Highly Efficient and Chemoselective Peptide Ubiquitylation**

K. S. Ajish Kumar, Mahmood Haj-Yahya, Diana Olschewski, Hilal A. Lashuel, and Ashraf Brik

Post-translational modifications play an important role in regulating protein structure and function in health and disease.[1] Ubiquitylation is one example for such a modification wherein both the extent (poly- versus mono-ubiquitylation) and the sequence position of this modification dictates the function and fate of the ubiquitylated protein.[2–4] In the ubiquitylation process, three distinct enzymes, known as the E1–E3 system, collaborate to achieve a site-specific tagging of the lysine residue(s) in target protein.[5] This condensation step generates an isopeptide linkage between the ε-NH$_2$ of the lysine residue and the activated C-terminal glycine of ubiquitin (Ub). The overwhelming majority of the studies in the field rely on the in vitro enzymatic reconstitution of this complex posttranslational modification for the protein of interest. This process is often challenged by the heterogeneity of the modified protein, the isolation of the specific ligase (E3), and obtaining reasonable quantities of the ubiquitylated protein. In principle, adopting chemical and semisynthetic strategies for the preparation of posttranslationally modified protein could overcome these challenges and allow for site-specific ubiquitylation of the modified protein in sufficient quantities for biochemical and structural studies. With respect to Ub, early efforts in these directions include the total synthesis of Ub using stepwise solid-phase peptide synthesis[6] (SPPS) and native chemical ligation (NCL),[7] semisynthesis of Ub analogues with C-terminal electrophile,[8] and chemical manipulation of the polyubiquitin chain.[9]

Recently, Muir and co-workers devised an elegant semisynthetic approach that allows site-specific peptide[10] and protein mono-ubiquitylation.[11] This strategy relies on attaching a photoremovable auxiliary[12] to a lysine residue to mediate isopeptide formation with the Ub-α-thioester. This concept was then applied to synthesize a homogeneous ubiquitylated histone H2B, which was used for the assembly of the nucleosome in vitro. This study paved the way to unravel the effect of mono-ubiquitylation on the mechanism of stimulation of the Dot1-catalyzed histone H3 methylation.[11] Despite the success in obtaining a purely modified ubiquitylated histone H2B, the crucial step in the synthesis, that is, the isopeptide formation mediated by the photoremovable auxiliary, proceeded slowly, and required five days for ubiquitylation and seven days for SUMOylation (small ubiquitin-like modification) to achieve a reasonable conversion to the mono-ubiquitylated product.[13] The slow ligation rate was attributed to the involvement of the secondary amine in the S-N acyl transfer, despite proceeding through a five-membered ring transition state.[14] Indeed, in sugar- and side-chain-assisted ligation, the rates are reasonable (6–48 h) despite the reaction proceeding through 14–15-membered ring intermediates for S-N acyl transfer. The involvement of a primary amine in the acyl transfer step, along with to the proximity effect, is a crucial factor.[15]
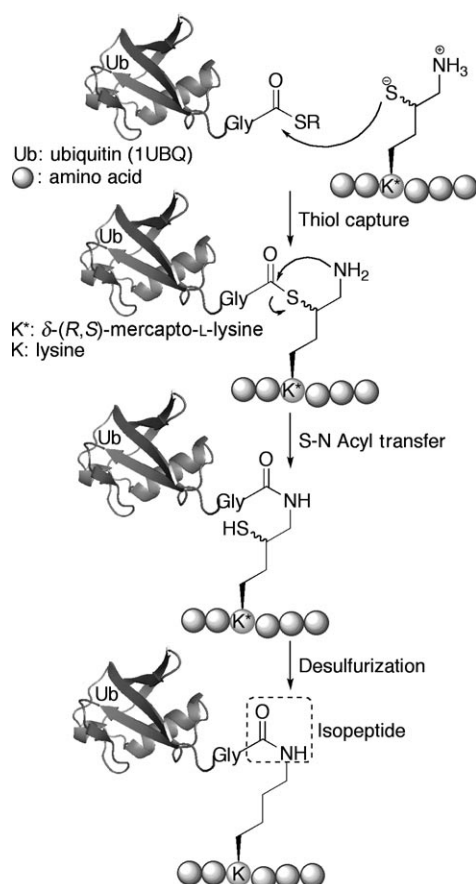
Inspired by NCL, the ligation method developed by Kent and co-workers,[16] we reasoned that the formation of five-membered-ring transition state involving the primary ε-NH$_2$ should have a major influence on the ligation rate. This method would require the installation of a thiol group at the lysine side-chain specifically on the δ-carbon generating a cysteine-like system (Scheme 1). With this in mind, we devised a new thiol-modified lysine analogue, that is, δ-mercaptolysine, to allow a thiol capture step with the Ub-α-thioester, followed by S-N acyl transfer, to form the isopeptide linkage. Moreover, this modification should be removable by applying the desulfurization reaction developed by Dawson and Yan[17] to furnish the unmodified lysine (Scheme 1).

Initially, we focused on the synthesis of the δ-mercaptolysine. We realized that the chiral center on the δ-carbon could be installed in its diastereomeric form, as it will be removed after the desulfurization, which simplifies the synthesis and purification steps. The synthesis of modified lysine started from commercially available L-glutamic acid, which was converted into aldehyde **1** in three steps according to a previously reported procedure (Scheme 2).[18] Subsequently, the Henry reaction was applied to **1** with nitromethane, in the presence of TBAF as a base, to give a diastereomeric mixture of nitro alcohol **2** in a 1:1 ratio (based on ^{1}H NMR spectroscopy; see the Supporting Information). The one-pot acetylation, followed by elimination of the alcohol functionality on the nitro alcohol **2** using acetic anhydride and 4-dimethylaminopyridine, afforded the *E/Z* mixture of the Michael acceptor **3** in 71% yield. To avoid any racemization at the α carbon, the reaction of lithium *tert*-butylsulfide with conjugated nitro olefin **3** was performed at −78°C to furnish a diastereomeric mixture of *tert*-butylmercapto nitro compound **4** in a ratio of 52:48 (see the Supporting Information). Next, the reduction of nitro group to the amine using NaBH$_4$ and NiCl$_2 <$ **M.** > 6 H$_2$O,[19] followed by protection with allyloxycarbonyl chloride, afforded the alloc-protected *tert*-butyl

[*] Dr. K. S. Ajish Kumar, M. Haj-Yahya, Dr. A. Brik
Department of Chemistry, Ben-Gurion University of the Negev
Beer Sheva 84105 (Israel)
Fax: (+972) 8-647-2943
E-mail: abrik@bgu.ac.il
Homepage: http://www.bgu.ac.il/~abrik
Dr. D. Olschewski, Prof. H. A. Lashuel
Brain Mind Institute
Ecole Polytechnique Federale de Lausanne (EPFL)
1015 Lausanne (Switzerland)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200902936.

Scheme 1. Proposed mechanism for peptide ubiquitylation using δ-(R,S)-mercaptolysine.

Scheme 2. Synthesis of δ-(R,S)-mercaptolysine for Boc-SPPS: a) MeNO₂, TBAF, −5 to 0°C, 30 min., 97%; b) Ac₂O, DMAP, ether, 0°C to RT, 4 h, 71%; c) *t*BuSH, BuLi (1.6 M), −10 to −78 0°C, 45 min., 85%; d) 1) NaBH₄, NiCl₂·6H₂O, THF/MeOH (1:1), −20 to −15°C, 20 min., 2) AllocCl, TEA, THF, 0°C to RT, 1 h, 69% for two steps; e) 1) HCl_g, ethylacetate (dry), −20°C, 1 h; 2) (Boc)₂O, Et₃N, MeOH, 0°C to RT, 2 h, 96% over two steps; f) LiOH, THF/H₂O (3:2), 0°C, 1 h, 95%. Boc = *tert*-butoxycarbonyl, TBAF = tetra-*n*-butylammonium fluoride, DMAP = 4-dimethylaminopyridine, Alloc = allyloxy carbonyl, TEA = triethylamine.

mercaptolysine **5** in 69% yield (over two steps). To avoid any racemization at the α carbon during the saponification step, the two Boc protecting groups were removed by purging hydrogen chloride gas through **5** in dry ethyl acetate, and the resultant amine hydrochloride was masked as NH-Boc derivative to give **6** in 96% yield (over two steps). This compound was hydrolyzed using LiOH in THF/H₂O at 0°C to afford the desired δ-(R,S)-mercaptolysine **7** in a quantitative yield.

To verify that the chiral integrity of the α carbon in δ-(R,S)-mercaptolysine was not affected during any step of the synthesis, we carried out the following study. Two peptides based on the sequence FKPSD (PheLysProSerAsp) were prepared bearing the lysine residue in the L- and D- form, generating F-L-LysPSD and F-D-LysPSD, which showed an excellent separation on the reverse-phase HPLC. The same sequence was then used wherein the lysine residue was replaced with the δ-(R,S)-mercaptolysine **7** and was subjected to a desulfurization reaction to generate the FKPSD peptide(s). HPLC analysis, under the same conditions that were used to separate the above diastereoisomer peptides, showed that only the F-L-LysPSD was observed, thus providing evidence that the chiral integrity of the α-carbon in the δ-(R,S)-mercaptolysine **7** was completely retained (see the Supporting Information).

With this amino acid at hand, we then turned our attention to the synthesis of the target peptide bearing the δ-mercaptolysine for site-specific ubiquitylation. For this purpose, we chose the N-terminal segment of α-synuclein(1–17) as a model peptide, driven by our interest to determine the effect of post-translational modifications, such as ubiquitylation, on α-synuclein aggregation and toxicity in Parkinson's disease and related synucleinopathies.[20] Although four lysine residues, at residues 21, 23, 32, and 34 of the α-synuclein N-terminal, were found to be liable to ubiquitylation in vitro, α-syn ubiquitylation in vivo occurs at lysine 6, lysine 10, and lysine 12.[20] The model peptide we selected (MDVFMKGLSKAKEGVVA) contains the desired lysine residues, from which lysine 6 was modified with δ-mercaptolysine and synthesized using Boc-SPPS. Upon completion of the synthesis, the alloc protecting group was removed using [Pd(PPh₃)₄]/TIS (TIS = triisopropylsilane), followed by treatment with trifluoroacetic acid/trifluoromethylsulfonic acid/thioanisol/1,4-dithiothreitol (80:8:8:4) for cleavage from the solid support and side-chain deprotection, to give the desired peptide **8** in about 22% yield of isolated product (see the Supporting Information). A recombinantly expressed HA-tagged Ub(1-76)-α-thioester **9** (HAUb-SR; HA = hemaglutinin) was prepared in 10 mg quantities by treating 50 mg intein-fusion protein with mercaptoethanesulfonic acid as previously described.[8,21]

With all the necessary building blocks at hand, we started our ligation studies between peptide **8** and **9** (HAUb-SR). Thus, HAUb-SR (2 mM) was treated with peptide **8** (2 equiv)

under NCL-like ligation conditions, that is, 6 M Gn·HCl, 200 mM phosphate buffer, pH 7.5 in the presence of 2% v/v thiophenol/benzylmercaptan.[22] The progress of the reaction was followed by analytical HPLC and ESI mass spectrometry. The ligation reaction proceeded rapidly, and a complete conversion of the HAUb-SR **9** was observed within 4 h to give the desired ligation product **10** in 75% yield of isolated product (Figure 1). Under similar conditions, the wild-type peptide of α-synuclein(1–17) failed to show any traces of the ligated product, which supports the proposed ligation mechanism shown in Scheme 1.

The ligation rate that was observed in this study, which is very similar to the rate of Gly-Cys junction formation,[23] is pleasantly surprising considering the differences in the mercaptoamine structures compared to cysteine and isocysteine. The isocysteine case studied by Seitz and Dose,[24] which shows more similarity to our mercaptoamine structure, had a ligation rate at pH 7.5 that is three times slower than for cysteine. Moreover, the calculated pK_a of the mercapto group in δ-mercaptolysine amide is $pK_a = 9.74$, which is higher than the calculated pK_a values of isocysteine amide ($pK_a = 8.2$) and cysteine amide ($pK_a = 9.1$).[25] This fact, along with steric factors of the secondary thiol, could influence the ligation rate, considering that the rate-determining step in these ligations is the thioesterification step. Nonetheless, we believe that the absence of the carbonyl group in our mercaptoamine structure, thus eliminating the inductive effect, combined with the higher flexibility in our system, enhance the nucleophilicity of the thiolate and compensate for the higher pK_a of the mercapto functionality.

Following the ligation step, the ubiquitylated α-syn(1–17) **10** was treated with the metal-free desulfurization conditions of Danishefsky and Wan[26] that were recently used to desulfurize other residues after assisting ligation at a non-cysteine junction.[27] Despite the presence of several methionine residues, we observed a clean, rapid, and full conversion of the starting material to the unmodified isopeptide ligation product **11** (Figure 2). As determined by HPLC and ESI mass spectrometry, the reaction was complete within 3–4 h, and a mass decrease of 32 Da from the starting material was observed, which is consistent with the loss of one sulfur atom (Figure 2). The desulfurized product was isolated in 85% yield and was found to be a good substrate for ubiquitin C-terminal hydrolase, UCH-L3, providing further evidence on the structure integrity of the isopeptide linkage and the folding of the tagged-Ub. In this experiment, the desulfurized product was treated with UCH-L3, as previously reported, to furnish both the hydrolyzed Ub and the α-syn(1–17) (see the Supporting Information). In certain cases, the use of δ-mercaptolysine could provide additional advantages: the desulfurization step could be preformed simultaneously by desulfurizing other specific residues (such as β-mercaptophenylalanine,[28] penicillamine,[27a] and cysteine[17]), which are used to assist ligation at non-cysteine junctions as in the case of the ubiquitylated histone H2B.[11] However, the presence of other cysteine residues could hamper the use of our method, which in principle could be overcome if temporary protection of the native cysteine is performed.
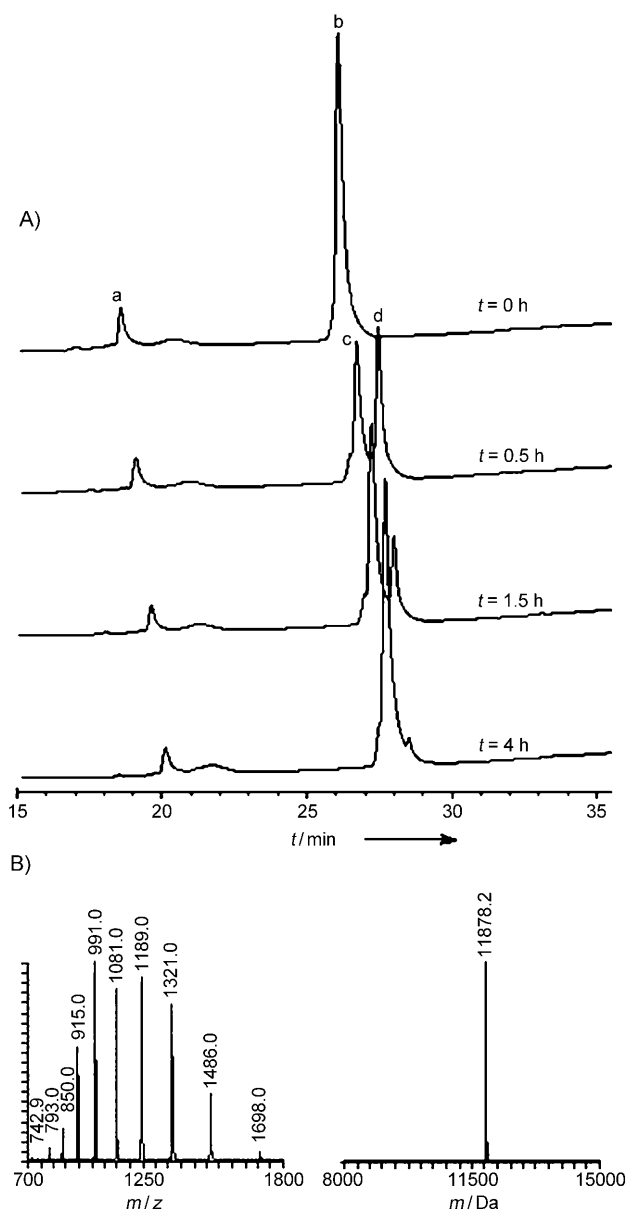

Figure 1. Peptide ubiquitylation: A) Analytical HPLC of the time-course for the ligation reaction of model peptide **8** and HAUb-SR **9**. Peak (a) corresponds to peptide **8** with the observed mass of 1805 Da (calcd 1806 Da); (b), HAUb-SR **9**, obs. 10215 Da (calcd 10216 Da); (c), ligation product **10** (total protein), obs. 11 878.2 Da (calcd 11 878.3 Da); (d), benzylmercaptan thiol exchange intermediate, obs. 10197 Da (calcd 10 198 Da). No HAUb-SR **9** starting material of was detected after 30 min. B) Deconvoluted ESI mass spectrum (right) of the ligation product **10** obtained from the unprocessed charge state spectra (left).

To achieve the synthesis of full-length ubiquitylated proteins, the synthetic strategy will often involve a second ligation step along with ubiquitylation. In this case, protecting-group manipulation of the mercaptolysine or the N-terminal cysteine will be required to achieve site-specific ubiquitylation and backbone ligation. To check the compatibility of our method in systems that require multiple ligations, we tested our strategy with model peptides **12** (Thz-α-syn(2-17); Thz = 1,3-thiazolidine-4-carboxyl) and **13**
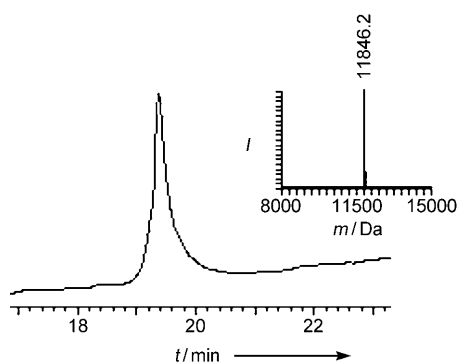
Figure 2. Analytical HPLC of the crude desulfurization reaction and the deconvoluted ESI mass spectrum (inset) of desulfurized product **11** (total protein) with the observed mass of 11 846.2 Da (calcd *m/z* 11 846.3 Da). The progress of the reaction was analyzed using C-4 analytical RP-HPLC employing a gradient of 5% B–60% B over 30 min.

(LYRAG-SR). Model peptide **12** (see the Supporting Information) was designed to include the δ-mercaptolysine and protected N-terminal cysteine residue (Scheme 3). Peptide **12**
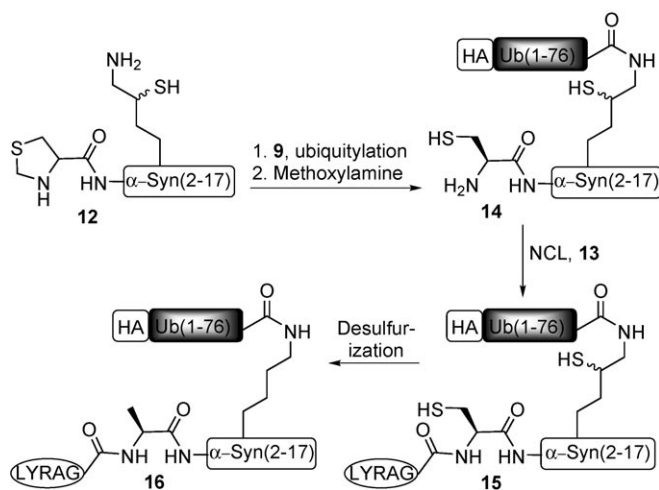
Scheme 3. The use of δ-mercaptolysine in peptide ubiquitylation combined with sequential ligation.

was first ubiquitylated with HAUb-SR **9** to furnish the ubiquitylated product, which was then subjected, without product isolation, to methoxylamine to unmask the N-terminal cysteine, furnishing peptide **14** in 75% yield of isolated product (Figure 3 A).[7] Product **14** was ligated with **13** under NCL conditions to allow for backbone assembly, affording the ligation product **15** (see the Supporting Information). Applying the desulfurization reaction on **15** led to the complete conversion of the cysteine residue to Ala, along with the full removal of the thiol handle from the δ-mercaptolysine residue, to give the desulfurized product **16** in 78% yield of isolated product (Figure 3 B). This model study emphasizes that our strategy could, in principle, be used in sequential ligation for the synthesis of ubiquitylated proteins.
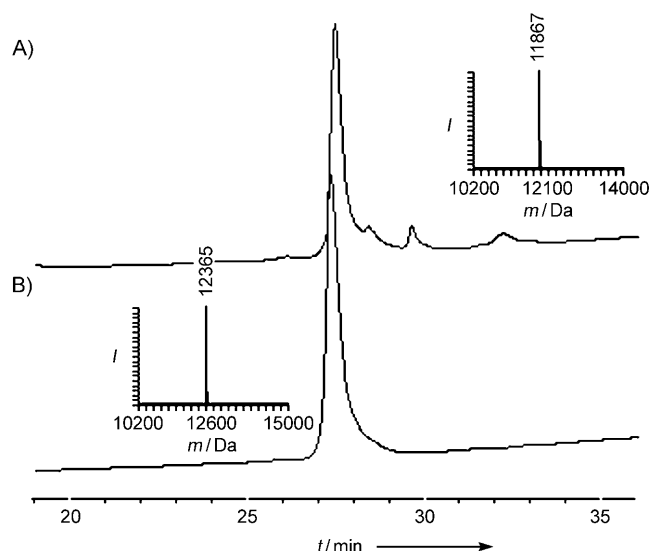

Figure 3. A) Analytical HPLC of the time-course for the ligation reaction of model peptide **12** and **9** followed by N-terminal protecting group removal. Major peak (inset) corresponds to the ligation product **14** (total protein), observed mass 11 867 Da (calcd *m/z* 11 866 Da). B) Analytical HPLC of the crude desulfurization reaction of **15** and the deconvoluted ESI mass spectrum (inset) of desulfurized product **16** (total protein) with the observed mass of 12 365 Da (calcd *m/z* 12 364 Da). The progress of the reaction was analyzed using C-4 analytical RP-HPLC employing a gradient of 5% B–60% B over 30 min.

In summary, we have shown a highly efficient and chemoselective method for ubiquitylation of peptides using the δ-mercaptolysine wherein the thiol functionality serves as a temporary handle in assisting isopeptide formation. Following the ligation step, the thiol moiety at the δ carbon was rapidly converted into the unmodified lysine using the desulfurization reaction. Our strategy, in combination with NCL and EPL, should facilitate the synthesis of homogenous ubiquitylated proteins for biochemical and structural studies. We believe that δ-mercaptolysine could find other uses in peptide and protein chemistry, such as in peptide cyclizations and decorating protein surfaces.[29] We are currently working to apply this approach to study the effect of ubiquitylation and SUMOylation on the function of a wide range of proteins, including α-Syn.

[1] C. T. Walsh, *Posttranslational Modification of Proteins: Expanding Nature's Inventory*, Roberts and Company, Englewood, **2006**.

[2] a) A. Ciechanover, Y. Hod, A. Hershko, *Biochem. Biophys. Res. Commun.* **1978**, *81*, 1100–1105; b) A. Hershko, A. Ciehanover, *Annu. Rev. Biochem.* **1998**, *67*, 425–479.

[3] H. Wang, L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R. S. Jones, Y. Zhang, *Nature* **2004**, *431*, 873–878.

Communications

[4] L. Fallon, C. M. L. Bélanger, A. T. Corera, M. Kontogiannea, E. Regan-Klapisz, F. Moreau, J. Voortman, M. Haber, G. Rouleau, T. Thorarinsdottir, A. Brice, P. M. van Bergen en Henegouwen, E. A. Fon, *Nat. Cell Biol.* **2006**, *8*, 834–842.

[5] C. M. Pickart, *Annu. Rev. Biochem.* **2001**, *70*, 503–533.

[6] R. Ramage, J. Green, T. W. Muir, O. M. Ogunjobi, S. Love, K. Shaw, *Biochem. J.* **1994**, *299*, 151–158.

[7] D. Bang, G. I. Makhatadze, V. Tereshko, A. A. Kossiakoff, S. B. Kent, *Angew. Chem.* **2005**, *117*, 3920–3924; *Angew. Chem. Int. Ed.* **2005**, *44*, 3852–3856.

[8] A. Borodovsky, H. Ovaa, N. Kolli, T. Gan-Erdene, K. D. Wilkinson, H. L. Ploegh, B. M. Kessler, *Chem. Biol.* **2002**, *9*, 1149–1159.

[9] C. M. Pickart, *Methods Enzymol.* **2005**, *399*, 21–36.

[10] C. Chatterjee, R. K. McGinty, J.-P. Pellois, T. W. Muir, *Angew. Chem.* **2007**, *119*, 2872–2876; *Angew. Chem. Int. Ed.* **2007**, *46*, 2814–2818.

[11] R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812–816.

[12] a) T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059–6061; b) C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorg. Med. Chem.* **2004**, *12*, 2749–2757.

[13] C. P. R. Hackenberger, *ChemBioChem* **2007**, *8*, 1221–1223.

[14] D. S. Kemp, *Biopolymers* **1981**, *20*, 1793–1804.

[15] a) A. Brik, Y.-Y. Yang, S. Ficht, C.-H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626–5627; b) A. Brik, C.-H. Wong, *Chem. Eur. J.* **2007**, *13*, 5670–5675; c) M. Y. Lutsky, N. Nepomniaschiy, A. Brik, *Chem. Commun.* **2008**, 1229–1231.

[16] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779; b) P. E. Dawson, A. Dirksen, *Curr. Opin. Chem. Biol.* **2008**, *12*, 760–766; c) C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182–10228; *Angew. Chem. Int. Ed.* **2008**, *47*, 10030–10074.

[17] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533.

[18] J. M. Padrón, G. Kokotos, T. Martín, T. Markidis, W. A. Gibbons, V. S. Martín, *Tetrahedron: Asymmetry* **1998**, *9*, 3381–3394.

[19] a) A. Nose, T. Kudo, *Chem. Pharm. Bull.* **1981**, *29*, 1159–1161.

[20] a) K. Beyer, *Acta Neuropathol.* **2006**, *112*, 237–251; b) T. Nonaka, T. Iwatsubo, M. Hasegawa, *Biochemistry* **2005**, *44*, 361–368.

[21] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705–6710.

[22] P. E. Dawson, M. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.

[23] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.

[24] C. Dose, O. Seitz, *Org. Biomol. Chem.* **2004**, *2*, 59–65.

[25] The p*K*a values were calculated using Sparc: http://ibmlc2.chem.uga.edu/sparc/.

[26] Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408–9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252.

[27] a) C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912–6915; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810; b) J. Chen, Q. Wan, Y. Yuan, J. Zhu, S. J. Danishefsky, *Angew. Chem.* **2008**, *120*, 8649–8652; *Angew. Chem. Int. Ed.* **2008**, *47*, 8521–8524.

[28] D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065.

[29] G. G. Kochendoerfer, S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884–887.