

Expanded Click Conjugation of Recombinant Proteins with Ubiquitin-Like Modifiers Reveals Altered Substrate Preference of SUMO2-Modified Ubc9**

Stefanie Sommer, Nadine D. Weikart, Andreas Brockmeyer, Petra Janning, and Henning D. Mootz*

Posttranslational modification of proteins with ubiquitin (Ub) and ubiquitin-like (Ubl) modifiers, such as SUMO (small ubiquitin-like modifiers), play an essential role in the regulation of numerous cellular pathways, which include proteasomal degradation, nuclear transport, transcription, and DNA repair.^[1–5] Ubl modifiers are typically conjugated by an isopeptide bond, which is formed between the C-terminal carboxyl group of the modifier and the ϵ -amino group of an acceptor lysine (Lys) residue in the substrate protein. Ubiquitin can form poly-ubiquitin chains through all of its seven lysine units.^[6] Humans encode four SUMO isoforms, namely SUMO1 to SUMO4. The highly homologous SUMO2 and SUMO3 isoforms can form poly-SUMO chains through Lys11.^[7–9] Enzymatic modification of a specific protein with a Ubl modifier is brought about by an E1-activating enzyme, an E2-conjugating enzyme, and an E3 ligase.^[1–5] While only a single E1 enzyme exists for each Ubl modifier, in the case of the ubiquitin-conjugation pathway, dozens of E2 and hundreds of E3 enzymes are found.^[1] This large number of proteins reflects the complexity required for a highly regulated ubiquitylation of thousands of different proteins. In contrast, there is only a single SUMO-specific E2 enzyme (Ubc9), and only a relatively small number of SUMO E3 ligases have been reported to date,^[5,10] even though over 750 proteins have been identified as targets of SUMOylation.^[3,11] Ubc9 exhibits a reasonable affinity to the canonical SUMOylation motif Ψ KxE (Ψ = hydrophobic residue, x = any amino acid, K = lysine, E = glutamic acid). By using purified components, substrate proteins can be modified in an E3-independent fashion,^[12] while E3 ligases are required for efficient SUMO transfer in vivo.^[13] An important question with regard to the comparably simple SUMOylation machi-

nery is how modification of the large and diverse set of proteins with this posttranslational modification is regulated. One potential mechanism is the reversion of Ubl conjugation by the action of Ubl-specific deconjugating enzymes. However, while 80–100 of these enzymes are known for ubiquitin,^[14] only seven have been identified to date for SUMO in humans.^[15] Another recently proposed mechanism suggests the modulation of Ubc9 preference for different substrates by SUMOylation of Ubc9 itself, as shown for Ubc9 modified with SUMO1 on Lys14.^[16]

A serious hurdle for studying the biochemical, structural, and cellular consequences of Ubl modification is the inability to prepare defined and homogenous Ubl conjugates by enzymatic means for any given Ubl target, that is, the site-specific introduction of a Ubl modifier at a desired site in the substrate protein.^[17,18] This problem arises from the incomplete understanding of the conjugation pathway, non-exclusive specificity of the involved E2 and E3 enzymes for certain sites in proteins, and often substoichiometric reaction yields. To solve this problem, several chemical approaches for Ubl conjugation have been introduced very recently. By using native chemical ligation (NCL), in these cases also referred to as isopeptide chemical ligation (ICL),^[19–26] unnatural amino acid incorporation,^[25,27] selective protein chemistry in combination with bioorthogonal conjugation reactions,^[28–30] alkylation,^[31,32] or disulfide bond formation,^[33,34] various approaches have been developed to form either the native isopeptide bonds or structurally similar artificial linkages. We have previously described the copper(I)-catalyzed cycloaddition between an azide and an alkyne (CuAAC; well known as the “click” reaction) to generate a triazole linkage between a Ubl modifier and a tag protein that bears a short SUMO-acceptor peptide sequence.^[28] This linkage is stable against isopeptidases.

Our goal in this study was to synthesize a defined Ubc9**SUMO2 conjugate at position 14 of Ubc9 in order to investigate the role of this modification in target discrimination of the E2 enzyme (** = triazole linkage). However, Ubc9 (158 amino acids) represents a very challenging protein and the conjugate could not easily be prepared with any of the aforementioned chemical approaches for the following reasons. Our previously reported CuAAC approach relied on the chemoselective modification of a single cysteine (Cys) residue at the site of the acceptor lysine (for the installation of an azide functional group),^[28] and is therefore limited to proteins that are free of additional cysteine units. However, Ubc9 contains several cysteine residues that cannot be removed by

[*] MSc. S. Sommer, MSc. N. D. Weikart, Prof. Dr. H. D. Mootz
Institute of Biochemistry—University of Münster
Wilhelm-Klemm-Strasse 2, 48149 Münster (Germany)
E-mail: henning.mootz@uni-muenster.de

Dipl.-Ing. A. Brockmeyer, Dr. P. Janning
Department of Chemical Biology
Max-Planck-Institute for Molecular Physiology
Otto-Hahn-Strasse 11, 44227 Dortmund (Germany)

[**] We thank Frauke Melchior and Andrea Pichler for discussions and Peter G. Schultz, F. Melchior, A. Pichler, and Christina Hecker for providing plasmids. We gratefully acknowledge financial support by Fonds der Chemischen Industrie and IMPRS Dortmund (Ph.D. fellowships to S.S. and N.D.W.).



Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.201102531>.

mutation without loss of activity; these residues include the essential catalytic Cys93, which forms the SUMO–thioester intermediate. To date, all reported Ubl isopeptide bond syntheses by ICL or chemical coupling involve strongly denaturing conditions, presumably to enable efficient reaction progress by high reactant concentrations and improved access of the functional groups.^[19–27] However, such protocols can only be applied to proteins that can be refolded, and indeed, to date these approaches have only been shown for such substrates, for example, for the synthesis of diubiquitins and a few monoubiquitinated target proteins.^[35] Our initial tests revealed that Ubc9 could be renatured only very poorly, thus ruling out this synthesis route. Finally, Ubl conjugation through disulfide bond formation is an attractive route, however, again this approach would require a single cysteine residue in the substrate protein. Moreover, the disulfide bond would not be stable to reducing conditions, which are necessary for the E2 enzyme Ubc9 in order to keep the catalytic cysteine in the active thiol form. In this work, we overcame these problems by an expanded protocol for click conjugation that included the incorporation of the azide moiety by using unnatural amino acid mutagenesis, and of the alkyne by using protein thioester aminolysis. This is the first report of the preparation of an active enzyme with a chemically conjugated Ubl modification.

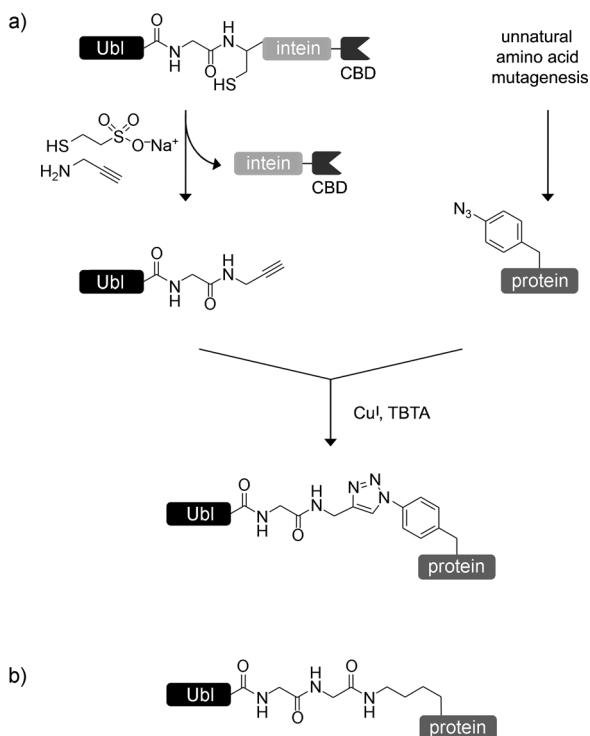
Scheme 1 A shows our expanded strategy for the chemical click conjugation of Ubl modifiers. The attachment of the alkyne group to the Ubl proteins is achieved by an intein fusion strategy with aminolysis of the Ubl thioester by propargylamine. To resemble the length of the native

isopeptide bond (Scheme 1 B) as closely as possible, only one residue of the conserved diglycine motif at the C terminus of the Ubl protein was deleted, in contrast to our previous report.^[28] To selectively introduce the azide functional group for the CuAAC reaction in the presence of cysteines, we turned to the incorporation of *p*-azidophenylalanine (AzF) by unnatural amino acid mutagenesis at the position of the acceptor lysine. The CuAAC reaction with a Ubl alkyne protein would then furnish the desired conjugate.

Incorporation of AzF was previously reported by Schultz and co-workers, and exploited for the attachment of small molecule labels or poly(ethylene glycol) with a subsequent CuAAC reaction.^[36,37] This strategy should be compatible with native cysteine residues in the protein.^[38] To test if this strategy was also useful for the efficient conjugation of the macromolecular Ubl alkyne as a protein modifier, we first sought to test the reaction on the AzF-functionalized model protein HA-GpD-PML₁₁/AzF-SBP (construct **1**; HA = hemagglutinin epitope, SBP = streptavidin binding peptide) as illustrated in Figure 1 A. In this conjugate, the λ -phage coat protein GpD served as a well-expressing tag protein, while the 11 aa peptide sequence PML₁₁/AzF corresponded to the SUMOylation site of the RING finger/promyelocytic leukemia protein PML^[39] in which the acceptor lysine was changed to AzF (PRKVI-AzF-MESEE). Construct **1** could be produced in good yields in *E. coli* by coexpression of the aminoacyl-tRNA synthetase and tRNA_{CUA} specific for AzF (the plasmid was kindly provided by Peter G. Schultz) in the presence of AzF (1 mM) in the growth medium (Figure 1 B). After cell lysis, **1** was purified on streptactin beads (see Figure S1 in the Supporting Information). For control purposes, the protein with the native lysine in the PML₁₁ peptide was also prepared (construct **2**). Figure S2 in the Supporting Information shows that His-tagged SUMO2(Δ G) alkyne (construct **3**) could be obtained in good purity. CuAAC reactions between **1** and **3** in a 2:1 molar ratio furnished the chemical SUMO2 conjugate **4** in good yields of 40 to 50 % after 30 min (calculated relative to the limiting reaction partner **3**, see Figure 1 C). The correct structure of purified conjugate **4** was further confirmed through immunoblotting by using anti-HA and anti-His antibodies (data not shown).

In order to investigate the triazole linkage by mass spectrometry, we introduced the mutation Q98R into SUMO2(Δ G) to obtain a smaller fragment after trypsin digestion. The peptide of the corresponding SUMO2(Δ G)-(Q98R) conjugate was identified by ESI-TOF-MS and successfully sequenced by MS/MS analysis (Figure 1 D, see also Figure S5 and Table S1 in the Supporting Information), thus unambiguously confirming the expected triazole linkage.

Similarly, by using an alkyne-modified SBP-tagged Ub(Δ G) (construct **5**) in the CuAAC reaction, the corresponding ubiquitin conjugate **6** could be prepared in similar yields (Figure 1 E, see also Figure S6 in the Supporting Information). Taken together, these results show that a Ubl modifier can be site-specifically conjugated by the CuAAC reaction to a recombinantly expressed protein that contains AzF, and that this approach is of general use for the family of Ubl proteins.



Scheme 1. General concept of the conjugation approach. a) Synthesis of the chemical conjugate. b) Native conjugate. CBD = chitin-binding domain, TBTA = tris(benzyltriazolylmethyl)amine.

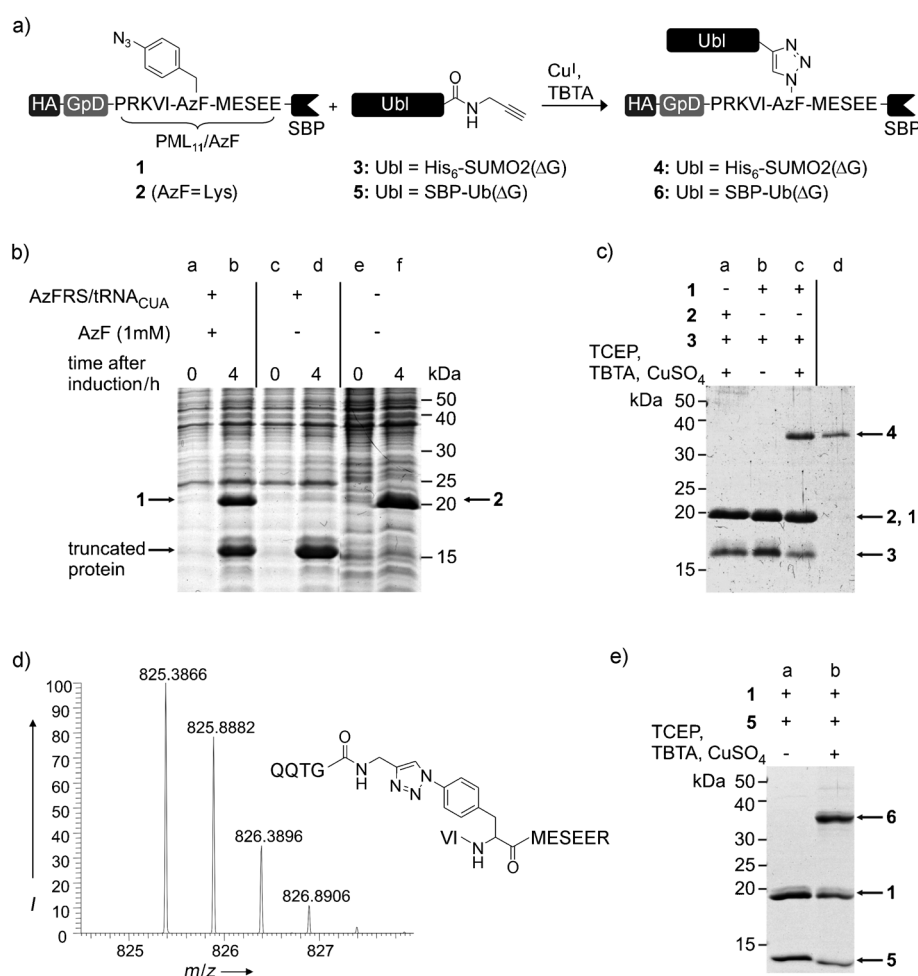


Figure 1. Expanded click conjugation by using a model protein. a) Reaction scheme. b) Selective incorporation of AzF by using amber stop codon suppression technology (lanes a–d). A coomassie-stained SDS-PAGE gel of *E. coli* total cell extracts is shown. The expression of the full length protein (1) in the presence of the orthogonal pAzFRS/tRNA_{CUA} pair was dependent on the addition of AzF (1 mM) in the growth medium. Lanes e, f show the expression of the control protein 2 with the codon for the regular acceptor lysine in the PML₁₁ sequence. c) Cu^I-catalyzed cycloaddition with SUMO2(ΔG) alkyne (3). A coomassie-stained SDS-PAGE gel of the purified proteins is shown. The reactions with the indicated protein components were performed at RT for 30 min. Lane d shows the purified synthetic conjugate 4. d) ESI-TOF-MS analysis of a tryptic digest of the SUMO2(ΔG)(Q98R) variant of 4. The signal for the fragment that contains the triazole linkage (see right panel) is shown. Signals for the doubly charged ion: $m/z_{calc.}$ 825.3892, $m/z_{obs.}$ 825.3866. e) Cu^I-catalyzed cycloaddition between model protein (1) and Ub(ΔG) alkyne (5). The reaction was performed at RT for 30 min. A coomassie-stained SDS-PAGE gel of the purified proteins is shown. TCEP = tris(2-carboxyethyl)phosphine.

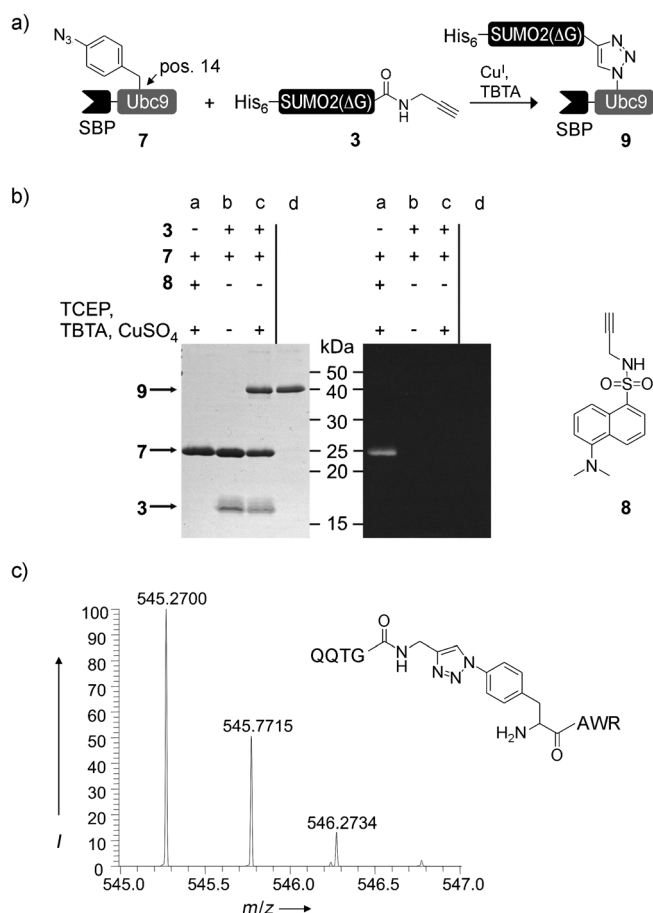
With the expanded protocol for click conjugation of Ubl modifiers established, we turned to the synthesis of the desired Ubc9**SUMO2 conjugate, which involves position 14 of Ubc9 (see Figure 2 A). Intriguingly, although the formation of an enzymatically generated Ubc9*SUMO2 conjugate (* = isopeptide linkage) was confirmed by mass spectrometry, a biochemical characterization of such a conjugate has not yet been reported, in contrast to its SUMO1 counterpart.^[40] In fact, the enzymatic preparation of such a defined mono-SUMOylated Ubc9 by enzymatic means is not straightforward because of poor self-SUMOylation of Ubc9, the concomitant SUMOylation on other positions of Ubc9 including Lys49 and Lys153,^[40] and the capability of

SUMO2 to form poly-SUMO2 chains. Thus, Ubc9**SUMO2 represents an ideal and challenging target for chemical conjugation.

To incorporate AzF into human Ubc9, the encoding gene was mutated to contain a TAG stop codon at position 14. An N-terminal SBP tag was added for facile protein purification. The full-length protein was efficiently produced in *E. coli* grown in minimal medium supplemented with AzF (1 mM), thus suggesting suppression of the amber codon with AzF, while in the absence of AzF no expression was observed (see Figure S7 in the Supporting Information). Pure SBP-Ubc9(AzF14) (construct 7) was obtained in a yield of about 6 mg L⁻¹ by affinity chromatography on a streptactin resin (see Figure S8 in the Supporting Information). The functional integrity and accessibility of AzF14 could be demonstrated by efficient attachment of alkyne-functionalized dansylamide (8; Figure 2 B, right-hand panel) under CuAAC conditions (Figure 2 B, lane a). More importantly, the CuAAC reaction between SBP-Ubc9(AzF14) and the macromolecular SUMO2(ΔG) alkyne (3) in a 2:1 molar ratio also proceeded in good yields of about 60% (calculated relative to the amount of 3; Figure 2 B, lane c). Importantly, this reaction could be carried out under non-denaturing conditions and with a short reaction time of only 30 min at room temperature, in order to preserve Ubc9 integrity. The resulting chemical conjugate Ubc9**SUMO2 (construct 9) was subsequently purified by affinity purification on Ni-NTA

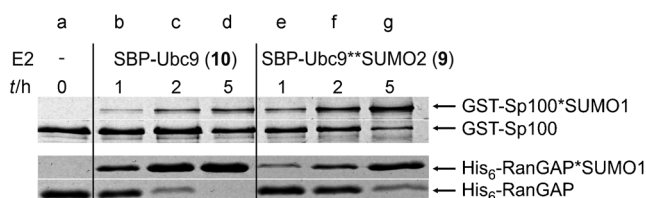
and streptactin resin followed by size-exclusion chromatography to apparent homogeneity (Figure 2 B, lane d). The identity of the desired linkage at position 14 of Ubc9 could be unambiguously verified by ESI-TOF-MS analysis of a tryptic digest (Figure 2 C; for the sequencing of the fragment by MS/MS see Figure S9 and Table S2 in the Supporting Information); again by using the corresponding conjugate with the SUMO2(ΔG)(Q98R) mutant.

Finally, we tested the catalytic activity of the chemical Ubc9**SUMO2 construct (9) and investigated the effect of the SUMO2 modification at position 14 on potential target discrimination of Ubc9. It was previously reported that the corresponding SUMO1 conjugate Ubc9*SUMO1 exhibited



higher SUMOylation activity for Sp100 and lower activity for RanGAP, when compared to the unmodified Ubc9.^[16] The increase in activity for Sp100 was shown to be dependent on the additional binding interface provided by the conjugated SUMO1 and the SUMO-interacting motif (SIM) of Sp100. Although it is known that in principle SUMO2 can engage in SIM-SUMO interactions in a similar fashion as SUMO1,^[41] it was unclear if a modification of Ubc9 with SUMO2 would have a similar effect. Alternatively, other interactions specific for SUMO1 and SUMO2 might play an important role in the molecular recognition, because of their partially distinct functions^[42] and sequence identity of only 44 %.

We assayed SUMOylation activity of equal amounts of native SBP-Ubc9 (**10**) and the chemical SBP-Ubc9**SUMO2 conjugate (**9**) in reactions that included Sp100 or RanGAP as substrate proteins as well as the E1 enzyme Aos1/Uba2, SUMO1, and ATP. Figure 3 shows that, in comparison to the



unmodified Ubc9, Ubc9**SUMO2 displayed decreased activity for RanGAP, but more efficiently modified Sp100. While the decreased activity for RanGAP alone could also be attributed to a general loss of activity because of our chemical conjugation procedure, the increased activity for Sp100 clearly excludes this possibility as the only cause and supports the regulatory effect of the attached SUMO2. Thus, Ubc9**SUMO2 displayed an altered substrate preference toward Sp100 and RanGAP that is very similar to the previously reported Ubc9*SUMO1.

In conclusion, we report a new triazole linkage for the conjugation of Ubl modifiers. Our results show that Ubc9 modification with SUMO at position 14 is a general mechanism to regulate target discrimination of this E2 enzyme. SUMO1 and SUMO2 act similarly in this regard. We show that an active Ubl-modified enzyme could be prepared by a combination of unnatural amino acid mutagenesis and click chemistry. This result implies that exposure of Ubc9, and in particular its catalytic Cys93, to the copper ions under the conditions of the CuAAC reaction was well tolerated, although we cannot rule out an effect of residual copper ions on other proteins.^[43] In terms of structural similarity between the triazole linkage and the native isopeptide bond, our approach is minimally invasive, as only the terminal glycine of the Ubl modifier and the acceptor lysine residue are altered. Importantly, our biochemical data suggests that this triazole linkage presents the Ubl modifier in a proper distance and orientation relative to the target protein. We therefore believe that click conjugation provides a suitable isopeptide-bond analogue useful for the biochemical characterization of many proteins conjugated to a Ubl modifier. The Ubc9**SUMO2 conjugate could not have been efficiently synthesized with any other of the existing approaches. Our approach holds the potential for the synthesis of other demanding Ubl-protein conjugates that cannot be refolded or are inaccessible by other techniques because of constraints in the amino acid composition, in particular the presence of cysteine residues.

Received: April 12, 2011

Revised: July 29, 2011

Published online: September 7, 2011

Keywords: bioorthogonal reactions · click chemistry · conjugating enzymes · protein modification · ubiquitin

- [1] M. Hochstrasser, *Nature* **2009**, 458, 422.
- [2] C. M. Pickart, *Annu. Rev. Biochem.* **2001**, 70, 503.
- [3] R. J. Dohmen, *Biochim. Biophys. Acta Mol. Cell Res.* **2004**, 1695, 113.
- [4] R. Geiss-Friedlander, F. Melchior, *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 947.
- [5] H. D. Ulrich, *Methods Mol. Biol.* **2009**, 497, 3.
- [6] D. Komander, *Biochem. Soc. Trans.* **2009**, 37, 937.
- [7] I. Matic, M. van Hagen, J. Schimmel, B. Macek, S. C. Ogg, M. H. Tatham, R. T. Hay, A. I. Lamond, M. Mann, A. C. Vertegaal, *Mol. Cell. Proteomics* **2008**, 7, 132.
- [8] H. D. Ulrich, *Mol. Cell* **2008**, 32, 301.
- [9] A. C. O. Vertegaal, *Biochem. Soc. Trans.* **2010**, 38, 46.
- [10] Y. Chu, X. Yang, *Oncogene* **2011**, 30, 1108.
- [11] F. Golebiowski, I. Matic, M. H. Tatham, C. Cole, Y. Yin, A. Nakamura, J. Cox, G. J. Barton, M. Mann, R. T. Hay, *Sci. Signaling* **2009**, 2, ra24.
- [12] T. Okuma, R. Honda, G. Ichikawa, N. Tsumagari, H. Yasuda, *Biochem. Biophys. Res. Commun.* **1999**, 254, 693.
- [13] E. S. Johnson, A. A. Gupta, *Cell* **2001**, 106, 735.
- [14] A. Y. Amerik, M. Hochstrasser, *Biochim. Biophys. Acta Mol. Cell Res.* **2004**, 1695, 189.
- [15] R. T. Hay, *Trends Cell Biol.* **2007**, 17, 370.
- [16] P. Knipscheer, A. Flotho, H. Klug, J. V. Olsen, W. J. van Dijk, A. Fish, E. S. Johnson, M. Mann, T. K. Sixma, A. Pichler, *Mol. Cell* **2008**, 31, 371.
- [17] L. J. Martin, R. T. Raines, *Angew. Chem.* **2010**, 122, 9226; *Angew. Chem. Int. Ed.* **2010**, 49, 9042.
- [18] T. Fekner, X. Li, M. K. Chan, *ChemBioChem* **2011**, 12, 21.
- [19] C. Chatterjee, R. K. McGinty, J. P. Pellois, T. W. Muir, *Angew. Chem.* **2007**, 119, 2872; *Angew. Chem. Int. Ed.* **2007**, 46, 2814.
- [20] R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, 453, 812.
- [21] R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu, C.-F. Liu, *J. Am. Chem. Soc.* **2009**, 131, 13592.
- [22] K. S. Ajish Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel, A. Brik, *Angew. Chem.* **2009**, 121, 8234; *Angew. Chem. Int. Ed.* **2009**, 48, 8090.
- [23] K. S. Ajish Kumar, L. Spasser, L. A. Erlich, S. N. Bavikar, A. Brik, *Angew. Chem.* **2010**, 122, 9312; *Angew. Chem. Int. Ed.* **2010**, 49, 9126.
- [24] F. El Oualid, R. Merckx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, T. K. Sixma, H. Ovaa, *Angew. Chem.* **2010**, 122, 10347; *Angew. Chem. Int. Ed.* **2010**, 49, 10149.
- [25] X. Li, T. Fekner, J. J. Ottesen, M. K. Chan, *Angew. Chem.* **2009**, 121, 9348; *Angew. Chem. Int. Ed.* **2009**, 48, 9184.
- [26] K. S. Ajish Kumar, L. Spasser, S. Ohayon, L. A. Erlich, A. Brik, *Bioconjugate Chem.* **2011**, 22, 137.
- [27] S. Virdee, Y. Ye, D. P. Nguyen, D. Kommander, J. W. Chin, *Nat. Chem. Biol.* **2010**, 6, 750.
- [28] N. D. Weikart, H. D. Mootz, *ChemBioChem* **2010**, 11, 774.
- [29] S. Eger, M. Scheffner, A. Marx, M. Rubini, *J. Am. Chem. Soc.* **2010**, 132, 16337.
- [30] A. Shanmugham, A. Fish, M. P. A. Luna-Vargas, A. C. Faesen, F. El Oualid, T. K. Sixma, H. Ovaa, *J. Am. Chem. Soc.* **2010**, 132, 8834.
- [31] N. S. Russell, K. D. Wilkinson, *Biochemistry* **2004**, 43, 4844.
- [32] J. E. Jung, H. P. Wollscheid, A. Marquardt, M. Manea, M. Scheffner, M. Przybylski, *Bioconjugate Chem.* **2009**, 20, 1152.
- [33] C. Chatterjee, R. K. McGinty, B. Fierz, T. W. Muir, *Nat. Chem. Biol.* **2010**, 6, 267.
- [34] J. Chen, Y. Ai, J. Wang, L. Haracska, Z. Zhuang, *Nat. Chem. Biol.* **2010**, 6, 270.
- [35] M. Hejjaoui, M. Haj-Yahya, K. S. Ajish Kumar, A. Brik, H. A. Lashuel, *Angew. Chem.* **2011**, 123, 425; *Angew. Chem. Int. Ed.* **2011**, 50, 405.
- [36] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, 124, 9026.
- [37] A. Deiters, A. Cropp, D. Summerer, M. Mukherji, P. G. Schultz, *Bioorg. Med. Chem. Lett.* **2004**, 14, 5743.
- [38] T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem.* **2009**, 121, 1661; *Angew. Chem. Int. Ed.* **2009**, 48, 1633.
- [39] M. H. Tatham, E. Jaffray, O. A. Vaughan, J. M. Desterro, C. H. Botting, J. H. Naismith, R. T. Hay, *J. Biol. Chem.* **2001**, 276, 35368.
- [40] S. Subramaniam, R. G. Mealer, K. M. Sixt, R. K. Barrow, A. Usiello, S. H. Synder, *J. Biol. Chem.* **2010**, 285, 20428.
- [41] J. Song, L. K. Durrin, T. A. Wilkinson, T. G. Krontiris, Y. Chen, *Proc. Natl. Acad. Sci. USA* **2004**, 101, 14373.
- [42] H. Saitoh, J. Hinchey, *J. Biol. Chem.* **2000**, 275, 6252.
- [43] S. Schoffelen, M. H. Lambermon, M. B. van Eldijk, J. C. van Hest, *Bioconjugate Chem.* **2008**, 19, 1127.