis and semisynthesis of glycoproteins.[34]

Chemical synthesis of monocyte chemotactic protein-3 glycoprotein. The preparation of glycosylated protein entails several challenges such as the synthesis of the oligosaccharide component(s) and the glycopeptide building blocks. The synthesis of the glycopeptide becomes even more complicated when both the thioester functionality and the glycan unit are present in the same peptide. This is due to the inherit instability of the saccharide part especially under Boc-SPPS conditions. The groups of Kajihara and Dawson overcame these obstacles by using Fmoc-SPPS to prepare such a precursor and applied it in the synthesis of monocyte chemotactic protein-3 (MCP-3), which consists of 76 residues and bearing two disulfide bonds.[35] This glycoprotein acts as chemotactic factor and activates a variety of inflammatory cells.[36] The need for accumulation of inflammatory cells requires a local production of a protein family called chemotactic cytokines (8–10 kDa), which induces directional migration of various cell types. For example, the family of MCP-1, MCP-2 and MCP-3 are known to chemottract and activate monocytes to the site of action.[37]

To synthesize the MCP-3 glycoprotein, the polypeptide sequence was divided into three fragments where the C-terminal and middle fragment bearing the thioester functionality, were synthesized applying Boc-SPPS chemistry (Scheme 3). The challenging step in this approach was the preparation of the N-terminal 10 mer peptide having both the thioester functionality and the *N*-glycosylated Asn. Two protocols for the synthesis of this peptide were examined. In the first approach, the authors protected the α-amine with the Boc-group while keeping the side-chains unprotected. To prevent coupling to the side-chains of Ser and Thr, mild coupling conditions were applied to give the desired thioester product in 13% isolated yield. The low yielding synthesis forced the group to adopt a second protocol based on Fmoc-SPPS with fully protected amino acids. In this strategy, the acid sensitive HMPB (4-(4-hydroxymethyl-3-methoxyphenoxy)-butyric acid) resin was used, which allowed the release of the peptide in the protected from. Subsequently, PyBOP/DIPEA was used to preform thioesterification with benzyl mercaptane, which upon a treatment with TFA, furnished the glycopeptide MCP-3(1–10)-thioester in 43% yield. Such approach may not be applicable for longer peptides as such peptides could have low solubility, which hampers further manipulations. The recent advances of preparing thioester peptides *via* Fmoc-SPPS, by applying for example the *N*-acyl urea chemistry overcome several limitations and should be applied successfully for such cases.[38]

With the three MCP-3 fragments in hand, the ligation steps were carried out smoothly followed by disulfide bonds formation to furnish the folded glycosylated MCP-3 (Scheme 3). The circular dichroism (CD) analysis of MCP-3 exhibited comparable spectra to the reported native form indicating a single α-helix and more poorly formed β-sheet, which characterizes this type of proteins. Chymotrypsin digestion followed by reduction of the disulfide bonds confirmed the correct protein sequence and structure. Moreover, the synthetic MCP-3 was recognized by the specific monoclonal antibody using ELISA assay, which further supports the structure integrity of the glycosylated MCP-3.

Semisynthesis of glycoprotein CAM-1. Using semisynthesis, Macmillan and Bertozzi prepared multi-glycosylated protein CAM-1, a mucin-like glycoprotein that functions as a ligand for the leukocyte adhesion molecule L-selectin.[39] This protein has two distinct motifs, which are separated by mucin domain located at the middle region, that undergo *N*-acetyl galactosamine (GalNAc) α-*O*-glycosylation on the hydroxyl groups of Ser/Thr residues. In this work, three glycosylated forms of CAM-1 were prepared having glycosylated residues at the N-terminal, C-terminal (Scheme 4) and simultaneously N- and C-termini regions (Scheme 5).



Scheme 3 Total chemical synthesis of MCP-3.

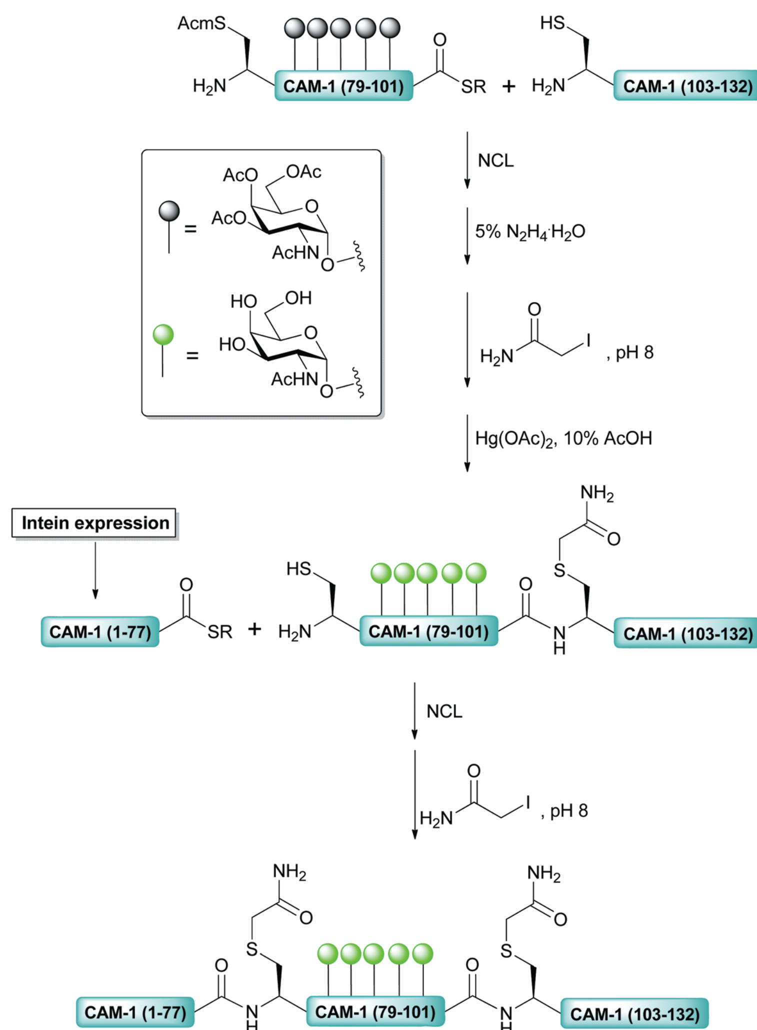
Notably, the middle fragment was prepared by the intein mediated thioester strategy in which the N-terminal Cys was masked with Ile-Glu-Gly-Arg extension to avoid unspecific ligations or cyclization. After ligation with the C-terminal fragment, the Xa protease was used to expose the N-terminal Cys, allowing for a second ligation step (Scheme 5). Notably, the glycosylated amino acids [Fmoc-Thr(α-*O*-GalNAc(OAc)$_3$-OH and Fmoc-Ser(α-*O*-GalNAc(OAc)$_3$-OH] were prepared in solution and coupled directly to the peptide bound resin. Another feature of the synthesis is the use of non-native Cys (originally Gln residues) in the middle and C-terminal fragments for the ligation steps. After ligation of the three fragments, the resulting glycosylated protein was subjected to removal of the glycan acetyl protecting groups followed by treatment with iodoacetamide to generate an alkylated Cys as a glutamine mimic.

Using chemical and semisynthetic approaches, several other glycosylated proteins such as diptericin ε protein,[18,40] bacterial immunity protein Im7,[41] ribonuclease C,[42] β-subunit of the human follicle-stimulating hormone,[43] erythropoietin analogue,[44] human glycosyl-interferon-β[45] and α-subunit of human glycoprotein hormones[46] were prepared, which further testifies to the great potential of these methods in this field.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

 This journal is © The Royal Society of Chemistry 2012

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C



Scheme 4 Semisynthesis of C-terminal multi-glycosylated CAM-1.

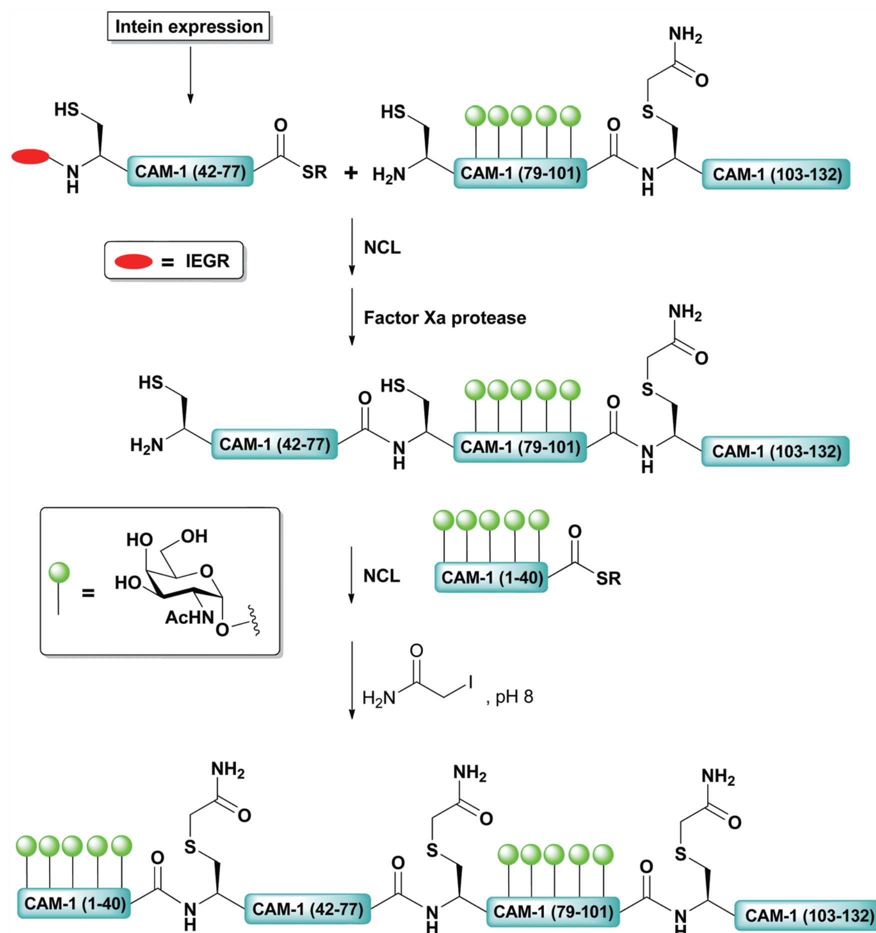
II. Phosphorylated proteins

Around one third of the human proteins are estimated to undergo phosphorylation, a PTM that is being catalyzed by around 500 kinases. The main residues that are involved in phosphorylation are Ser, Thr and Tyr. However, other amino acids such as Arg, Glu and His, could also undergo phosphorylation, albeit to a lesser extent. In the His case, for example, it was found that the nitrogen of the imidazole ring could be targeted for phosphorylation, acting as a catalyst for transferring the phosphate group to other side chains and affecting several cellular processes.[47] Phosphorylation can modify the protein by a single phosphate or as much as eleven phosphate groups as in the case of ABL tyrosine kinases.[48]

Semisynthesis of phosphovariants of Smad2 protein. The Smad protein family is involved in the activation cycle of transforming growth factor-β (TGF-β) and its family members, resulting in proliferation, differentiation, migration and apoptosis of tissue in fruits, flies up to human.[49] This cycle of activation starts with a complex formation between the dimeric TGF-β (active form) and tetrameric cell surface receptor that consists of TβRII and TβRI kinases. A third protein called the Smad anchor for receptor activation (SARA) further stabilizes this complex. Here, the binding of TGF-β activates TβRII to phosphorylate TβRI in multiple Ser and Thr residues. Upon phosphorylation, TβRI phosphorylate receptor activated-Smad (R-Smad), which is a common motif found in all Smads. Subsequently, R-Smad detaches from this complex and is translocated into the nucleus where it regulates transcriptional processes. Yet, several questions are still open regarding to the mechanism of Smads phosphorylation, the effect of phosphorylation on the detachment of Smads form TβRI protein and the effect of Ser phosphorylations on the interactions of Smads with the heterocomplex.

To shed light on these questions, Muir and coworkers adopted a semisynthetic approach to prepare phosphorylated Smad2 analogues.[50] Smads consist of a conserved region known as MH1 (residues 241–462) and a small C-terminal fragment known as MH2 (CSSXS) that varies at the X position within the Smad family, being Met in Smad2. The MH1 region was prepared using intein mediated protein thioester formation, while the C-terminal fragment was prepared using SPPS to introduce

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C



Scheme 5 Preparation of simultaneously N- and C-termini multi-glycosylated CAM-1 using EPL.

phosphoserine at the desired sites. By ligating MH1 and MH2 peptides, using NCL, the group was able to prepare three phosphovariants and non-phosphorylated form (Scheme 6).

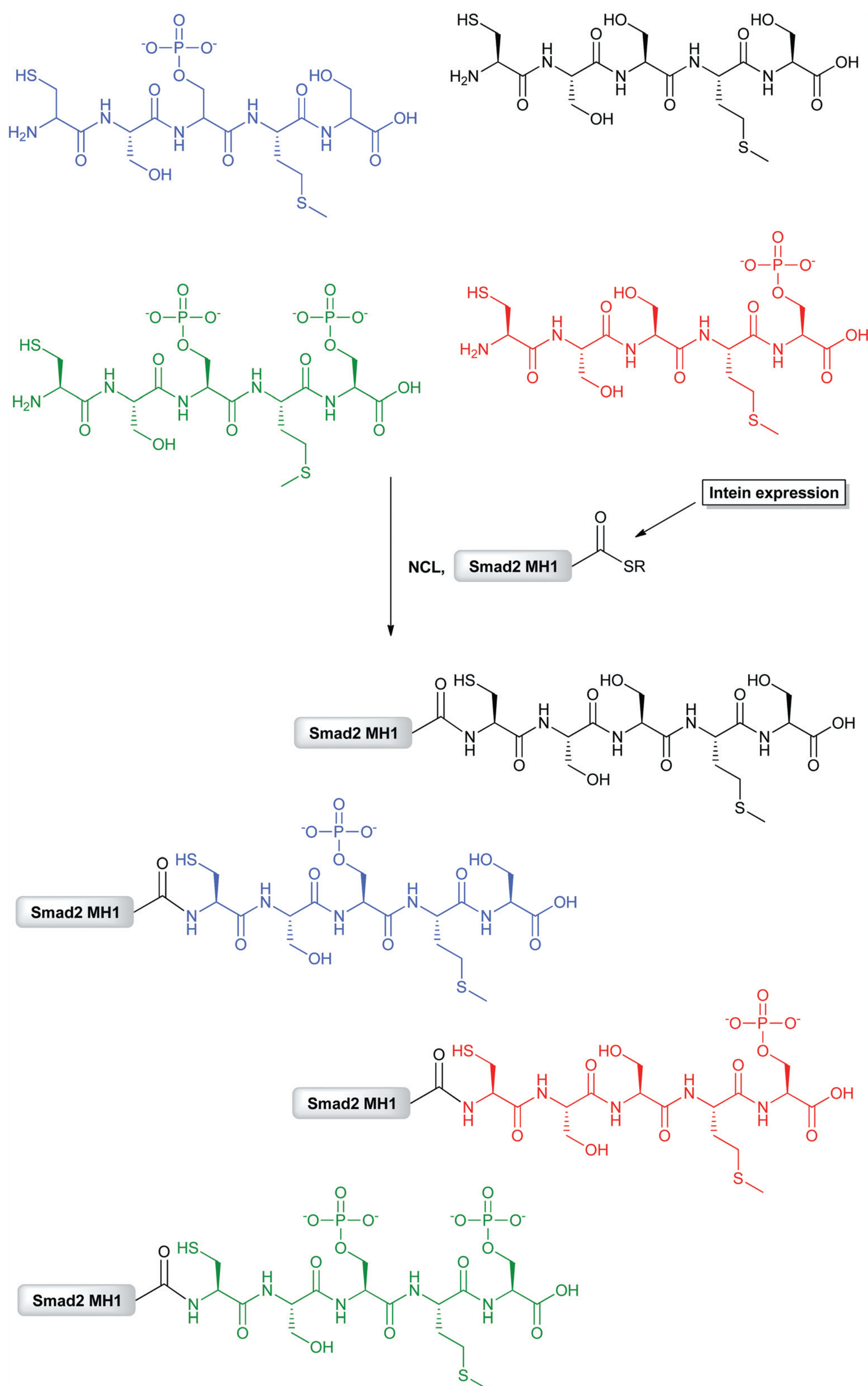
This interesting study revealed that phosphorylation at Ser465 accelerates the second phosphorylation step at Ser467 within Smad2. In addition, it was found that in order for Smad2 to detach from the complex it should form a stable oligomer where double phosphorylation drives it towards the oligomeric form for the translocation step. Notably, purification after EPL was assisted by SARA protein, which interacted strongly with the Smad2 allowing its fishing through ultracentrifugation method. These results showed that this strong interaction assists in stabilizing the *in vivo* complex to prevent premature release of the mono-phosphorylated Smad2.

Semisynthesis of serotonin *N*-acetyltransferase enzyme. In another interesting study, semisynthesis of protein was applied to prepare the phosphorylated enzyme serotonin *N*-acetyltransferase, also known as arylalkylamine *N*-acetyltransferase (AANAT).[51] Changes in the activity of this enzyme affect the levels of melatonin hormone in the brain. AANAT undergoes phosphorylation at two sites, Thr31 and Ser205, which was proposed to affect the interactions with the regulatory protein of the family 14-3-3 (14-3-3ζ) that is also known to interact with other signaling proteins. This interaction appears to stabilize AANAT against degradation and helps in modulating its acetyltransferase activity. However, it is not clear how these phosphorylations affect this interaction at the molecular level. Hence, several phosphorylated analogues of AANAT and fluorescently labeled protein 14-3-3ζ were prepared in order to study the AANNAT–14-3-3ζ interaction. To prepare AANAT protein, the polypeptide sequence was divided to three fragments; AANAT(1–33), AANAT(34–199) and AANAT(200–207). The middle fragment bearing C-terminus thioester was prepared using intein based expression. This peptide was also equipped with N-terminal extension of Ile-Glu-Gly-Arg peptide or SUMO to prevent polymerization or cyclization during the ligation step. However, two additional residues (Gly-Cys) were inserted into the AANAT sequence, between residues 33–34, to allow NCL. Using this strategy, the Cole group was able to obtain five analogues of AANAT with GC insertion: unmodified AANAT, pS205–AANAT, pT31–AANAT, (pT31 + pS205)–AANAT and the labeled AANAT.

Interestingly, this study revealed that phosphorylation of AANAT either at Thr31 or at both sites *i.e.* Thr31 + Ser205 competed efficiently with a synthetic peptide derived from AANAT, which is known to bind to 14-3-3ζ with K_D of 3.2 μM. However, mono-phosphorylated AANAT at Ser205 exhibited reduced affinity. In addition, they were able to show that the doubly phosphorylated AANAT interacts simultaneously with the two

 This journal is © The Royal Society of Chemistry 2012

View Online



Scheme 6 Semisynthesis of phosphorylated Smad2 analogues.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

monomers in the dimer 14-3-3 ζ protein while in the case of mono-phosphorylation, two AANAT molecules are required. Notably, phosphorylation at Ser205 resulted in a tighter interaction with 14-3-3ζ compared to modification at Thr31, despite that this residue is located in a canonical 14-3-3ζ binding sequence.

Applying chemical and semisynthetic approaches several other groups were able to prepare other phosphorylated proteins

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

such as Src kinase,[24a] histone H2B,[52] histone H3,[53] tau protein,[54] and more recently α-synuclein.[55]

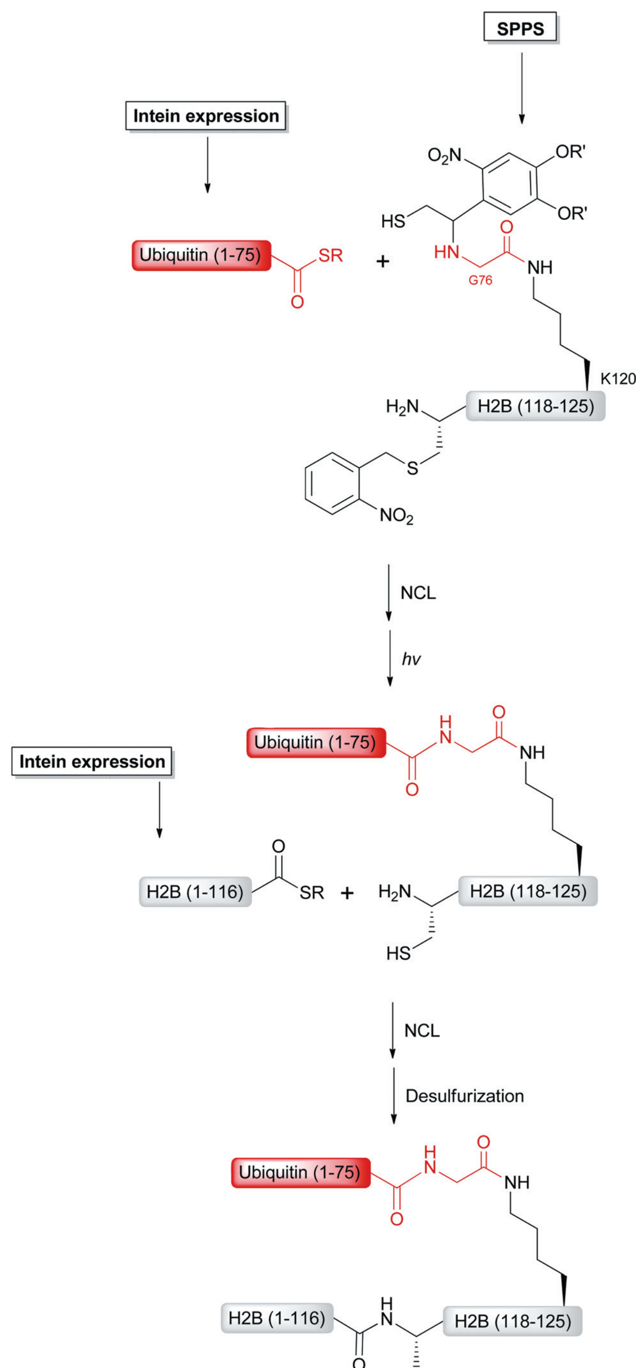
III. Synthesis and semisynthesis of mono-ubiquitinated proteins

PTMs could occur *via* the addition of a small molecule *e.g.* phosphate, acetyl or a protein as in the case of ubiquitination—the attachment of ubiquitin protein to a specific target.[56] Ubiquitination is a remarkable PTM that affects a variety of biological pathways such as transcription, protein degradation and endocytosis. Ubiquitin can be conjugated to a protein in the form of a monomer or as chains of different lengths and linkage types. Specifically, ubiquitin is attached to a target protein using a highly controlled enzymatic machinery of three enzymes known as the E1–E3, resulting in an isopeptide linkage between the ubiquitin C-terminal Gly and the Lys side-chain of a protein substrate. This process can repeat itself to from a polyubiquitin chain, in which any one of the seven Lys in ubiquitin (*i.e.* K6, K11, K27, K29, K33, K48, K63) can be used to anchor the consecutive ubiquitin. Interestingly, the signaling outcome depends on the type of the chain attached to a specific protein. For example, while K48-linked chain signals for protein degradation,[57] the K33-chain is involved in immune regulation.[58]

In order to achieve highly homogeneous ubiquitin bioconjugates such as ubiquitinated peptides and proteins with ubiquitin chains of varying lengths and types, it is crucial to replicate the enzymatic process *in vitro*. Hence, it is important to identify and isolate the specific E2/E3, which is often a formidable task. Moreover, these enzymes when functioning *in vitro* often operate in a non-specific manner and give the product in very small amounts, hindering studies aiming at dissecting the role of ubiquitin on protein structure and function. As a result, there has always been a great demand to develop methods to overcome these challenges in order to reconstitute homogeneous ubiquitinated peptides or proteins. These challenges have been a major driving force for the development of chemical strategies to complement biological methods in these efforts. The field of chemical synthesis of ubiquitin bioconjugates has made important progress in recent years. Several ubiquitinated peptides and proteins as well as anchored and unanchored ubiquitin chains were prepared for a variety of structural and biochemical studies.[8b,15,59]

Semisynthesis of ubiquitinated histone H2B. The massive DNA structure is wrapped and packed into a small and compact size unit within the nucleus. This process is facilitated by nucleosome proteins, *e.g.* histones, in which their assembly along with the DNA forms the chromatin. In a typical nucleosome, the DNA is wrapped into two turns around the histone octamer core, which consists of four histone partners; H3–H4 tetramer and two H2A–H2B dimers. The histone proteins are known to undergo a variety of PTMs such as phosphorylation, methylation, acetylation and ubiquitination, which play crucial roles in regulating chromatin dynamics, gene expression and DNA repair.[60]

Muir and coworkers have recently reported the semisynthesis of monoubiquitinated H2B to shed light on the role of H3 Lys79 methylation executed by Lys79-specific methyltransferase DotLl (disrupter of telomeric silencing-like).[59b] The group



Scheme 7 Semisynthesis of mono ubiquitinated H2B.

successfully prepared monoubiquitinated H2B assisted by an auxiliary mediated ligation between ubiquitin(1–75)-thioester to H2B (117–124) fragment (Scheme 7).[59a] The ligation step using this approach required 5–7 days due to the involvement of secondary amine in the acyl transfer step, which afforded the ligation product in a low yield and limited the generality of this method. After the ligation step, the auxiliary was removed by UV irradiation, simultaneously with the deprotection of 2-nitrobenzyl protecting group of the N-terminal Cys. Subsequently, this was ligated with N-terminal H2B(1–116)-thioester, obtained *via* intein based expression, to give after a desulfurization step

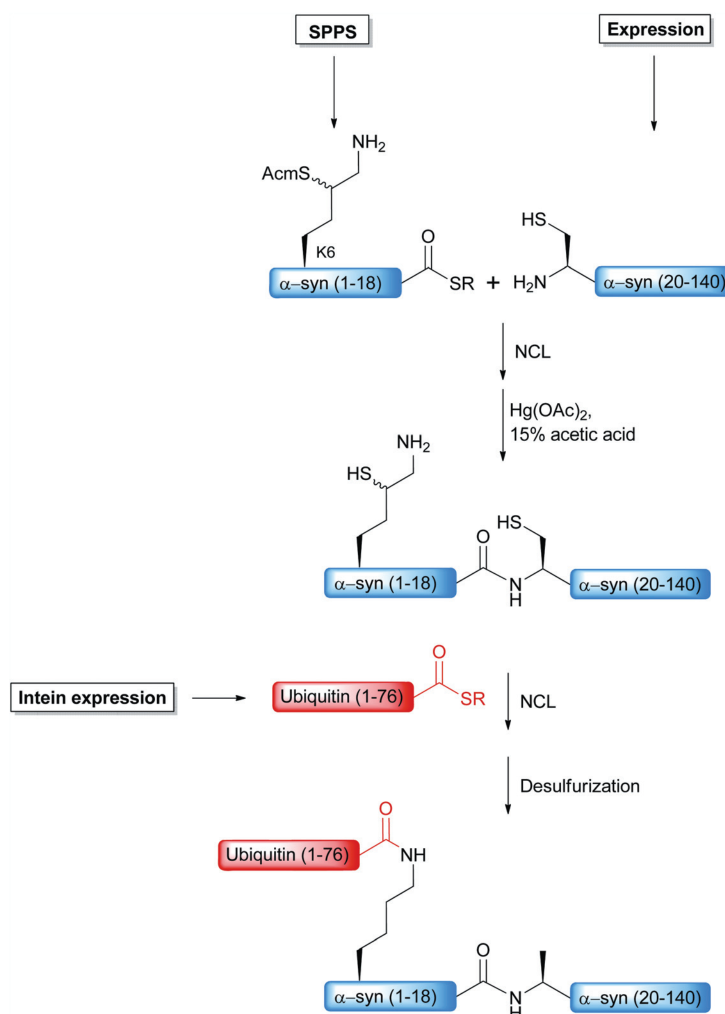
 This journal is © The Royal Society of Chemistry 2012

the desired monoubiquitinated H2B. The semisynthetic version of ubiquitinated H2B was successfully incorporated into the core histone octamers. In addition, the group prepared histone octamers without the ubiquitinated H2B as well as dinucleosomes with asymmetric insertion of ubiquitinated H2B and H3 (K79R). Studying these constructs revealed that ubiquitinated H2B directly activates methylation of H3 K79 by Dot1L and that efficient methylation of H3 K79 requires the presence of ubiquitinated H2B in the same nucleosome.

Semisynthesis of ubiquitinated α-synuclein protein. In a joint effort with Lashuel's group, we recently reported the semisynthesis of site-specifically mono-ubiquitinated α-synuclein (α-syn).[61] This protein constructs most of the Lewy bodies (LBs), which are responsible for the loss of dopaminergic neurons, resulting in several neurodegenerative diseases such as Parkinson's disease (PD). It has been reported also that these LBs contain mono or di-ubiquitinated α-syn at multiple sites; however, the exact role of ubiquitination on α-syn structure, function and its aggregation is still unclear. Hence, we reasoned that obtaining highly homogeneous ubiquitinated α-syn would be the first step to answer some of the open questions related to the role of ubiquitination on α-syn.

The semisynthesis of monoubiquitinated α-syn at Lys6 is shown in Scheme 8. In our approach, the C-terminal α-syn (19–140) fragment was prepared recombinantly. Due to the absence of Cys residues in the sequence of α-syn, Ala19 was replaced with Cys to allow NCL. On the other hand, the N-terminal fragment, α-syn(1–18) bearing δ-mercaptolysine[15a,62] at position 6, was prepared using Boc-SPPS. The thiol handle in the δ-mercaptolysine was protected with Acm[63] group to allow ligation with Ub-thioester following the full assembly of α-syn polypeptide. Final desulfurization step to remove the thiol handle from the δ-mercaptolysine along with the conversion of Cys19 to Ala gave the ubiquitinated α-syn in an excellent yield and quantity for biophysical characterizations. Interestingly, we found that ubiquitination significantly inhibited aggregation by stabilizing the monomeric form of this protein and without affecting its folding and membrane interactions. Moreover, we found that this modification did not affect phosphorylation at Ser87 and Ser129 by several kinases known to phosphorylate α-syn at these residues.

In addition to the methods that were introduced to prepare the native isopeptide bond,[15,59a,f,h] several other strategies were developed to construct isopeptide bond mimics.[29h,64] In this regard, click chemistry and disulfide bond formation were used



Scheme 8 Preparation of ubiquitinated α-synuclein protein using EPL.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

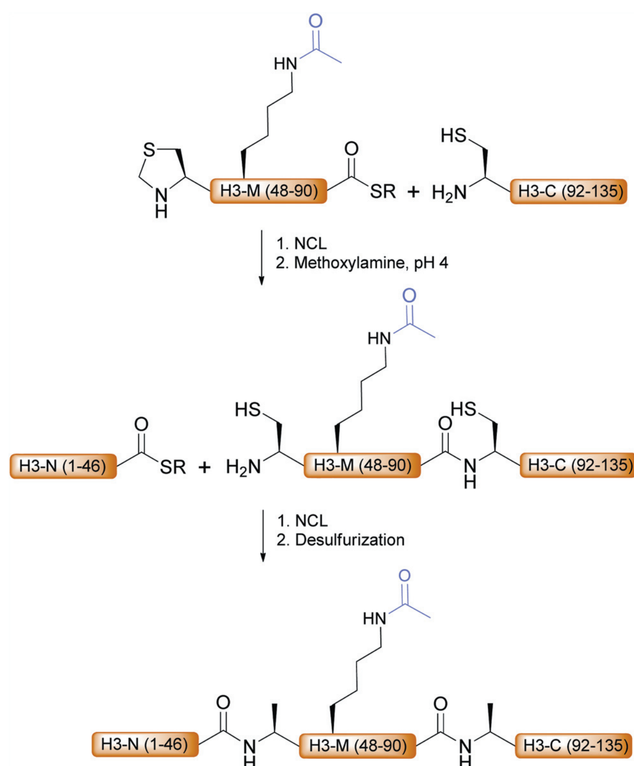
to prepare ubiquitinated H2B,[64c] proliferating cell nuclear antigen (PCNA)[64d,65] and SUMOylated SUMO-specific E2 enzyme (Ubc9).[29h,66]

IV. Acetylated proteins

Acetylation is referred to the addition of C_2 moiety (acetate group) to ε-amine of Lys. Acetylation can also be generalized to include lipidation, which will be described in the following section. Introduction of such group neutralizes the Lys charge, hence affecting protein function and/or its interaction.

Chemical synthesis of Lys56 acetylated histone H3. The work reported recently by Ottesen and coworkers to prepare acetylated H3 protein at Lys56 (the largest among other histones) emphasizes the importance of such a modification over the accessibility to DNA in chromatin.[45] Acetylation causes unpacking of the chromatin structure resulting in unwrapping of DNA to allow different actions to take place *e.g.* DNA repair, maintenance of genomic stability, and transcriptional regulation.

The H3 sequence lacks Cys residues necessary for the chemical synthesis of this protein using NCL. As a result, the group introduced Cys substitutions into the sequence to assist such synthesis. Initially, two Cys residues were introduced into H3 at positions Arg40 and Ser96 based on homology alignments that found H3(R40C) in *Cairina moschata* and H3(S96C) in the H3.1 variants in Homo sapiens, *Mus musculus*, and *Caenorhabditis elegans*. However, these substitutions affected nucleosome structure and dynamic. This led the researchers to modify their strategy to include substitution of Ala residues at positions 47 and 91 with Cys to enable ligation followed by desulfurization



Scheme 9 Chemical synthesis of acetylated H3 protein.

(Scheme 9). Accordingly, three different proteins; native H3, H3 (K56ac) and H3 (K56Q) were prepared. Using reconstituted nucleosome with these analogues and DNA labeled with FRET pair at the entry–exit region, the study revealed that Lys56 is a key residue in the mechanism of unwrapping DNA, resulting in opening this highly condensed structure of chromatin to allow translation of the DNA. This was attributed to the minimization of the interactions between this Lys and DNA base pairs adjacent to the entry region. One particular advantage of applying total synthesis in preparing any of the histone proteins is the ability to introduce multiple PTMS to study cross talks between these modifications in regulating chromatin structure and function.

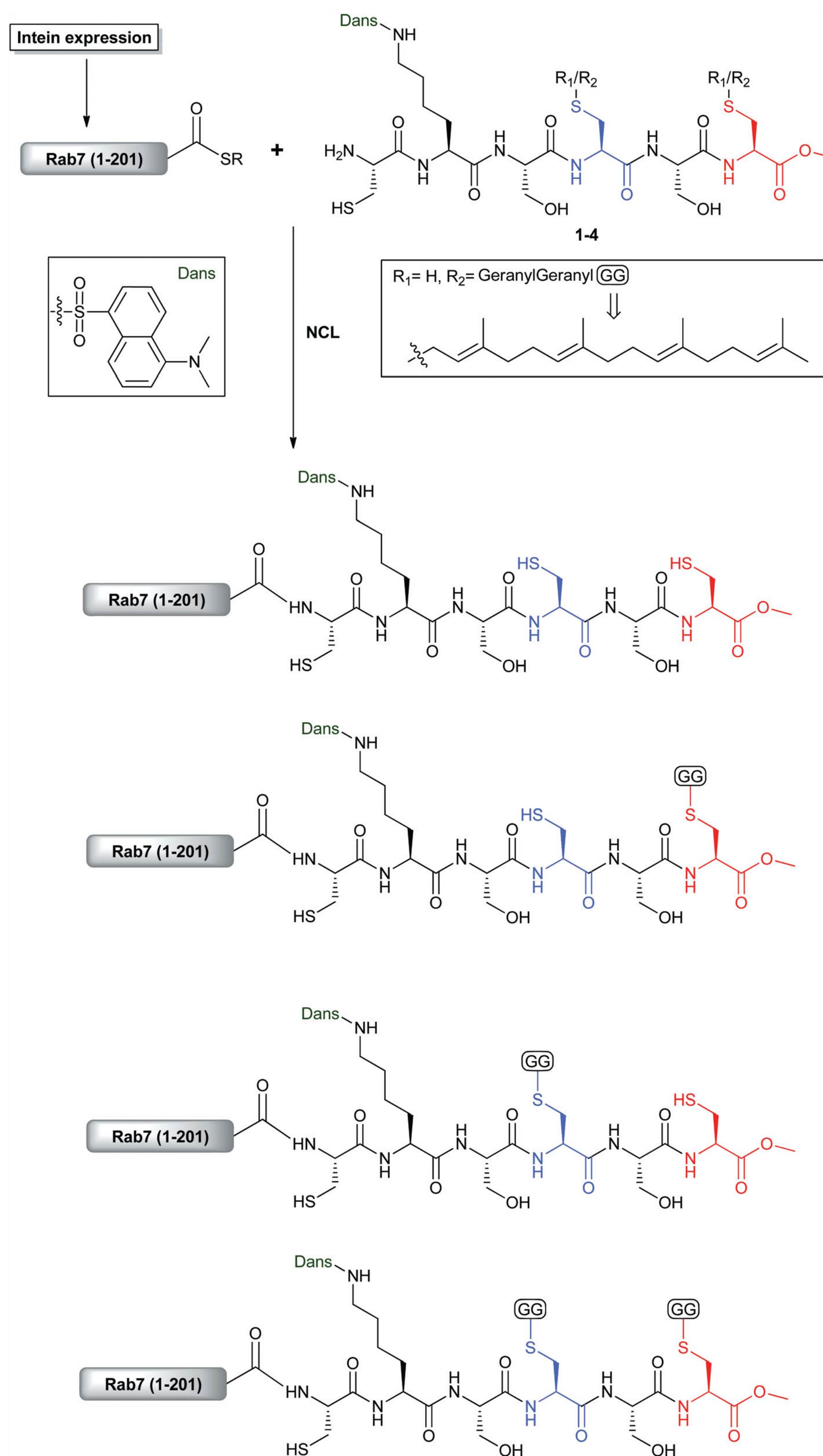
Semisynthesis of acetylated histone H3 at the nucleosome dyad. In the same spirit of the above discussed work, Ottesen and coworkers explored the role of lysine acetylation on the stability of the DNA-histone *i.e.* chromosome.[67] The researchers introduced acetylation into two sites of H3, Lys115 and Lys122, which are located at the dyad of H3–H4 and DNA. The study revealed that the nucleosome position within the genome was not altered by H3-Lys115 and/or H3-Lys122 acetylation. However, these modifications were found to reduce the free energy of binding between the histone octamer and a well-defined nucleosome positioning DNA sequence. This supports that this modification influences also the DNA unwrapping from the histone octamer resulting in an indirect effect over the entry–exit region. Notably, this study also revealed that substitution of a Lys residue by Gln as acetylation mimic in nucleosome reflects well the charge state of acetylated Lys but not the steric hindrance.[68]

VI. Lipidated proteins

Lipidation of proteins is processed through enzymatic machinery that introduces alkyl groups varying form C_1–C_{15} and C_{20} isoprenyl (*e.g.* farnesyl and geranylgeranyl) groups on Lys, Arg and Cys side-chains.[69] These alkyl groups serve mainly to increase the hydrophobicity of a protein and neutralize the amine positive charge. Hence, such a hydrophobic chain directs the target protein into the membrane insertion and localization. McCafferty and coworkers were the first to obtain site-specific methylation and/or acetylation of H3 and H4 proteins using NCL combined with EPL[70] which paved the way for preparing other modified histone proteins. A unique type of lipidation occurs by modifying the Cys side-chain with C_{15} or C_{20} fatty acid referred as S-prenylation as in Ras GTPase protein.[1,71]

Synthesis of mono- and di-prenylated Rab7 GTPase protein. Rab7 GTPase belongs to the Rab family of proteins, which is responsible for regulating several processes related to membrane trafficking. Prenylation of Rab7 GTPase involves the attachment of either farnesyl (C_{15}) or geranylgeranyl (C_{20}) moiety, *via* a thioether linkage, to one or two Cys residues located at the C-terminal of the target protein. This modification is catalyzed by three different prenyl transferases including protein farnesyltransferase (FTase), protein geranylgeranyl transferase-I (GGTase-I) and Rab geranylgeranyl transferase (RabGGTase or GGTase-II). However, the synthesis of prenylated proteins through Cys thiolate proved to be a very challenging task due to side reactions,

 This journal is © The Royal Society of Chemistry 2012



Scheme 10 Preparation of lipidated Rab7 protein using EPL.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

which could occur on the thioether linkage.[72] Waldmann and co-workers accomplished the synthesis of fluorescently labeled Rab7 GTPase[73] known to undergo prenylation at two Cys residues (205 and 207). This synthesis aimed to explore Rab7 GTPase function, the role of prenyl modification on their biogenesis and to determine the effect of the Rab escort protein (Rep), which forms a complex with Rab GTPase during regulation processes. To prepare this protein, a semisynthesis approach was

applied along with solution synthesis of hexapeptide bearing the two alkylated Cys residues (**1–4**), (Scheme 10). The hexapeptide with mono/di-prenyl groups on the Cys205/207 was also equipped with N-terminal Cys for ligation with the N-terminal recombinant fragment bearing C-terminus thioester moiety. Notably, the addition of an appropriate detergent with the appropriate size, such as sodium dodecyl sulfate (SDS) or cetyltrimethylammonium bromide (CTAB), was necessary to enhance the solubility of the alkylated hexapeptide analogues (**1–4**) and increasing the effective concentration of the reactants to achieve efficient ligation and enable the preparation of the Rab GTPase analogues.

The prepared lipidated proteins were subjected to biophysical and biochemical analyses to explore the interactions with RabGGTas. This study revealed that the prenylated protein at Cys207 interacts with RabGGTase in a higher affinity compared to the Cys205 prenylated analogue. These findings were attributed to the association of the conjugated isoprenoid with the active site of RabGGTase, which is less sterically favorable when it is located at Cys205. This is due to the requirement to initially accommodate the C-terminal Ser and Cys207 residues in the active site of RabGGTase, which leads to the fixation of the unprenylated Cys205 in the active site resulting in acceleration of its prenylation.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C

Summary and outlook

In this review, we have presented several examples of various posttranslationally modified proteins wherein synthetic and semisynthesis methods were applied to obtain highly homogeneous targets. These studies enabled dissecting the role of the PTMs on biological function as well as on protein structure. With the growing evidence of the importance of PTMs in health and disease, studying their effect at the molecular levels will continue to be one of the most exciting areas in chemical biology. Such knowledge could lead to the identification of new therapeutic targets that are involved in PTM regulation. The next decade will witness the synthesis of more challenging posttranslationally modified proteins and new mysteries of these modifications will be unraveled. In such a journey the development of new chemical methods will remain to be an important aspect of the field and chemists will be more challenged to accomplish these targets.

References

1 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto, *Angew. Chem., Int. Ed.*, 2005, **44**, 7342–7372.
2 L. N. Johnson and R. J. Lewis, *Chem. Rev.*, 2001, **101**, 2209–2242.
3 A. M. Weissman, N. Shabek and A. Ciechanover, *Nat. Rev. Mol. Cell Biol.*, 2011, **12**, 605–620.
4 M. E. M. Noble, J. A. Endicott and L. N. Johnson, *Science*, 2004, **303**, 1800–1805.
5 C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413–444.
6 (*a*) P. E. Dawson and S. B. H. Kent, *Annu. Rev. Biochem.*, 2000, **69**, 923–960; (*b*) C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030–10074; (*c*) K. S. A. Kumar and A. Brik, *J. Pept. Sci.*, 2010, **16**, 524–529.
7 P. E. Dawson, T. W. Muir, l. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776–779.
8 (*a*) V. Y. Torbeev and S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2007, **46**, 1667–1670; (*b*) K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon and A. Brik, *Angew. Chem., Int. Ed.*, 2011, **50**, 6137–6141.
9 (*a*) G. S. Beligere and P. E. Dawson, *J. Am. Chem. Soc.*, 1999, **121**, 6332–6333; (*b*) E. Saxon, J. I. Armstrong and C. R. Bertozzi, *Org. Lett.*, 2000, **2**, 2141–2143; (*c*) B. L. Nilsson, L. L. Kiessling and R. T. Raines, *Org. Lett.*, 2000, **2**, 1939–1941; (*d*) D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent and P. Botti, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 6554–6559; (*e*) J. Offer, C. N. C. Boddy and P. E. Dawson, *J. Am. Chem. Soc.*, 2002, **124**, 4642–4646; (*f*) J. W. Bode, R. M. Fox and K. D. Baucom, *Angew. Chem., Int. Ed.*, 2006, **45**, 1248–1252; (*g*) A. Brik and C. H. Wong, *Chem.–Eur. J.*, 2007, **13**, 5670–5675; (*h*) R. J. Payne, S. Ficht, W. A. Greenberg and C.-H. Wong, *Angew. Chem.*, 2008, **120**, 4483–4487; (*i*) M.-Y. Lutsky, N. Nepomniaschiy and A. Brik, *Chem. Commun.*, 2008, 1229–1231; (*j*) R. Okamoto, S. Souma and Y. Kajihara, *J. Org. Chem.*, 2009, **74**, 2494–2501; (*k*) F. B. Dyer, C.-M. Park, R. Joseph and P. Garner, *J. Am. Chem. Soc.*, 2011, **133**, 20033–20035.
10 L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526–533.
11 D. Crich and A. Banerjee, *J. Am. Chem. Soc.*, 2007, **129**, 10064–10065.
12 (*a*) C. Haase, H. Rohde and O. Seitz, *Angew. Chem., Int. Ed.*, 2008, **47**, 6807–6810; (*b*) J. Chen, Q. Wan, Y. Yuan, J. Zhu and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2008, **47**, 8521–8524.
13 (*a*) Z. Harpaz, P. Siman, K. S. A. Kumar and A. Brik, *ChemBioChem*, 2010, **11**, 1232–1235; (*b*) Z. Tan, S. Shang and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2010, **49**, 9500–9503.
14 J. Chen, P. Wang, J. Zhu, Q. Wan and S. J. Danishefsky, *Tetrahedron*, 2010, **66**, 2277–2283.
15 (*a*) K. S. A. Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel and A. Brik, *Angew. Chem., Int. Ed.*, 2009, **48**, 8090–8094; (*b*) R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu and C.-F. Liu, *J. Am. Chem. Soc.*, 2009, **131**, 13592–13593.
16 (*a*) S. Shang, Z. Tan, S. Dong and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2011, **133**, 10784–10786; (*b*) H. Ding, A. Shigenaga, K. Sato, K. Morishita and A. Otaka, *Org. Lett.*, 2011, **13**, 5588–5591.
17 P. Siman, S. V. Karthikeyan and A. Brik, *Org. Lett.*, 2012, **14**, 1520–1523.
18 Y. Y. Yang, S. Ficht, A. Brik and C. H. Wong, *J. Am. Chem. Soc.*, 2007, **129**, 7690–7701.
19 Q. Wan and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248–9252.
20 P. Siman, O. Blatt, T. Moyal, T. Danieli, M. Lebendiker, H. A. Lashuel, A. Friedler and A. Brik, *ChemBioChem*, 2011, **12**, 1097–1104.
21 D. M. Coltart, *Tetrahedron*, 2000, **56**, 3449–3491.
22 A. Brik, S. Ficht, Y.-Y. Yang and C.-H. Wong, *J. Am. Chem. Soc.*, 2006, **128**, 15026–15033.
23 T. W. Muir, *Annu. Rev. Biochem.*, 2003, **72**, 249–289.
24 (*a*) T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705–6710; (*b*) T. C. Evans, Jr., J. Benner and M.-Q. Xu, *Protein Sci.*, 1998, **7**, 2256–2264.
25 T. J. Tolbert and C.-H. Wong, *Angew. Chem., Int. Ed.*, 2002, **41**, 2171–2174.
26 S. D. Weeks, M. Drinker and P. J. Loll, *Protein Expression Purif.*, 2007, **53**, 40–50.
27 I. E. Gentle, S. D. P. De and M. Baca, *Bioconjugate Chem.*, 2004, **15**, 658–663.
28 N. Singla, H. Erdjument-Bromage, J. J. P. Himanen, T. W. Muir and D. B. Nikolov, *Chem. Biol.*, 2011, **18**, 361–371.
29 (*a*) D. Macmillan, R. M. Bill, K. A. Sage, D. Fern and S. L. Flitsch, *Chem. Biol.*, 2001, **8**, 133–145; (*b*) K. Kuhn, D. J. Owen, B. Bader, A. Wittinghofer, J. Kuhlmann and H. Waldmann, *J. Am. Chem. Soc.*, 2001, **123**, 1023–1035; (*c*) I. Wei ping Zheng, Z. Zhongsen, S. Ganguly, J. L. Weller, D. C. Klein and P. A. Cole, *Nat. Struct. Biol.*, 2003, **10**, 1054–1057; (*d*) S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105–1109; (*e*) N. Floyd, B. Vijayakrishnan, J. R. Koeppe and B. G. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 7798–7802; (*f*) Y. A. Lin, J. M. Chalker and B. G. Davis, *J. Am. Chem. Soc.*, 2010, **132**, 16805–16811; (*g*) J.-M. Kee, B. Villani, L. R. Carpenter and T. W. Muir, *J. Am. Chem. Soc.*, 2010, **132**, 14327–14329; (*h*) N. D. Weikart and H. D. Mootz, *ChemBioChem*, 2010, **11**, 774–777; (*i*) L. Yi, M. Abootorabi and Y.-W. Wu, *ChemBioChem*, 2011, **12**, 2413–2417; (*j*) F. Li, A. Allahverdi, R. Yang, G. B. J. Lua, X. Zhang, Y. Cao, N. Korolev, L. Nordenskiöld and C.-F. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 9611–9614; (*k*) G. J. L. Bernardes,

This journal is © The Royal Society of Chemistry 2012

L. Linderoth, K. J. Doores, O. Boutureira and B. G. Davis, *ChemBioChem*, 2011, **12**, 1383–1386; (*l*) N. D. Weikart, S. Sommer and H. D. Mootz, *Chem. Commun.*, 2012, **48**, 296–298.
30 (*a*) A. Varki, *Glycobiology*, 1993, **3**, 97–130; (*b*) R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683–720; (*c*) R. G. Spiro, *Glycobiology*, 2002, **12**, 43R–56R.
31 D. H. Dube and C. R. Bertozzi, *Nat. Rev. Drug Discovery*, 2005, **4**, 477–488.
32 (*a*) B. G. Davis, *Chem. Rev.*, 2002, **102**, 579–602; (*b*) A. Brik, S. Ficht and C.-H. Wong, *Curr. Opin. Chem. Biol.*, 2006, **10**, 638–644.
33 (*a*) J. R. Rich and S. G. Withers, *Nat. Chem. Biol.*, 2009, **5**, 206–215; (*b*) F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi and L.-X. Wang, *Nat. Chem. Biol.*, 2010, **6**, 264–266.
34 (*a*) H. Hojo and Y. Nakahara, *Biopolymers (Peptide Science)*, 2007, **88**, 308–324; (*b*) D. P. Gamblin, E. M. Scanlan and B. G. Davis, *Chem. Rev.*, 2008, **109**, 131–163; (*c*) R. J. Payne and C.-H. Wong, *Chem. Commun.*, 2010, **46**, 21–43.
35 N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson and Y. Kajihara, *J. Am. Chem. Soc.*, 2007, **130**, 501–510.
36 M. Baggiolini, B. Dewald and B. Moser, *Adv. Immunol.*, 1994, **55**, 97–179.
37 D. D. Taub, P. Proost, W. J. Murphy, M. Anver, D. L. Longo, J. van Damme and J. J. Oppenheim, *J. Clin. Invest.*, 1995, **95**, 1370–1376.
38 (*a*) J. B. Blanco-Canosa and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2008, **47**, 6851–6855; (*b*) F. Mende and O. Seitz, *Angew. Chem., Int. Ed.*, 2011, **50**, 1232–1240.
39 D. Macmillan and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 1355–1359.
40 Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 1999, **121**, 11684–11689.
41 M. M. Chen, A. I. Bartlett, P. S. Nerenberg, C. T. Friel, C. P. R. Hackenberger, C. M. Stultz, S. E. Radford and B. Imperiali, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 22528–22533.
42 (*a*) C. Piontek, P. Ring, O. Harjes, C. Heinlein, S. Mezzato, N. Lombana, C. Pöhner, M. Püttner, D. Varón Silva, A. Martin, F. X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1936–1940; (*b*) C. Piontek, D. Varón Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1941–1945.
43 P. Nagorny, N. Sane, B. Fasching, B. Aussedat and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2011, **51**, 975–979.
44 M. Murakami, R. Okamoto, M. Izumi and Y. Kajihara, *Angew. Chem., Int. Ed.*, 2012, **51**, 3494, DOI: 10.1002/anie.201200881.
45 I. Sakamoto, K. Tezuka, K. Fukae, K. Ishii, K. Taduru, M. Maeda, M. Ouchi, K. Yoshida, Y. Nambu, J. Igarashi, N. Hayashi, T. Tsuji and Y. Kajihara, *J. Am. Chem. Soc.*, 2012, **134**, 5428, DOI: 10.1021/ja2109079.
46 B. Aussedat, B. Fasching, E. Johnston, N. Sane, P. Nagorny and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3532–3541.
47 J.-M. Kee and T. W. Muir, *ACS Chem. Biol.*, 2011, **7**, 44–51.
48 (*a*) O. Hantschel and G. Superti-Furga, *Nat. Rev. Mol. Cell Biol.*, 2004, **5**, 33–44; (*b*) J. Colicelli, *Sci. Signal.*, 2010, **3**, re6.
49 J. Massague and D. Wotton, *EMBO J.*, 2000, **19**, 1745–1754.
50 J. J. Ottesen, M. Huse, M. D. Sekedat and T. W. Muir, *Biochemistry*, 2004, **43**, 5698–5706.
51 L. M. Szewczuk, M. K. Tarrant, V. Sample, W. J. Drury, J. Zhang and P. A. Cole, *Biochemistry*, 2008, **47**, 10407–10419.
52 K. P. Chiang, M. S. Jensen, R. K. McGinty and T. W. Muir, *ChemBioChem*, 2009, **10**, 2182–2187.
53 M. A. Shogren-Knaak, C. J. Fry and C. L. Peterson, *J. Biol. Chem.*, 2003, **278**, 15744–15748.
54 M. Broncel, E. Krause, D. Schwarzer and C. P. R. Hackenberger, *Chem.–Eur. J.*, 2012, **18**, 2488–2492.
55 M. Hejjaoui, S. Butterfield, B. Fauvet, F. Vercruysse, J. Cui, I. Dikiy, M. Prudent, D. Olschewski, Y. Zhang, D. Eliezer and H. A. Lashuel, *J. Am. Chem. Soc.*, 2012, **134**, 5196, DOI: 10.1021/ja210866j.
56 A. Hershko and A. Ciechanover, *Annu. Rev. Biochem.*, 1998, **67**, 425–479.
57 (*a*) A. Hershko and H. Heller, *Biochem. Biophys. Res. Commun.*, 1985, **128**, 1079–1086; (*b*) M. H. Glickman and A. Ciechanover, *Physiol. Rev.*, 2002, **82**, 373–428.
58 H. Huang, M.-s. Jeon, L. Liao, C. Yang, C. Elly, J. R. Yates Iii and Y.-C. Liu, *Immunity*, 2010, **33**, 60–70.
59 (*a*) C. Chatterjee, R. K. McGinty, J. P. Pellois and T. W. Muir, *Angew. Chem., Int. Ed.*, 2007, **46**, 2814–2818; (*b*) R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder and T. W. Muir, *Nature*, 2008, **453**, 812–816; (*c*) K. S. A. Kumar, L. Spasser, L. A. Erlich, S. N. Bavikar and A. Brik, *Angew. Chem., Int. Ed.*, 2010, **49**, 9126–9131; (*d*) R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu and C.-F. Liu, *Chem. Commun.*, 2010, **46**, 7199–7201; (*e*) F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, H. Hilkmann, T. K. Sixma and H. Ovaa, *Angew. Chem., Int. Ed.*, 2010, **49**, 10149–10153; (*f*) S. Virdee, Y. Ye, D. P. Nguyen, D. Komander and J. W. Chin, *Nat. Chem. Biol.*, 2010, **6**, 750–757; (*g*) S. Virdee, P. B. Kapadnis, T. Elliott, K. Lang, J. Madrzak, D. P. Nguyen, L. Riechmann and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 10708–10711; (*h*) K. S. A. Kumar, L. Spasser, S. Ohayon, L. A. Erlich and A. Brik, *Bioconjugate Chem.*, 2011, **22**, 137–143; (*i*) C. Castañeda, J. Liu, A. Chaturvedi, U. Nowicka, T. A. Cropp and D. Fushman, *J. Am. Chem. Soc.*, 2011, **133**, 17855–17868; (*j*) C. A. Castaneda, J. Liu, T. R. Kashyap, R. K. Singh, D. Fushman and T. A. Cropp, *Chem. Commun.*, 2011, **47**, 2026–2028.
60 T. Jenuwein and C. D. Allis, *Science*, 2001, **293**, 1074–1080.
61 M. Hejjaou, M. Haj-Yahya, K. S. A. Kumar, A. Brik and H. A. Lashuel, *Angew. Chem., Int. Ed.*, 2010, **50**, 405–409.
62 K. S. A. Kumar and A. Brik, *Isr. J. Chem.*, 2011, **51**, 900–907.
63 M. Haj-Yahya, K. S. Ajish Kumar, L. A. Erlich and A. Brik, *Biopolymers (Peptide Science)*, 2010, **94**, 504–510.
64 (*a*) X. Li, T. Fekner, J. J. Ottesen and M. K. Chan, *Angew. Chem., Int. Ed.*, 2009, **121**, 9348–9351; (*b*) A. Shanmugham, A. Fish, M. P. A. Luna-Vargas, A. C. Faesen, F. El Oualid, T. K. Sixma and H. Ovaa, *J. Am. Chem. Soc.*, 2010, **132**, 8834–8835; (*c*) C. Chatterjee, R. K. McGinty, B. Fierz and T. W. Muir, *Nat. Chem. Biol.*, 2010, **6**, 267–269; (*d*) J. Chen, Y. Ai, J. Wang, L. Haracska and Z. Zhuang, *Nat. Chem. Biol.*, 2010, **6**, 270–272.
65 S. Eger, B. Castrec, U. Hübscher, M. Scheffner, M. Rubini and A. Marx, *ChemBioChem*, 2011, **12**, 2807–2812.
66 S. Sommer, N. D. Weikart, A. Brockmeyer, P. Janning and H. D. Mootz, *Angew. Chem., Int. Ed.*, 2011, **50**, 9888–9892.
67 M. Manohar, A. M. Mooney, J. A. North, R. J. Nakkula, J. W. Picking, A. Edon, R. Fishel, M. G. Poirier and J. J. Ottesen, *J. Biol. Chem.*, 2009, **284**, 23312–23321.
68 X. Wang and J. J. Hayes, *Mol. Cell. Biol.*, 2008, **28**, 227–236.
69 G. Triola, H. Waldmann and C. Hedberg, *ACS Chem. Biol.*, 2011, **7**, 87–99.
70 S. He, D. Bauman, J. S. Davis, A. Loyola, K. Nishioka, J. L. Gronlund, D. Reinberg, F. Meng, N. Kelleher and D. G. McCafferty, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 12033–12038.
71 D. Kadereit, J. Kuhlmann and H. Waldmann, *ChemBioChem*, 2000, **1**, 144–169.
72 L. Brunsveld, J. Kuhlmann, K. Alexandrov, A. Wittinghofer, R. S. Goody and H. Waldmann, *Angew. Chem., Int. Ed.*, 2006, **45**, 6622–6646.
73 T. Durek, K. Alexandrov, R. S. Goody, A. Hildebrand, I. Heinemann and H. Waldmann, *J. Am. Chem. Soc.*, 2004, **126**, 16368–16378.

Downloaded by University of California - San Diego on 01 September 2012
Published on 27 March 2012 on http://pubs.rsc.org | doi:10.1039/C2OB25149C