Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Synthetic and semi-synthetic strategies to study ubiquitin signaling

Gabriëlle BA van Tilburg[1,2], Angela F Elhebieshy[1,2] and Huib Ovaa[1,2]

The post-translational modification ubiquitin can be attached to the ε-amino group of lysine residues or to a protein's N-terminus as a mono ubiquitin moiety. Via its seven intrinsic lysine residues and its N-terminus, it can also form ubiquitin chains on substrates in many possible ways. To study ubiquitin signals, many synthetic and semi-synthetic routes have been developed for generation of ubiquitin-derived tools and conjugates. The strength of these methods lies in their ability to introduce chemo-selective ligation handles at sites that currently cannot be enzymatically modified. Here, we review the different synthetic and semi-synthetic methods available for ubiquitin conjugate synthesis and their contribution to how they have helped investigating conformational diversity of diubiquitin signals. Next, we discuss how these methods help understanding the ubiquitin conjugation–deconjugation system by recent advances in ubiquitin ligase probes and diubiquitin-based DUB probes. Lastly, we discuss how these methods help studying post-translational modification of ubiquitin itself.

Addresses
[1] Department of Cell Biology II, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands
[2] Department of Chemical Immunology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands

Corresponding author: Ovaa, Huib (h.ovaa@nki.nl)

Current Opinion in Structural Biology 2016, **38**:92–101

This review comes from a themed issue on **New constructs and expression of proteins**

Edited by **Rob Meijers** and **Anastassis Perrakis**

For a complete overview see the Issue and the Editorial

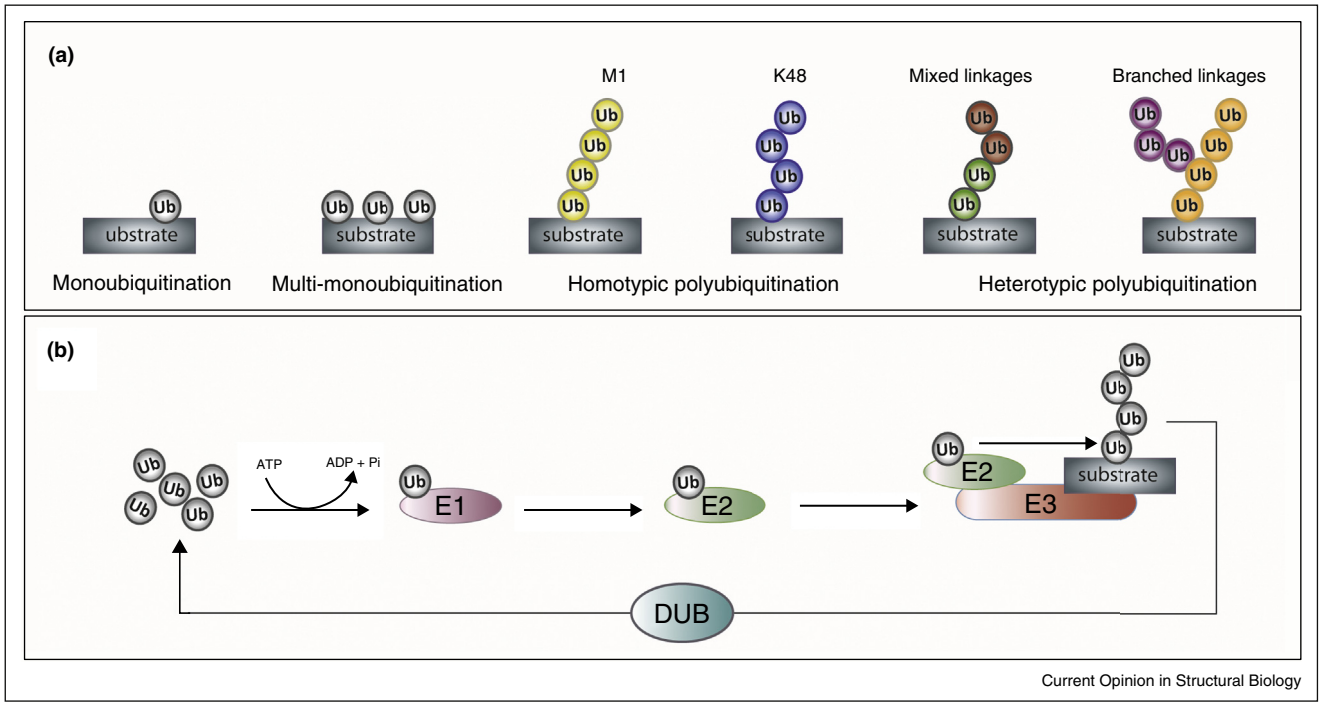
Available online 15th June 2016

http://dx.doi.org/10.1016/j.sbi.2016.05.022

Introduction

Ubiquitin (Ub) is a 76-amino acid post-translational modifier fundamental to cellular homeostasis. Cellular processes regulated by ubiquitin modification range from classically known Ub-mediated proteasomal degradation to DNA repair, cell division, endoplasmic reticulum-associated degradation (ERAD), mRNA stability and even regulation of the innate immune system [1]. Not surprisingly, deregulation of the ubiquitin system is associated with various serious illnesses such as cancer, immunological disorders and neurodegenerative diseases [2]. The post-translational modifier can be covalently attached to substrate proteins at the ε-amino group of lysine residues or at the protein's N-terminal residue [3,4]. Since ubiquitin harbors seven intrinsic lysine residues, it can also be conjugated to another Ub moiety. In this manner, homotypical ubiquitin chains of a single linkage type consisting of M1, K6, K11, K27, K29, K33, K48 or K63 can be formed, all of which are known to exist *in vivo* [5]. In addition, heterotypical chains of multiple ubiquitin linkage types can be formed, opening up an even more complex layer of post-translational modification (Figure 1a). Conjugation of Ub to a substrate protein is carried out by a cascade of three enzymatic activities: E1 ubiquitin-activating, E2 ubiquitin-conjugating and E3 ubiquitin-ligating activities (Figure 1b). To date, 2 human E1s, about 40 E2s and over 600 E3 enzymes are known. The combination of E2 and E3 enzymes dictates what type of ubiquitin chain is formed and which substrate protein becomes ubiquitinated. In addition, the ubiquitination status of a protein can be regulated by removal or editing of ubiquitin chains, which is carried out by a family of approximately 100 deubiquitinating enzymes (DUBs) [6]. For some of these DUBs, linkage specificity has also been observed. To study the properties of differentially linked ubiquitin chains, traditionally one has relied on generating chains enzymatically by usage of linkage specific E2–E3 pairs. Not all E2-E3 pairs are known however, and some have only been uncovered recently. This resulted in a less extensive knowledge of the 'a-typical' K6-linked, K11-linked, K27-linked, K29-linked and K33-linked ubiquitin chains. Therefore, much effort has been put into making differentially linked ubiquitin derivatives through synthetic and semi-synthetic methods, since these methods allow site specific incorporation of a specific chemoselective ligation handle. In addition, ubiquitin-based DUB probes and fluorescent ubiquitin-based enzyme substrates have seen an enormous boost, producing ubiquitin-based tools in all sorts of different flavors [7]. Two other highly investigated areas of synthetic and semi-synthetic ubiquitin–conjugate research focus on ubiquitinated histones [8,9] and ubiquitinated α-synuclein [10,11]. We will briefly introduce the current synthetic and semi-synthetic methods that can be applied to generate ubiquitin and Ub-based tools and then focus on three other rapidly developing areas. First, we discuss how semi-synthetic diubiquitin synthesis has aided the structural characterization of all differentially linked diubiquitin modules, which further enhanced our understanding of

Figure 1



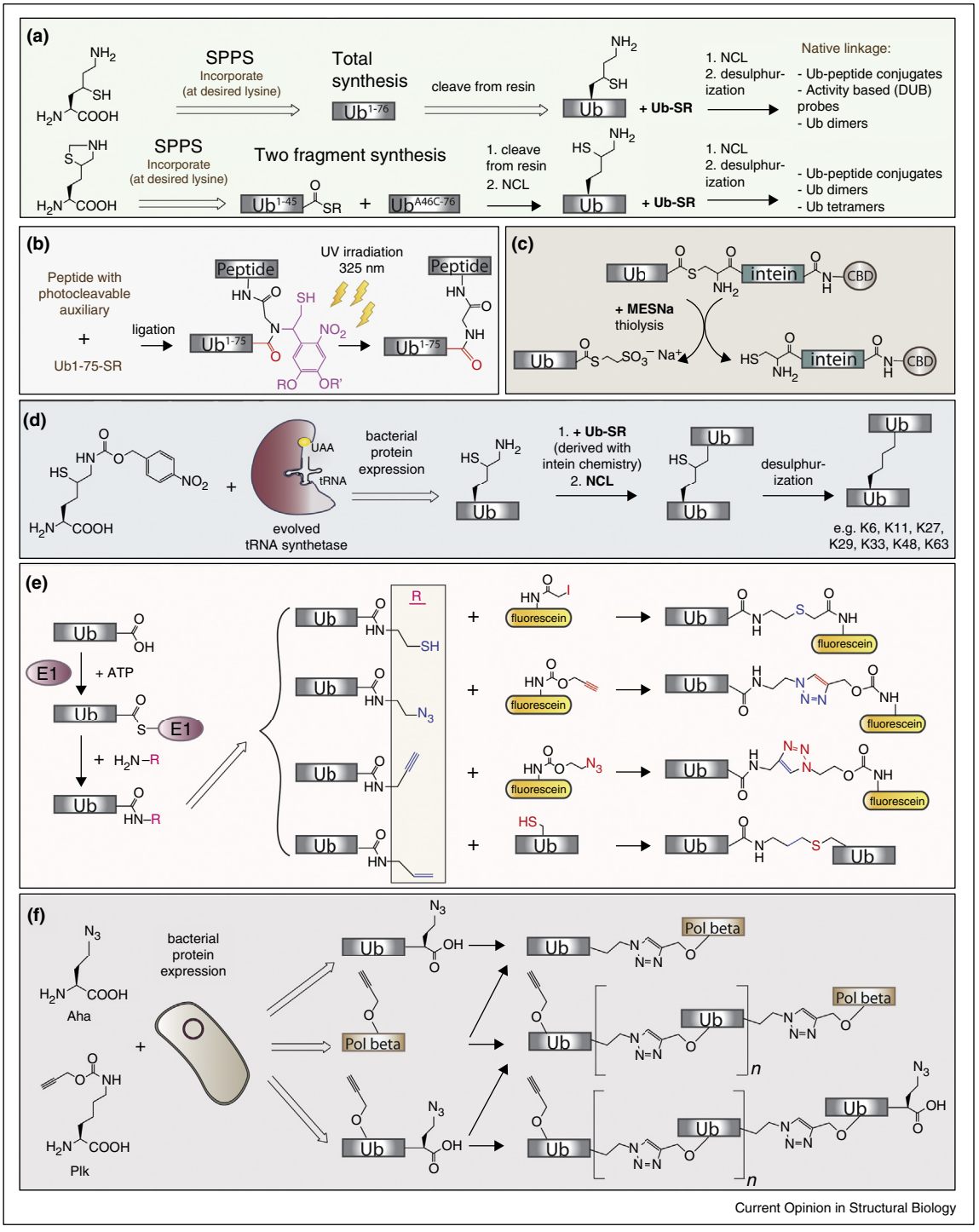
The ubiquitin system. **(a)** Types of ubiquitin conjugation: ubiquitin (UB) can be conjugated as a monomer on one site, or on multiple sites of the substrate protein (multi-monoubiquitination). It can also form homotypical ubiquitin polymers through its N-terminus (M1-linked) or either one of its seven lysine residues (e.g. K48-linked). Mixing of different linkage types gives rise to heterotypic polyubiquitin chains. **(b)** General overview of ubiquitin conjugation and deubiquitination by E1, E2, E3 enzymes and deubiquitinating enzymes (DUBs). Ubiquitin is activated by an E1 enzyme, transferred to a specific E2 enzyme and conjugated to a substrate protein with the help of an E3 ligase.

how different linkages can confer specific signals. Next, we focus on recent advances in diubiquitin-based DUB probes and the development of ubiquitin ligase probes that open up new interesting areas of ubiquitin structural biology. Finally, we discuss advances in making synthetic and semi-synthetic ubiquitin molecules containing other post-translational modifications such as phosphorylation and acetylation.

Chemical strategies

Native chemical ligation (NCL) has been an extremely useful tool to make ubiquitin, ubiquitin–peptide conjugates, ubiquitin dimers and ubiquitin tetramers, as reviewed recently by Pham *et al.* [11]. Chemical synthesis of ubiquitinated peptides was first established by Muir and co-workers and utilizes a ligation auxiliary where the auxiliary group is removed under photolytic conditions; yielding a natively linked Ub-peptide conjugate [12]. Today, most reported methods rely on the incorporation of a γ-thiolysine or δ-thiolysine moiety at a designated lysine residue to allow NCL with a thioester moiety. The thiol containing ubiquitin module can be synthesized with total, linear synthesis [13], or from two fragments [14]. During the total, linear fluorenylmethyloxycarbonyl (Fmoc)-based solid phase peptide synthesis (SPPS) approach, the growing peptide chain is stabilized by the incorporation of special building blocks that prevent the formation of aggregates as the ubiquitin chain grows. In the two segment approach, an N-terminal Ub(1–45)-SR fragment is synthesized and ligated to a synthetic C-terminal Ub[(46–76)-A46C] fragment. In the latter fragment, alanine 46 is replaced by N-methylcysteine to allow NCL with the first fragment and is afterwards converted into the native alanine residue through a desulphurization step. To make ubiquitin dimers and other conjugates, a thioester needs to be introduced at the ubiquitin C-terminus for NCL with the Ub thiolysine-containing module. The thioester functionality can be incorporated by E1-mediated enzymatic conversion with sodium 2-mercaptoethane sulfonate (MESNa) [13] or during Fmoc-based SPPS [15]. Next, NCL can be performed and a subsequent desulphurization step results in a ubiquitin conjugate that bears the native isopeptide linkage. The above described methods are schematically represented in Figure 2a,b. In addition, other strategies yielding a non-native isopeptide linkage have been reported and include oxime-based ligation for non-hydrolysable ubiquitin–conjugate synthesis [16] and thioether based ligation to prepare diubiquitin [17], branched triubiquitin [18] and polyubiquitin [19] modules that retain

Figure 2



Overview of described synthetic methods that yield a native isopeptide bond (**a** and **b**) and described semi-synthetic ligation methods (**c**–**f**). (a) Native chemical ligation using γ-thiolysine (top) or δ-mercaptolysine (bottom) auxiliaries. (b) Photo-crosslinking ligation using a photolytically (UV 325 nm) removable auxiliary. (c) MESNa mediated thiolysis of a ubiquitin–intein construct generates a ubiquitin thioester which can be used for ligation purposes. (d) Semi-synthetic incorporation of δ-thiol-L-lysine with genetic code expansion by addition of δ-thiol-Nε-(p-nitrocarbobenzyloxy)lysine to bacterial cell cultures. (e) Semi-synthetic functionalization of the ubiquitin C-terminus with different moieties utilizing the E1-enzyme. Thioether-based linkages and triazole-based linkages can also be made using synthetic methods. (f) Semi-synthetic incorporation of Aha and Plk allows click chemistry to make several ubiquitin–peptide conjugates and ubiquitin polymers. NCL, native chemical ligation; MESNa, 2-mercaptoethane sulfonate; CBD, chitin binding domain; Aha, azidohomoalanine, Plk, propargyl-derivatized lysine.

a sulfur atom in the forged isopeptide bond. It is clear that synthetic protein chemistry has much to offer to the biological community, but one must keep in mind that the chemistry is not always straightforward. In addition, protein folding must always be checked, although this is less of an issue with the very stable ubiquitin protein.

Semi-synthetic strategies

Next to synthetic strategies, protein semi-synthesis has also become a very convenient strategy for the production of large peptides and small proteins since it was first introduced. Using expressed protein as the starting material, a number of strategies have been invented to equip the protein with chemical reactive groups, labels or fluorogenic dyes. For instance, the use of inteins has become a useful tool for the formation of protein thioesters through MESNa-mediated thiolysis (Figure 2c) [20]. Expansion of the genetic code with unnatural amino acids (UAAs) has further aided the field of protein semi-synthesis [21]. Briefly, an UAA can be incorporated by a modified transfer RNA (tRNA) at the site of a rarely used codon, such as the amber stop codon in *E. coli*. To this end, the active site of the complementary aminoacyl tRNA synthetase (aaRS) needs to be mutated to allow selective recognition of the UAA, which will then be transferred to the modified tRNA and eventually incorporated into the newly translated protein. While genetic code expansion-based methods are clearly useful, most do require certain expertise that can only be found in specialized labs and often require specific *E. coli* strains and tRNA pairs. Below, we will discuss genetic code expansion strategies employed on ubiquitin, while a more extensive overview of incorporation of chemoselective unnatural amino acids in general is discussed by Lang *et al.* [22].

The GOPAL approach (genetically encoded orthogonal protection and activated ligation) uses genetic code expansion for site specific incorporation of tert-butyloxycarbonyl (Boc)-protected lysine in ubiquitin with a specific *Mb*PylRS/*Mb*PylRS$_{CUA}$ pair. Initially, two bioorthogonal protection groups were used that allowed selective deprotection of the desired lysine residue on the proximal ubiquitin module [23]. Later, a specific tRNA pair was found that allows site-specific incorporation of δ-thio-L-lysine and δ-hydroxy-L-lysine without the Boc protecting group, circumventing the need of the previously reported extensive protection and deprotection strategy (Figure 2d) [24].

Another semi-synthetic strategy for making diubiquitin mimics was reported in 2014 and utilizes E1 enzyme to equip the ubiquitin C-terminus with several reactive groups through an amidation reaction (Figure 2e). The advantage of this method is that it does not require extensive knowledge of peptide chemistry nor of genetic code expansion. In this case, allylamine was used to equip the Ub C-terminus with an alkene. Expressed ubiquitin lysine to cysteine mutants at the desired lysine conjugation site were converted into ubiquitin dimers by UV irradiation [25•]. Through this procedure, a C-terminal alkyne moiety was also introduced, which in turn can be used as a reactive group to form non-hydrolysable ubiquitin conjugates with Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC, or 'click chemistry').

The group of Marx has also developed a strategy to achieve non-hydrolysable ubiquitin–conjugate synthesis through bacterial expression and click chemistry (Figure 2f) [26••]. For click chemistry, both an azide and an alkyne are needed. The methionine analog azidohomoalanine (Aha) was incorporated on the ubiquitin C-terminus as the azide containing unnatural amino acid. The alkyne functionality was introduced at one of the desired lysine conjugation positions by using propargyl-derivatized pyrrolysine (Plk) and the PylRS/tRNA^{Pyl} pair of *Methanosarcina barkeri* [27]. After introduction of both azide and alkyne functionalities, non-hydrolysable ubiquitin conjugates can be formed by CuAAC. Such non-hydrolysable ubiquitin-based tools can be utilized to address the selective nature of ubiquitin binding domains towards a certain diubiquitin molecule [28], to study the impact of a monoubiquitination event such as on Polβ [27] and PCNA [29], or to study the effect of non-hydrolysable polyubiquitin chains such as K11 on cell cycle related events [30].

Using synthetic and semi-synthetic tools to study ubiquitin chain structure and biochemistry

Over the recent years, much effort has gone into structural and biochemical characterization of di-ubiquitin and polyubiquitin molecules representing all eight different homogenously linked ubiquitin types. In these efforts, people have been hampered by the lack of specific E2 and E3 enzymes to generate the so-called atypical (K6, K11, K27, K29, K33) chains. The K6 diubiquitin structure was first solved by using the semi-synthetic GOPAL approach [23]. Only recently an enzymatic approach for making K6-polyubiquitin chains was reported [31]. K11 diubiquitin structure elucidation was achieved by two groups independently in 2010 [32,33] after a K11-specific ligase was identified in 2009 [34]. Recently, K29-linked and K33-linked diubiquitins have been structurally characterized after a combinatorial ligation–deubiquitination approach was developed to generate such linkages enzymatically [35,36]. Currently, only K27-linked ubiquitin cannot be prepared fully enzymatically, although interestingly a recent study reports K27 polyubiquitination by the E3 ligase RNF168 [37]. Nevertheless, semi-synthetic and synthetic strategies have been an indispensable tool for studying the structural and biophysical properties of atypical diubiquitins.

Two recent papers describe NMR structural conformation of all atypical diubiquitins in solution. With the exception of K11-linked diubiquitin, all were prepared by a variant of the semi-synthetic GOPAL method. Based on SANS (small-angle neutron scattering), 15N relaxation and RDC (residual dipolar coupling) NMR data, Castañeda *et al.* found that the available diubiquitin crystal structure data poorly overlaps with their in-solution NMR data, which suggests a great flexibility in conformation of Ub chains [38••]. Multiple crystal conformations for the same diubiquitin molecules have indeed been reported, which is consistent with this notion. The same approach was next used to further characterize the atypical K27 diubiquitin linkage in particular [39•]. Interestingly, K27 diubiquitin was processed very inefficiently by the DUBs tested, including more promiscuous ones. It must be noted however that in other studies, DUBs have been described that do cleave K27-linked diubiquitin quite similarly to other diubiquitin molecules, albeit not specifically [40,41]. The study by Castañeda *et al.* provides the first structural analysis of K27 diubiquitin: it shows that the K27 residue is almost completely buried in the diubiquitin structure and that it has very low solvent accessibility. In addition, the C-terminal Gly75–76 region of the distal Ub moiety is more ordered compared to the other known diubiquitin structures. This structural aspect might explain why K27 diubiquitin is processed less efficiently by some DUBs.

Using ubiquitin-based activity probes for structural and biochemical characterization of the ubiquitin conjugation–deconjugation machinery

Ubiquitin-based probes have been used extensively in the past to investigate the mechanism of DUB activation and to trap deubiquitinating enzymes in their active state. Numerous structures were solved with ubiquitin aldehyde (Ubal), the vinyl methyl ester (VME) warhead or the propargyl (PA) warhead (Table 1). While activity-based probes have been recently reviewed [7], we would like to highlight a few recently emerged probes which open up new interesting areas of ubiquitin biology.

The introduction of diubiquitin probes with a warhead in-between the distal and proximal ubiquitin module has allowed a more selective investigation of diubiquitin recognition by DUBs [42–45]. These covalent probes certainly allow more detailed structural investigation of diubiquitin-specific DUB recognition, but do not allow investigation of potential additional binding sites. A recent report was made on DUB probes that do allow investigation of supplementary binding sites, such as the so-called DUB S2 site, by utilizing click chemistry and incorporation of a C-terminal propargyl warhead [46••]. Generated fluorogenic substrates allowed for detailed studies of enzymatic turnover and in a more recent report [47], one of these probes was used to structurally characterize the K48 polyubiquitin cleaving mechanism of the SARS DUB PLpro.

While there is a plethora of DUB probes available, until recently the E1–E2–E3 machinery has been falling behind in this respect. An AMP-based E1 probe has been described recently and could be of use for studying E1 enzymes for other ubiquitin-like proteins as well [63•]. In addition, a probe that reacts with purified E2 and E3 ligases has been reported and was used to study activation of the Parkin E3 ligase in detail [64••]. Lastly, a recent report was made on a Ub-dehydroalanine-based ligase probe, which can be passed on in the ligation cascade to E1, E2 and E3 ligases of the HECT/RBR type [65••].

Using synthetic and semi-synthetic tools to study post-translational modifications on ubiquitin

An exciting emerging area currently under investigation is the post-translational modification of ubiquitin itself. In 2010, a report was made about deamidation of Gln40 on ubiquitin by the bacterial effector protein CHBP, which effectively impaired ubiquitin chain synthesis [66]. Recently, it was found that ubiquitin can be acetylated on Lys6, Lys48 and Lys63 by mass spectrometric efforts [67]. The effects of acetylation were studied by incorporation of N-acetyllysine on positions 6 and 48 of ubiquitin through genetic code expansion. Interestingly, acetylation at Lys6 and Lys48 on mono ubiquitin prevented ubiquitin chain elongation by several E2 enzymes *in vitro*. Histone H2B was identified as a substrate for acetylated mono Ub, further suggesting a chain terminating effect of the acetylation event. In addition to the three acetylation sites, Ohtake et al. also detected phosphorylation on Thr14 and Ser65 in conjugated ubiquitin species. Ub has eight potential phosphorylation sites: Thr7, Thr12, Thr14, Ser20, Ser57, Tyr59, Ser65 and Thr66. A general overview of post translational modifications on Ub is presented in Figure 3a.

The best studied example of ubiquitin phosphorylation is Ser65 phosphorylated Ub. Phosphorylation on Ser65 is mediated by PTEN-induced kinase 1 (PINK1) and activates the RING-in-between-RING (RBR) E3 ligase Parkin. This in turn leads to clearance of damaged mitochondria in a process called 'mitophagy' [68]. Mutations in both Parkin and PINK1 have been implicated in the early onset of autosomal-recessive Parkinson's disease and players in this pathway are therefore heavily investigated for therapeutic purposes [69]. Although Ser65 phosphorylated Ub is easily made enzymatically using recombinant PINK1, this is restricted to this residue only. Not surprisingly, several efforts have been put into making phosphorylated ubiquitin available through synthetic and semi-synthetic approaches that allow investigation of other potential phosphorylation sites. Phosphorylated

Table 1

Overview of DUB structures with Ub-based activity-based probes and key findings summarized. Ubal, ubiquitin aldehyde; Ub-VME, ubiquitin vinyl methyl ester; Ub-PA, ubiquitin propargyl

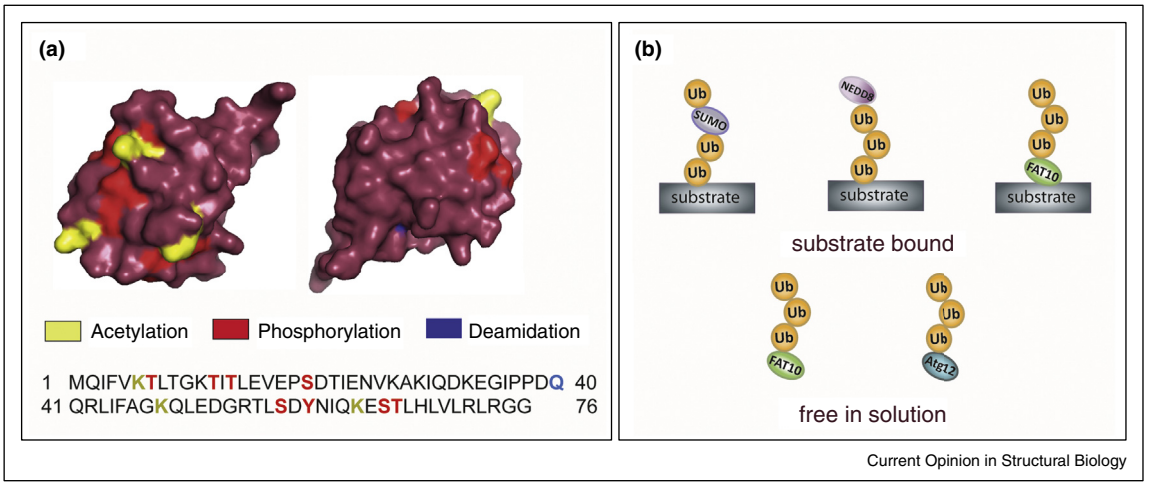
Probe	Ref.	DUB	PDB code	Key findings
Ubal	[48]	Otubain-1	4LDT	Binding of free ubiquitin to Otubain-1 (OTUB1) triggers conformational changes in the OTU domain of OTUB1 and allosterically increases the affinity for Ubc13∼Ub.
	[49]	SAGA DUB module	3MHS	Integrity of the SAGA module (Ubp8, Sgf11, Sus1 and Sgr73) is essential for proper Ubp8 catalytic activity. Although the Ubp8 active site is well ordered, binding of Ubal shows important structural rearrangements in the vicinity of the active site. The intact SAGA module stabilizes Ubp8 so it is catalytically competent and able to bind Ub.
	[50]	SARS-CoV PLpro	4MM3	Structure elucidation allowed characterization of the binding surface while subsequent mutational analysis and structural modeling suggest the existence of a second binding site to provide K48 and ISG15 specificity of PLpro.
	[51]	USP14 (catalytic domain)	1XD3	In the apo USP14 structure, the active site is well formed *before* substrate binding. Ubiquitin binding is blocked by two nearby surface loops (BL1 and BL2). Ub binding translocates the two loops to allow access of its C-terminus to the USP14 active site.
	[52]	USP7/HAUSP (catalytic domain)	1NBF	In the apo USP7 structure, the catalytic histidine is nearly 10 Å away from the catalytic cysteine. The C-terminal segment of Ubal induces backbone conformational changes in the vicinity of the active site, leading to alignment of the catalytic triad.
	[53]	Yuh1	1CMX	Ubal binding is coupled to ordering of a 21-residue loop which blocks the active-site cleft in the absence of substrate.
Ub-VME	[54]	Legionella DUB module	5CRA	The $SDeA_{DUB}$ modules engage Gln40 of Ub instead of the common Ile44 hydrophobic patch. The sparse use of molecular contacts likely allows dual specificity of the DUB module towards Ub and NEDD8.
	[55]	$M48^{USP}$	2J7Q	An acidic cleft on $M48^{USP}$ accommodates the complementary basic C-terminal Ub stretch. Specificity is additionally mediated by a β-hairpin to recognize Ub over other UbL's such as ISG15.
	[56]	UCHL1	3KW5, 3IFW (S18Y), 3KVF (I93M)	Binding of the N-terminal ubiquitin β-hairpin results in placement of the UCHL1 active site histidine in the correct location, which in the apo structure is located 7.7 Å from the catalytic cysteine.
	[57]	hUCHL3	1XD3	UCH-L3 features a crossover loop that is disordered in the substrate-free structure and ordered above the active-site cleft when Ub is bound.
	[58]	*Pf*UCHL3	2WDT	The *Plasmodium falciparum (Pf)* UCHL3 structure provides a basis for dual recognition of both NEDD8 and Ub, but also shows distinct differences in the Ub binding site compared to human UCHL3.
	[59]	USP37	4I6N, 4IG7	The UCH37-like domain (ULD) contacts Ub and stabilizes an unusual salt bridge between Lys48 and Glu51 of Ub. This results in an inhibited enzyme state in the proteasome free form.
Ub-PA	[60]	UCH-L5	4UEL (RPN13)	UCH-L5 can be activated by RPN13 and deactivated by INO80G. Ub binding to RPN13 is identical to the canonical ubiquitin-binding mode found in all UCH family members. Ub binds via its C-terminal tail close to the UCH-L5 active site and via its core relatively far from the active site in a series of three specific exosites that lead to several structural rearrangements.
	[61]	vOTU	3ZNH	Structure elucidation of viral (v)OTU with Ub-propargyl closely resembles previous vOTU-Ub complexes and shows that the Ub-propargyl group forms a vinyl thioether linkage in complex.
Ub-Br_2	[62]	Otubain-2	4FJV	Ub binding is different from vOTU due to two extra β-strands in vOTU. Upon Ub binding, the OTUB2 catalytic residues show only subtle movements within the catalytic center. On either side of bound Ub however, large conformational changes occur in OTUB2. The Ub C-terminus is extended to reach the OTUB2 catalytic core and forms the main points of contact. N-terminal domain swapping reveals how OTUB1 can cleave different Ub-linkages than OTUB2.

ubiquitin variants other than Ser65 are indeed currently commercially available.

Chemical synthesis of phosphorylated ubiquitin

Last year, the total chemical synthesis of Ser65 phosphorylated mono and Lys63-linked diubiquitin was reported [70•]. Lys63-linked diubiquitin was prepared with a phospho Ser65 at the distal, proximal or both moieties of the diubiquitin module. Characterization of biological consequences of enzymatically prepared Ser65 phospho-Ub was reported that year and showed that the majority of tested DUBs were impaired in hydrolyzing phosphorylated

Figure 3



Crosstalk between ubiquitin and other post-translational modifications. **(a)** Reported post-translational modifications on ubiquitin shown on the Ub surface structure and the amino acid sequence of Ub. **(b)** Representation of *in vivo* found Ub/UbL heterologous chains.

ubiquitin chains [71]. An additional layer of DUB regulation was implied by using synthetic Lys63-linked distal, proximal and dual phospho-Ser65 diubiquitin, which showed that diminished DUB activity could be affected by position of the phosphorylation site at either the proximal or distal ubiquitin [70•]. Biological relevance of this observation is currently unknown.

Semi-synthesis of phosphorylated ubiquitin

A recently published semi-synthetic strategy to obtain phosphorylated ubiquitin involves native chemical ligation of an expressed N-terminal ubiquitin–intein fusion fragment and a synthetic C-terminal ubiquitin fragment [72•]. This allows introduction of a phosphorylated residue at 4 out the 8 known phosphorylation sites on ubiquitin in the C-terminal Ub(46–76) fragment.

Crosstalk with other post-translational modifications

Many reports have already been made implying crosstalk between ubiquitin and ubiquitin-like (UbL) proteins: a class of proteins that share high structural similarity and a common β-grasp fold with Ub. Best studied is the crosstalk between ubiquitin and SUMO [73], but ubiquitinated-NEDD8 chains have also been reported [74], as well as the existence of ubiquitinated FAT10 [75], ISGylated ubiquitin [76] and ubiquitinated Atg12 [77]. The protein (semi)-synthesis community is moving rapidly to catch up with this phenomenon and semi-synthetic strategies for obtaining ubiquitinated SUMO-2 [24] and ubiquitinated Rub1, the yeast NEDD8 homolog, have already been reported [78•]. Synthetic strategies for obtaining full length UbL proteins such as Nedd8 [61], SUMO-1 [79,80•] and Ufm1 [81] are also starting to emerge. It is not hard to imagine the existence of other PTMs on ubiquitin such as methylation or glycosylation, both of which could be incorporated through (semi-)synthetic approaches. Plenty of opportunities are available in this area and we expect that future research will provide many more interesting insights into this exciting new field.

Future directions for the field

Although much has been accomplished already, there are many more future challenges in the ubiquitin and ubiquitin-like field for (semi)-synthetic scientists. The synthesis of lysine-specific, homotypical polyubiquitin chains is such an example; as is the (semi-)synthesis of heterotypical ubiquitin chains and mixed Ub/UbL chains. With regard to probes, much more can be achieved by extending the scope towards ubiquitin-like proteins, which in general is a much less characterized class of proteins. Many opportunities also still lie in the (semi-)synthesis of these UbL proteins, which would allow incorporation of fluorogenic dyes, warheads, ligation sites or other useful chemical handles. Finally, the seemingly limitless emerging post-translational crosstalk between protein modifiers reveals a whole new layer of interesting protein regulation. Here, (semi)-synthetic biology would certainly add great value to help understanding the expanded ubiquitin code. Not only for ubiquitin itself, but also for ubiquitin-like proteins and possibly even their decoding sequences: (ubiquitin) binding domains.

Summary

The post-translational modifier ubiquitin is an essential constituent of the cell and participates in many cellular processes. Synthetic and semi-synthetic strategies to obtain ubiquitin and ubiquitin-based derivatives have provided indispensable tools to study ubiquitin biochemistry. Structural NMR-based characterization of semi-synthetic

atypical diubiquitins has revealed great flexibility of the ubiquitin signal and helped further understanding of how different Ub linkages can confer specific signals. Ubiquitin-based probes have helped understanding DUB activation mechanisms and the coming of new diubiquitin-based DUB probes will shed more insight into polyubiquitin recognition. Ub-based ligase probes are starting to emerge and will provide more mechanistic investigation of the ubiquitin ligation process. Additionally, techniques are starting to emerge to study an even more intricate layer of the ubiquitin signal: post-translational modification of ubiquitin itself. Deamidation, acetylation and phosphorylation of ubiquitin have been reported and in addition, crosstalk between ubiquitin-like proteins and Ub is known to exist *in vivo*. We expect that the use of synthetic and semi-synthetic strategies will be very useful to study this exciting new field of ubiquitin regulation as they allow control over the exact site of the modification.

Conflict of interest

H.O. is founder and shareholder of the company UbiQ that markets reagents in the ubiquitin field. H.O. is part of the DUB Alliance that includes Cancer Research Technology and FORMA Therapeutics

Acknowledgments

This work was supported by a grant from the European Research Council (ERC) to HO (Grant Agreement Number 281699) and by the Institute of Chemical Immunology (ICI), an NWO Gravitation project funded by the Ministry of Education, Culture and Science of the government of the Netherlands. We thank Gerbrand van der Heden-van Noort and Dennis Flierman for critical reading of the manuscript and for suggestions.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Akutsu M, Dikic I, Bremm A: **Ubiquitin chain diversity at a glance**. *J Cell Sci* 2016, **129**:875-880.

2. Chaugule VK, Walden H: **Specificity and disease in the ubiquitin system**. *Biochem Soc Trans* 2016, **44**:212-227.

3. Ciechanover A, Ben-Saadon R: **N-terminal ubiquitination: more protein substrates join in**. *Trends Cell Biol* 2004, **14**:103-106.

4. Hershko A, Ciechanover A: **The ubiquitin system**. *Annu Rev Biochem* 1998, **67**:425-479.

5. Peng J *et al.*: **A proteomics approach to understanding protein ubiquitination**. *Nat Biotechnol* 2003, **21**:921-926.

6. Komander D, Rape M: **The ubiquitin code**. *Annu Rev Biochem* 2012, **81**:203-229.

7. Ekkebus R *et al.*: **Catching a DUB in the act: novel ubiquitin-based active site directed probes**. *Curr Opin Chem Biol* 2014, **23**:63-70.

8. Fischle W, Mootz HD, Schwarzer D: **Synthetic histone code**. *Curr Opin Chem Biol* 2015, **28**:131-140.

9. Maity SK, Jbara M, Brik A: **Chemical and semisynthesis of modified histones**. *J Pept Sci* 2016, **22**:252-259 http://dx.doi.org/10.1002/psc.2848 [Epub 2016 Jan 18].

10. Pratt MR, Abeywardana T, Marotta NP: **Synthetic proteins and peptides for the direct interrogation of alpha-synuclein posttranslational modifications**. *Biomolecules* 2015, **5**:1210-1227.

11. Pham GH, Strieter ER: **Peeling away the layers of ubiquitin signaling complexities with synthetic ubiquitin–protein conjugates**. *Curr Opin Chem Biol* 2015, **28**:57-65.

12. Chatterjee C *et al.*: **Auxiliary-mediated site-specific peptide ubiquitylation**. *Angew Chem Int Ed Engl* 2007, **46**:2814-2818.

13. El Oualid F *et al.*: **Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin**. *Angew Chem Int Ed Engl* 2010, **49**:10149-10153.

14. Kumar KS *et al.*: **Total chemical synthesis of di-ubiquitin chains**. *Angew Chem Int Ed Engl* 2010, **49**:9126-9131.

15. Erlich LA *et al.*: **N-methylcysteine-mediated total chemical synthesis of ubiquitin thioester**. *Org Biomol Chem* 2010, **8**:2392-2396.

16. Shanmugham A *et al.*: **Nonhydrolyzable ubiquitin–isopeptide isosteres as deubiquitinating enzyme probes**. *J Am Chem Soc* 2010, **132**:8834-8835.

17. Jung JE *et al.*: **Functional ubiquitin conjugates with lysine-epsilon-amino-specific linkage by thioether ligation of cysteinyl-ubiquitin peptide building blocks**. *Bioconjug Chem* 2009, **20**:1152-1162.

18. Valkevich EM *et al.*: **Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases**. *J Am Chem Soc* 2012, **134**:6916-6919.

19. Trang VH *et al.*: **Nonenzymatic polymerization of ubiquitin: single-step synthesis and isolation of discrete ubiquitin oligomers**. *Angew Chem Int Ed Engl* 2012, **51**:13085-13088.

20. Mootz HD: **Split inteins as versatile tools for protein semisynthesis**. *Chembiochem* 2009, **10**:2579-2589.

21. Wals K, Ovaa H: **Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins**. *Front Chem* 2014, **2**:p15.

22. Lang K, Chin JW: **Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins**. *Chem Rev* 2014, **114**:4764-4806.

23. Virdee S *et al.*: **Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase**. *Nat Chem Biol* 2010, **6**:750-757.

24. Virdee S *et al.*: **Traceless and site-specific ubiquitination of recombinant proteins**. *J Am Chem Soc* 2011, **133**:10708-10711.

25. • Wang XA *et al.*: **E1-catalyzed ubiquitin C-terminal amidation for the facile synthesis of deubiquitinase substrates**. *Chembiochem* 2014, **15**:37-41.
This paper describes a relatively easy, semi-synthetic method to generate (fluorescent) DUB substrates via an E1-mediated amidation of the Ub C-terminus. Ub is linked to a fluorescent dye via a triazole or thioether linkage. In addition, thioether linked diubiquitin molecules can be made with this method.

26. •• Rosner D *et al.*: **Click chemistry for targeted protein ubiquitylation and ubiquitin chain formation**. *Nat Protoc* 2015, **10**:1594-1611.
This nature protocols paper describes how to incorporate azide and alkyne functionalities into a single ubiquitin moiety through genetic code expansion. This can be used to make triazole linked, non-hydrolysable polyubiquitin molecules and is, apart from [19], the only currently known method to make Ub polymers non-enzymatically. The ability to introduce azide, alkyne or both functionalities also allows site specific incorporation of (poly-)Ub on a substrate (e.g. on polymerase bèta); a process for which the concurrent E3 enzymes are presently unknown.

27. Schneider D *et al.*: **Improving bioorthogonal protein ubiquitylation by click reaction**. *Bioorg Med Chem* 2013, **21**:3430-3435.

28. Weikart ND, Sommer S, Mootz HD: **Click synthesis of ubiquitin dimer analogs to interrogate linkage-specific UBA domain binding**. *Chem Commun (Camb)* 2012, **48**:296-298.

29. Eger S *et al.*: **Generation of a mono-ubiquitinated PCNA mimic by click chemistry**. *Chembiochem* 2011, **12**:2807-2812.

30. Schneider T *et al.*: **Dissecting ubiquitin signaling with linkage-defined and protease resistant ubiquitin chains**. *Angew Chem Int Ed Engl* 2014, **53**:12925-12929.

31. Hospenthal MK, Freund SM, Komander D: **Assembly, analysis and architecture of atypical ubiquitin chains**. *Nat Struct Mol Biol* 2013, **20**:555-565.

32. Bremm A, Freund SM, Komander D: **Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne**. *Nat Struct Mol Biol* 2010, **17**:939-947.

33. Matsumoto ML *et al.*: **K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody**. *Mol Cell* 2010, **39**:477-484.

34. Williamson A *et al.*: **Identification of a physiological E2 module for the human anaphase-promoting complex**. *Proc Natl Acad Sci U S A* 2009, **106**:18213-18218.

35. Michel MA *et al.*: **Assembly and specific recognition of k29- and k33-linked polyubiquitin**. *Mol Cell* 2015, **58**:95-109.

36. Kristariyanto YA *et al.*: **K29-selective ubiquitin binding domain reveals structural basis of specificity and heterotypic nature of k29 polyubiquitin**. *Mol Cell* 2015, **58**:83-94.

37. Gatti M *et al.*: **RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage**. *Cell Rep* 2015, **10**:226-238.

38. •• Castaneda CA *et al.*: **Linkage-specific conformational ensembles of non-canonical polyubiquitin chains**. *Phys Chem Chem Phys* 2016, **18**:5771-5788.

This paper describes the structural diversity of all diubiquitin molecules by solution NMR. Structural determination of (di)ubiquitin (chain) conformation helps our understanding of how differently linked ubiquitin molecules can confer specific signals in the cell. Polyubiquitin is conformationally heterogenous, but each chain type exhibits unique conformations for that linkage. A number of diubiquitin molecules adopt conformation similar to previously solved K48-linked or K63-linked diubiquitin, suggesting potential overlap of biological functions between different lysine linkages.

39. • Castaneda CA *et al.*: **Linkage via K27 bestows ubiquitin chains with unique properties among polyubiquitins**. *Structure* 2016, **24**:423-436.

This paper describes the first complete characterization of K27 diubiquitin in particular. Interestingly the authors find that K27 diubiquitin is processed least by the DUBs they tested and also find that K27 diubiquitin can act as a competitive inhibitor of DUB activity towards other diubiquitin linkages. The paper provides a structural basis for why K27-linked diubiquitin is the most resistant diubiquitin linkage to deubiquitinating enzymes.

40. Faesen AC *et al.*: **The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types**. *Chem Biol* 2011, **18**:1550-1561.

41. Mevissen TE *et al.*: **OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis**. *Cell* 2013, **154**:169-184.

42. McGouran JF *et al.*: **Deubiquitinating enzyme specificity for ubiquitin chain topology profiled by di-ubiquitin activity probes**. *Chem Biol* 2013, **20**:1447-1455.

43. Mulder MP *et al.*: **A native chemical ligation handle that enables the synthesis of advanced activity-based probes: diubiquitin as a case study**. *Chembiochem* 2014, **15**:946-949.

44. Li G *et al.*: **Activity-based diubiquitin probes for elucidating the linkage specificity of deubiquitinating enzymes**. *Chem Commun (Camb)* 2014, **50**:216-218.

45. Haj-Yahya N *et al.*: **Dehydroalanine-based diubiquitin activity probes**. *Org Lett* 2014, **16**:540-543.

46. •• Flierman D *et al.*: **Non-hydrolyzable diubiquitin probes reveal linkage-specific reactivity of deubiquitylating enzymes mediated by S2 pockets**. *Cell Chem Biol* 2016, **23**:472-482 http://dx.doi.org/10.1016/j.chembiol.2016.03.009 [Epub 2016 Apr 7].

This paper describes a diubiquitin-based, non-hydrolysable DUB probe with a covalent warhead at the C-terminus which can be used to discriminate S2 binding sites on DUBs. DUBs can either remove ubiquitin chains as a whole or cleave the Ub moieties one by one. This probe allows investigation of DUBs that recognize ubiquitin chains and cleave ubiquitin moieties larger than a diubiquitin module.

47. Bekes M *et al.*: **Recognition of Lys48-linked di-ubiquitin and deubiquitinating activities of the SARS coronavirus papain-like protease**. *Mol Cell* 2016, **62**:572-585.

48. Wiener R *et al.*: **The mechanism of OTUB1-mediated inhibition of ubiquitination**. *Nature* 2012, **483**:618-622.

49. Samara NL *et al.*: **Structural insights into the assembly and function of the SAGA deubiquitinating module**. *Science* 2010, **328**:1025-1029.

50. Ratia K *et al.*: **Structural basis for the ubiquitin-linkage specificity and deISGylating activity of SARS-CoV papain-like protease**. *PLoS Pathog* 2014, **10**:pe1004113.

51. Hu M *et al.*: **Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14**. *EMBO J* 2005, **24**:3747-3756.

52. Hu M *et al.*: **Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde**. *Cell* 2002, **111**:1041-1054.

53. Johnston SC *et al.*: **Structural basis for the specificity of ubiquitin C-terminal hydrolases**. *EMBO J* 1999, **18**:3877-3887.

54. Sheedlo MJ *et al.*: **Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination**. *Proc Natl Acad Sci U S A* 2015, **112**:15090-15095.

55. Schlieker C *et al.*: **Structure of a herpesvirus-encoded cysteine protease reveals a unique class of deubiquitinating enzymes**. *Mol Cell* 2007, **25**:677-687.

56. Boudreaux DA *et al.*: **Ubiquitin vinyl methyl ester binding orients the misaligned active site of the ubiquitin hydrolase UCHL1 into productive conformation**. *Proc Natl Acad Sci U S A* 2010, **107**:9117-9122.

57. Misaghi S *et al.*: **Structure of the ubiquitin hydrolase UCH-L3 complexed with a suicide substrate**. *J Biol Chem* 2005, **280**:1512-1520.

58. Artavanis-Tsakonas K *et al.*: **Characterization and structural studies of the *Plasmodium falciparum* ubiquitin and Nedd8 hydrolase UCHL3**. *J Biol Chem* 2010, **285**:6857-6866.

59. Morrow ME *et al.*: **Stabilization of an unusual salt bridge in ubiquitin by the extra C-terminal domain of the proteasome-associated deubiquitinase UCH37 as a mechanism of its exo specificity**. *Biochemistry* 2013, **52**:3564-3578.

60. Sahtoe DD *et al.*: **Mechanism of UCH-L5 activation and inhibition by DEUBAD domains in RPN13 and INO80G**. *Mol Cell* 2015, **57**:887-900.

61. Ekkebus R *et al.*: **On terminal alkynes that can react with active-site cysteine nucleophiles in proteases**. *J Am Chem Soc* 2013, **135**:2867-2870.

62. Altun M *et al.*: **The human otubain2-ubiquitin structure provides insights into the cleavage specificity of poly-ubiquitin-linkages**. *PLoS One* 2015, **10**:pe0115344.

63. • An H, Statsyuk AV: **Facile synthesis of covalent probes to capture enzymatic intermediates during E1 enzyme catalysis**. *Chem Commun (Camb)* 2016, **52**:2477-2480.

This paper describes a covalent probe to capture E1 enzymes via a tetrahedral intermediate mimick of the E1–UbL–AMP complex. The strategy was also applied on the ubiquitin-like (UbL) protein LC3 and is of use for labeling E1 enzymes of other UbL proteins as well.

64. •• Pao KC *et al.*: **Probes of ubiquitin E3 ligases enable systematic dissection of parkin activation**. *Nat Chem Biol* 2016, **12**:324-331 http://dx.doi.org/10.1038/nchembio.2045 [Epub 2016 Mar 7].

This study describes the elegant development of an E3 ligase probe based on structural and mechanistic knowledge. The probe is subsequentially used to investigate the E3 ligase Parkin in detail: it allows quantification of Parkin transthiolation activity and detection of endogenous phosphorylated Parkin. Its potential use as a biomarker for early onset of Parkinson's Disease is also investigated.

65. •• Mulder MP *et al.*: **A cascading activity-based probe sequentially targets E1–E2–E3 ubiquitin enzymes**. *Nat Chem Biol* 2016 http://dx.doi.org/10.1038/nchembio.2084. [Epub ahead of print].
This paper describes the development of a Ub-dehydroalanine-based ligase probe which can be passed on in the ligation cascade to E1 and E2 enzymes, in addition to E3 ligases of the HECT/RBR type. The advantage is that this probe can label E1, E2 and E3 enzymes, a feature not possible with other probes. The authors demonstrate the utility of this probe by its electroporation into cells to show specific labeling of a panel of Ub ligases *in vivo*.

66. Cui J *et al.*: **Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family**. *Science* 2010, **329**:1215-1218.

67. Ohtake F *et al.*: **Ubiquitin acetylation inhibits polyubiquitin chain elongation**. *EMBO Rep* 2015, **16**:192-201.

68. Eiyama A, Okamoto K: **PINK1/Parkin-mediated mitophagy in mammalian cells**. *Curr Opin Cell Biol* 2015, **33**:95-101.

69. Chin LS, Li L: **Ubiquitin phosphorylation in Parkinson's disease: implications for pathogenesis and treatment**. *Transl Neurodegener* 2016, **5**:p1.

70. • Bondalapati S *et al.*: **Chemical synthesis of phosphorylated ubiquitin and diubiquitin exposes positional sensitivities of e1–e2 enzymes and deubiquitinases**. *Chemistry* 2015, **21**:7360-7364.
This paper describes, for the first time, the complete synthesis of phosphorylated ubiquitin and phosphorylated K63-linked diubiquitin. DUB cleavage assays show that DUBs may exhibit different specificities towards distal or proximal phosphorylated diubiquitin molecules.

71. Wauer T *et al.*: **Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis**. *EMBO J* 2015, **34**:307-325.

72. • Han C *et al.*: **A versatile strategy for the semisynthetic production of Ser65 phosphorylated ubiquitin and its biochemical and structural characterisation**. *Chembiochem* 2015, **16**:1574-1579.
This paper describes a method to make phosphorylated ubiquitin via a semi-synthetic approach by ligation of a N-terminal, expressed ubiquitin(1–45) thioester moiety to a synthetic ubiquitin(46–75) module. The synthetic C-terminal part can be readily synthesized by a commercial service with differently incorporated phosphorylated amino acids, making investigation of potential C-terminal phosphorylation sites feasible.

73. Nie M, Boddy MN: **Cooperativity of the SUMO and ubiquitin pathways in genome stability**. *Biomolecules* 2016, **6**.

74. Leidecker O *et al.*: **The ubiquitin E1 enzyme Ube1 mediates NEDD8 activation under diverse stress conditions**. *Cell Cycle* 2012, **11**:1142-1150.

75. Buchsbaum S, Bercovich B, Ciechanover A: **FAT10 is a proteasomal degradation signal that is itself regulated by ubiquitination**. *Mol Biol Cell* 2012, **23**:225-232.

76. Fan JB *et al.*: **Identification and characterization of a novel ISG15-ubiquitin mixed chain and its role in regulating protein homeostasis**. *Sci Rep* 2015, **5**:p12704.

77. Haller M *et al.*: **Ubiquitination and proteasomal degradation of ATG12 regulates its proapoptotic activity**. *Autophagy* 2014, **10**:2269-2278.

78. • Singh RK, Sundar A, Fushman D: **Nonenzymatic rubylation and ubiquitination of proteins for structural and functional studies**. *Angew Chem Int Ed Engl* 2014, **53**:6120-6125.
This paper describes a semi-synthetic method similar to GOPAL to make different heterlogous ubiquitin–Rub1 conjugates and can be readily used in any biochemical laboratory. It uses a chemical condensation reaction between a thioester on the C-terminus of one Ub(-like) protein and the ε-amine of a lysine residue on the other Ub(-like) protein by using two orthogonal protecting groups: Boc and Alloc. Since both proteins are made via bacterial expression they can be ^{15}N labeled and easily characterized by NMR. This strategy could also be applied for (structural) characterization of other Ub/UbL conjugates.

79. Boll E *et al.*: **A novel PEG-based solid support enables the synthesis of > 50 amino-acid peptide thioesters and the total synthesis of a functional SUMO-1 peptide conjugate**. *Chem Sci* 2014, **5**:2017-2022.

80. • Boll E *et al.*: **One-pot chemical synthesis of small ubiquitin-like modifier protein-peptide conjugates using bis(2-sulfanylethyl)amido peptide latent thioester surrogates**. *Nat Protoc* 2015, **10**:269-292.
This paper describes the one-pot, synthetic route to make functional SUMO-peptide conjugates. It uses a SAE PEG-based resin which allows synthesis of longer SAE latent thioesters and thioester peptides of approx. 50 amino acids by Fmoc-SPPS. The advantage of this method is the relative ease with which longer peptides can be synthesized; as long as they have internal cysteine residues that are up to maximally 50 amino acids apart. The cysteine residue is used as acceptor site for the thioester containing-moiety in native chemical ligation.

81. Ogunkoya AO, Pattabiraman VR, Bode JW: **Sequential alpha-ketoacid-hydroxylamine (KAHA) ligations: synthesis of C-terminal variants of the modifier protein UFM1**. *Angew Chem Int Ed Engl* 2012, **51**:9693-9697.