



Chemical Protein Synthesis

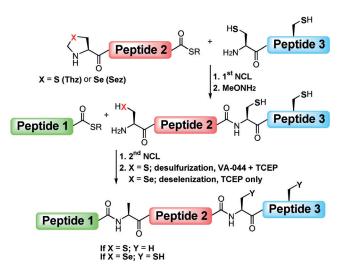
Deutsche Ausgabe: DOI: 10.1002/ange.201509378 Internationale Ausgabe: DOI: 10.1002/anie.201509378

Chemical Synthesis of Proteins with Non-Strategically Placed Cysteines Using Selenazolidine and Selective Deselenization

Post Sai Reddy, Shahar Dery, and Norman Metanis*

Abstract: Although native chemical ligation has enabled the synthesis of hundreds of proteins, not all proteins are accessible through typical ligation conditions. The challenging protein, 125-residue human phosphohistidine phosphatase 1 (PHPT1), has three cysteines near the C-terminus, which are not strategically placed for ligation. Herein, we report the first sequential native chemical ligation/deselenization reaction. PHPT1 was prepared from three unprotected peptide segments using two ligation reactions at cysteine and alanine junctions. Selenazolidine was utilized as a masked precursor for Nterminal selenocysteine in the middle segment, and, following ligation, deselenization provided the native alanine residue. This approach was used to synthesize both the wild-type PHPT1 and an analogue in which the active-site histidine was substituted with the unnatural and isosteric amino acid β thienyl-L-alanine. The activity of both proteins was studied and compared, providing insights into the enzyme active site.

Chemical protein synthesis (CPS)^[1] offers chemists complete control over proteins at the atomic level. Using solidphase peptide synthesis (SPPS) and selective chemical ligation of unprotected peptides in aqueous conditions,[1] CPS affords access to proteins with unnatural amino acids, [2,3] site-specific post-translational modifications [4] as well as mirror-image proteins.^[5,6] The most widely used chemical ligation method is the native chemical ligation (NCL) reaction^[7] between a C-terminal thioester of one peptide and an N-terminal cysteine (Cys, C) of another. Chemical synthesis of moderate-sized proteins (ca. 150 aa) often requires more than a single NCL reaction (Scheme 1), and works best when Cys residues are located strategically within the protein sequence (ca. 40 residues apart). In such a sequential synthetic scheme, the middle peptide segment must contain both an N-terminal Cys and a C-terminal thioester, which prompted the use of several protecting groups for N-terminal Cys to prevent cyclization. [3,8,9] However, these protecting groups suffer from cumbersome deprotection procedures and lower yields, and became less utilized after the development of thiazolidine (Thz).[10] Currently, Thz has been adapted for the synthesis of many proteins owing to its smooth deprotection procedure and high efficiency.[10]



Scheme 1. Utilizing Thz/Sez together with desulfurization/deselenization for chemical protein synthesis.

Perhaps the only limitation of NCL is its dependence on a Cys residue at the ligation junction. As Cys is a rare amino acid in proteins, many research groups aimed to expand chemical ligation technology to overcome this limitation by developing new ligation reactions, [11,12] using removable auxiliaries, [13-15] or utilizing selective post-ligation modifications. [16-24] One example of the latter is the NCL/desulfurization approach developed by Yan and Dawson, [16] in which NCL is expanded to an alanine (Ala, A) residue at the ligation junction (or other amino acids when a thiolated side-chain is installed). [16-24] A key weakness in this approach is its lack of selectivity during the desulfurization step (Scheme 1) [16,25] which leads to global desulfurization of unprotected native Cys groups. [23,26,27]

An alternative approach to NCL/desulfurization is the NCL/deselenization technique. [28] Here, NCL is performed at selenocysteine (Sec, U;[29-31] or a selenolated side-chain of any residue)[23,27,32,33] followed by a selective deselenization reaction, which converts the Sec at the ligation site into Ala in the presence of unprotected Cys residues (Scheme 1). Recently, through a detailed mechanistic study of the proposed radical deselenization reaction, we found that Sec can also be selectively converted into Ser if the deselenization reaction is performed under oxygen-saturated conditions.^[34] This result was recently supported by Payne and co-workers.[35] These reactions open up new directions for the synthesis of proteins, especially when non-strategically positioned Cys residues are present in a protein sequence. Alternatively, Sec can be retained either to facilitate oxidative protein folding,[36-40] or to be used as a handle for site-selective modifi-

^[*] Dr. P. S. Reddy, S. Dery, Dr. N. Metanis The Institute of Chemistry The Hebrew University of Jerusalem Edmond J. Safra, Givat Ram, Jerusalem 91904 (Israel) E-mail: Metanis@mail.huji.ac.il

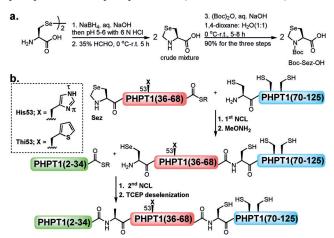
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201509378.





In light of these clear advantages, we explored the utility of the selenazolidine (Sez) building block for the synthesis of challenging proteins. Key elements to our approach were the facile and smooth conversion of the Sez-peptide to Secpeptide using $MeONH_2$ and selective deselenization after NCL.

As a proof-of-concept, we chose the human enzyme phosphohistidine phosphatase 1 (PHPT1; Scheme 2). Histi-



Scheme 2. a. Chemical synthesis of Boc-Sez-OH. b. Utilizing Sez for the chemical synthesis of PHPT1 by ligation of three unprotected peptide segments. Residue His 53 in the wildtype PHPT1 active site was also replaced with the unnatural amino acid, Thi.

dine phosphorylation (pHis) is a poorly studied post-translational modification (PTM), even though it has been known since the $1960s^{[42]}$ and accounts for $\approx\!6\,\%$ of all phosphorylation PTMs. $^{[43,44]}$ PHPT1, discovered in 2002, has a proposed neuronal function as phosphatase acting on pHis. $^{[45]}$ The relatively late discovery is attributed to the lability of the P–N bond, which makes this PTM challenging to study, although recent advances in proteomics, $^{[46]}$ synthesis of stable pHis analogues, $^{[44]}$ and mAb development $^{[47]}$ have been reported recently.

PHPT1 contains 125 amino acids and three non-strategically positioned Cys residues (Cys 69, 71, and 73), and cannot be easily prepared through typical Cys-NCL techniques. Mutagenesis, X-ray, and NMR studies [48,49] suggest that the active site histidine (His 53) catalyzes the activation of a water molecule, which attacks the phosphate group of the pHis substrate. We utilized Sez for the synthesis of the wild-type PHPT1 and His53Thi analogue, in which the active site His53 is substituted with the isosteric, unnatural amino acid β -thienyl-L-alanine (Thi). [3] Our results provide further insight into the critical role of His53 in the catalytic activity of PHPT1.

The synthesis of Boc-protected selenazolidine carboxylic acid (Boc-Sez-OH) is summarized in Scheme 2a. Starting from selenocystine (see the Supporting Information), it was reduced in aqueous NaBH₄ and then treated with formaldehyde. [50] The crude mixture was reacted with (Boc)₂O, which provided the desired Boc-Sez-OH in 90% overall yield.

We first checked the enantiopurity of the Sez building block and its sensitivity to racemization during standard Fmoc synthesis, cleavage, and purification steps using three synthetic hexapeptides (X-LKFAG; X = Sez (1), L-Ala (2), and D-Ala (3); Supporting Information). After cleavage, peptide 1 was treated with MeONH₂ at pH 4, which converted Sez-LKFAG (1) into Sec-LKFAG (1a; isolated as a dimer). Deselenization by TCEP produced a product that co-eluted with 2 (L-Ala) with no observation of the diastereomer 3 (D-Ala). Since the deselenization reaction is enantioselective, [28] this result suggests that Boc-Sez-OH was enantiopure and the chiral center was retained during the synthesis, cleavage, and handling steps (Supporting Information, Figures S2–S4). Next, we compared the ring opening of Sez to that of Thz by synthesizing Thz-LKFAG (4) and treating peptides 1 and 4 with a MeONH₂ solution at pH 4, as previously described. [10] The reaction was followed by HPLC to show a slightly faster Thz opening (2 h) than Sez opening (6 h; Figures S5–S7).

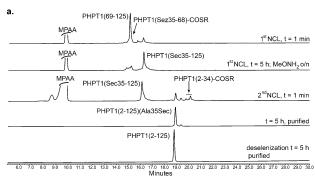
We then demonstrated the usefulness of the Sez building block in the chemical synthesis of PHPT1. Our synthetic route was based on three segments with two sequential NCL reactions at Gly68–Cys69 and Gly34–Ala35, where Ala35 was substituted with Sez, followed by a single deselenization step (Scheme 2b). All peptide syntheses were performed using standard stepwise Fmoc-SPPS (Supporting Information). PHPT1(Sez35–Gly68) and PHPT1(Ala2–Gly34) segments were synthesized with a C-terminal thioester precursor, Dbz^[51] (Supporting Information).

The ligation between PHPT1(Sez35–68)-COSR and PHPT1(Cys69–125) was performed under traditional conditions and was completed in 5 h (Figure 1a). The crude mixture was then treated overnight with 0.2 m MeONH₂ at pH 4–5 affording PHPT1(Sec35–125) (Figure S12) in 51 % (5.8 mg) yield. Prior to purification, the solution was treated with either dithiothreitol (DTT, after adjusting to pH 7), or a mixture of 1:3 tris(2-carboxyethyl)phosphine (TCEP) and sodium ascorbate (pH 4) to reduce any Se–S bonds and simplify product isolation.

Slow ligations are usually observed at Sec-NCL due to the sensitivity of the selenol groups to air oxidation. [28,31] Previously, aryl thiols, such as 4-mercaptophenylacetic acid (MPAA), were found both to act as a suitable reductant to the oxidized Sec, keeping it in the reactive reduced selenol form, and to activate the thioester. However, MPAA is oxidized during the reaction, which can slow the ligation, and may form MPAA-adducts of product isomers. Reducing agents, such as TCEP, can be advantageous in keeping MPAA reduced, but may deselenize the Sec as a side reaction.^[31] Recently, we were excited to find that the radical quencher, sodium ascorbate, completely inhibits the undesired deselenization reaction, even in the presence of excess TCEP.[34] Therefore, the second ligation between PHPT1(2– 34)-COSR and PHPT1(Sec35-125) was performed in ligation buffer containing TCEP and sodium ascorbate (Supporting Information), and was completed within 5 h (Figure 1a) with $\approx 60\%$ (2.8 mg) yield. Prior to deselenization, the ligated product was either purified by preparative HPLC or desalted using a PD-10 column. The deselenization reaction [28,34] was performed using TCEP (42 equiv) in the presence of DTT (4.9 equiv) and was completed in 5 h. The product was then purified by preparative HPLC (Figure 1a; Supporting Information, Figure S13) to give $\approx 70\%$ yield (1.4 mg), and







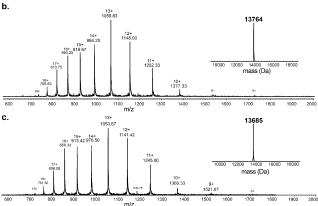


Figure 1. Preparation of PHPT1. a. Analytical HPLC of NCL and deselenization reactions. b. ESI-MS of PHPT1(2–125) (Ala35Sec) (obs. 13764 Da, calc. 13762 Da). c. ESI-MS of PHPT1(2–125) (obs. 13685 Da, calc. 13683 Da). The mass difference between (b) and (c) is 79 Da, indicating deselenization.

characterized by ESI-MS (Figure 1b,c) and HR-MS (Figure S14). The same synthetic approach was used for the synthesis of the PHPT1(His53Thi) analogue (Figure S15–S17), using PHPT1(Sez 35–68)(His53Thi)-COSR (Figure S10) as the middle segment.

To investigate the proposed role of His as a general base catalytic residue in the enzyme, [45,48,49] the two synthetic proteins, PHPT1 and PHPT1(His53Thi), were separately dissolved in Tris buffer in the presence of DTT for one hour to fold, upon which the solution was taken for further studies. First, the proteins were characterized by circular dichroism (CD, Figure S18). The CD spectra of PHPT1 and PHPT1-(His53Thi) (Figure S18) suggest they are highly composed of β-sheet structures. [48,49] Nevertheless, the His53Thi analogue showed small changes in comparison to PHPT1. This may suggests that the imidazole ring of His 53 is not only important for catalysis, but also for the correct folding of the active site through, for example, hydrogen bonding (see below). This may support the previous observation with PHPT1-(His53Ala). [48,49]

Next, we tested the enzymatic activity of the two proteins using p-nitrophehylphosphate (pNPP) as a substrate. ^[49] We found that our synthetic PHPT1 was active with comparable steady-state kinetic parameters ($K_{\rm M}$, $k_{\rm cat}$, and $k_{\rm cat}/K_{\rm M}$; Figure S19) to the published values (Table 1). On the other hand, PHPT1(His53Thi) activity was below the detection limit, that is, at least 3600-fold lower than PHPT1. These results support

Table 1: Steady-state kinetic parameters for human PHPT1 proteins using pNPP as substrate.

Protein	$K_{\rm M}$ [тм]	$k_{\rm cat} [s^{-1}]$	$k_{\rm cat}/K_{\rm M} \ [{\rm M}^{-1}{\rm s}^{-1}]$
Wt PHPT1 ^[a]	7.4 ± 0.1	$\textbf{0.35} \pm \textbf{0.02}$	47±3
PHPT1	6.1 ± 0.4	$\textbf{0.22} \pm \textbf{0.01}$	36 ± 2
PHPT1 (His53Thi) ^[b]	n.d.	n.d.	< 0.01

[a] Published values. [49] [b] The enzymatic activity was at least 3600-fold lower than PHPT1.

the crucial catalytic role of His53 in PHPT1, such that when the basic imidazole $N\pi$ is replaced with S, the catalytic activity of the enzyme is severely diminished. Additionally, His 53 may play an important role in the pre-organization of the active site through, for example, hydrogen bonding. Indeed, careful examination of the active site of PHPT1 (based on NMR structure of the Pi-bound form)^[49] revealed that $N\tau$ -H of His 53 forms a hydrogen bond (1.75-1.81 Å in 14 out of the 21 NMR structures) with the carbonyl oxygen of Gly77 (found in the Gly-rich domain), which is in the β4-strand underneath His 53 (Figure S20). Gly 77 is a conserved residue among PHPTs^[52] (Figure S21), and it has been reported that a Gly77Ala mutant is inactive.^[52] These results support the critical role of Gly77 in the active site, which, through hydrogen bonding with His 53, ensures that the imidazole ring is in the right conformation to activate the catalytic water molecule. In the His53Thi analogue, the imidazole ring is replaced with thienyl, and as a result, the Nτ-H is replaced with a methylene, resulting in the elimination of this critical hydrogen bond (Figure S22). This might also contribute to the observed small changes in the CD and diminished activity. Future studies in which His 53 is replaced with more functional His analogues such as 1,2,3-triazole-3-alanine, or 2- or 4-fluorohistidine are currently underway.

In summary, we performed the first chemical synthesis using selenazolidine and deselenization to access PHPT1, a protein with non-strategically placed cysteine residues. Both wild-type PHPT1 and an analogue in which the active-site histidine (His) was substituted with the unnatural and isosteric amino acid β-thienyl-L-alanine (Thi) were prepared from two ligations of three peptide segments, followed by a single, selective deselenization reaction. Through CD and enzymatic activity studies of the two proteins and the examination of the published NMR structures, we propose that the carbonyl oxygen of Gly77 forms a hydrogen bond with the imidazole ring of His 53, ensuring the imidazole ring is in the right conformation to activate a catalytic water molecule. As a result of the elimination of the hydrogen bond with Gly 77 (Nτ-H is replaced with C-H), and/or the replacement of the basic imidazole $N\pi$ with S, PHPT1(His53Thi) was found to be inactive.

The synthetic work reported here may serve as a useful methodology for the synthesis of more complex protein targets. In principle, a similar approach could be used to assemble larger proteins from more than three unprotected peptide segments. Such studies are currently underway in our research group.





Acknowledgements

We thank Dr. Dana Reichmann for helping with HR-MS data collection, Reem Mousa and Linoy Dery for peptide purifications, and Rebecca Notis Dardashti for input on the manuscript. N.M. acknowledges the Israel Science Foundation and the German–Israeli Foundation for Scientific Research and Development (GIF). P.S.R. is supported by the PBC fellowship, Council for Higher Education, Israel.

Keywords: native chemical ligation \cdot phosphohistidine \cdot phosphohistidine phosphatase 1 \cdot post-translational modifications \cdot selenocysteine

How to cite: Angew. Chem. Int. Ed. **2016**, 55, 992–995 Angew. Chem. **2016**, 128, 1004–1007

- [1] S. B. H. Kent, Chem. Soc. Rev. 2009, 38, 338-351.
- [2] N. Metanis, A. Brik, P. E. Dawson, E. Keinan, J. Am. Chem. Soc. 2004, 126, 12726–12727.
- [3] T. M. Hackeng, J. H. Griffin, P. E. Dawson, Proc. Natl. Acad. Sci. USA 1999, 96, 10068–10073.
- [4] K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, Angew. Chem. Int. Ed. 2011, 50, 6137–6141; Angew. Chem. 2011, 123, 6261–6265.
- [5] R. C. D. Milton, S. C. F. Milton, S. B. H. Kent, Science 1992, 256, 1445–1448.
- [6] K. Mandal, M. Uppalapati, D. Ault-Riche, J. Kenney, J. Lowitz, S. S. Sidhu, S. B. H. Kent, *Proc. Natl. Acad. Sci. USA* 2012, 109, 14779–14784.
- [7] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, Science 1994, 266, 776 – 779.
- [8] D. Bang, N. Chopra, S. B. H. Kent, J. Am. Chem. Soc. 2004, 126, 1377 – 1383.
- [9] E. Boll, J. P. Ebran, H. Drobecq, O. El-Mahdi, L. Raibaut, N. Ollivier, O. Melnyk, *Org. Lett.* 2015, 17, 130–133.
- [10] D. Bang, S. B. H. Kent, Angew. Chem. Int. Ed. 2004, 43, 2534–2538; Angew. Chem. 2004, 116, 2588–2592.
- [11] J. W. Bode, R. M. Fox, K. D. Baucom, Angew. Chem. Int. Ed. 2006, 45, 1248–1252; Angew. Chem. 2006, 118, 1270–1274.
- [12] B. L. Nilsson, L. L. Kiessling, R. T. Raines, Org. Lett. 2000, 2, 1939–1941.
- [13] J. Offer, C. N. C. Boddy, P. E. Dawson, J. Am. Chem. Soc. 2002, 124, 4642–4646.
- [14] P. Botti, M. R. Carrasco, S. B. H. Kent, Tetrahedron Lett. 2001, 42, 1831–1833.
- [15] D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, Proc. Natl. Acad. Sci. USA 2001, 98, 6554–6559.
- [16] L. Z. Yan, P. E. Dawson, J. Am. Chem. Soc. 2001, 123, 526-533.
- [17] D. Crich, A. Banerjee, J. Am. Chem. Soc. 2007, 129, 10064– 10065.
- [18] M. Y. Lutsky, N. Nepomniaschiy, A. Brik, Chem. Commun. 2008, 1229–1231.
- [19] J. Chen, Q. Wan, Y. Yuan, J. L. Zhu, S. J. Danishefsky, Angew. Chem. Int. Ed. 2008, 47, 8521–8524; Angew. Chem. 2008, 120, 8649–8652.
- [20] C. Haase, H. Rohde, O. Seitz, Angew. Chem. Int. Ed. 2008, 47, 6807–6810; Angew. Chem. 2008, 120, 6912–6915.
- [21] Z. Harpaz, P. Siman, K. S. A. Kumar, A. Brik, *ChemBioChem* 2010, 11, 1232–1235.
- [22] S. Y. Shang, Z. P. Tan, S. W. Dong, S. J. Danishefsky, J. Am. Chem. Soc. 2011, 133, 10784–10786.
- [23] A selective desulfurization of β-mercapto-Asp has been reproted. R. E. Thompson, B. Chan, L. Radom, K. A. Jolliffe,

- R. J. Payne, *Angew. Chem. Int. Ed.* **2013**, *52*, 9723 9727; *Angew. Chem.* **2013**, *125*, 9905 9909.
- [24] P. E. Dawson, Isr. J. Chem. 2011, 51, 862-867.
- [25] Q. Wan, S. J. Danishefsky, Angew. Chem. Int. Ed. 2007, 46, 9248–9252; Angew. Chem. 2007, 119, 9408–9412.
- [26] B. L. Pentelute, S. B. H. Kent, Org. Lett. 2007, 9, 687-690.
- [27] Y.-Y. Yang, S. Ficht, A. Brik, C. H. Wong, J. Am. Chem. Soc. 2007, 129, 7690 – 7701.
- [28] N. Metanis, E. Keinan, P. E. Dawson, Angew. Chem. Int. Ed. 2010, 49, 7049 – 7053; Angew. Chem. 2010, 122, 7203 – 7207.
- [29] R. J. Hondal, B. L. Nilsson, R. T. Raines, J. Am. Chem. Soc. 2001, 123, 5140 – 5141.
- [30] R. Quaderer, A. Sewing, D. Hilvert, Helv. Chim. Acta 2001, 84, 1197–1206.
- [31] M. D. Gieselman, L. L. Xie, W. A. van der Donk, *Org. Lett.* 2001, 3, 1331 – 1334.
- [32] L. R. Malins, R. J. Payne, Org. Lett. 2012, 14, 3142-3145.
- [33] S. D. Townsend, Z. P. Tan, S. W. Dong, S. Y. Shang, J. A. Brailsford, S. J. Danishefsky, J. Am. Chem. Soc. 2012, 134, 3912–3916.
- [34] S. Dery, P. S. Reddy, L. Dery, R. Mousa, R. Notis Dardashti, N. Metanis, *Chem. Sci.* 2015, 6, 6207 6212.
- [35] L. R. Malins, N. J. Mitchell, S. McGowan, R. J. Payne, Angew. Chem. Int. Ed. 2015, 54, 12716–12721; Angew. Chem. 2015, 127, 12907–12912.
- [36] A. Dantas de Araujo, B. Callaghan, S. T. Nevin, N. L. Daly, D. J. Craik, M. Moretta, G. Hopping, M. J. Christie, D. J. Adams, P. F. Alewood, *Angew. Chem. Int. Ed.* 2011, 50, 6527–6529; *Angew. Chem.* 2011, 123, 6657–6659.
- [37] N. Metanis, D. Hilvert, Angew. Chem. Int. Ed. 2012, 51, 5585–5588; Angew. Chem. 2012, 124, 5683–5686.
- [38] N. Metanis, D. Hilvert, Curr. Opin. Chem. Biol. 2014, 22, 27-34.
- [39] N. Metanis, D. Hilvert, Chem. Sci. 2015, 6, 322-325.
- [40] A. M. Steiner, K. J. Woycechowsky, B. M. Olivera, G. Bulaj, Angew. Chem. Int. Ed. 2012, 51, 5580-5584; Angew. Chem. 2012, 124, 5678-5682.
- [41] D. T. Cohen, C. Zhang, B. L. Pentelute, S. L. Buchwald, J. Am. Chem. Soc. 2015, 137, 9784–9787.
- [42] P. D. Boyer, M. Deluca, K. E. Ebner, D. E. Hultquist, J. B. Peter, J. Biol. Chem. 1962, 237, PC3306 – PC3308.
- [43] H. R. Matthews, Pharmacol. Ther. 1995, 67, 323-350.
- [44] J. M. Kee, T. W. Muir, ACS Chem. Biol. **2012**, 7, 44–51.
- [45] P. Ek, G. Pettersson, B. Ek, F. Gong, J. P. Li, O. Zetterqvist, Eur. J. Biochem. 2002, 269, 5016–5023.
- [46] R. C. Oslund, J. M. Kee, A. D. Couvillon, V. N. Bhatia, D. H. Perlman, T. W. Muir, J. Am. Chem. Soc. 2014, 136, 12899 – 12911.
- [47] S. R. Fuhs, J. Meisenhelder, A. Aslanian, L. Ma, A. Zagorska, M. Stankova, A. Binnie, F. Al-Obeidi, J. Mauger, G. Lemke, J. R. Yates, T. Hunter, *Cell* 2015, 162, 198–210.
- [48] R. D. Busam, A. G. Thorsell, A. Flores, M. Hammarstrom, C. Persson, B. M. Hallberg, J. Biol. Chem. 2006, 281, 33830 – 33834.
- [49] W. B. Gong, Y. F. Li, G. F. Cui, J. C. Hu, H. M. Fang, C. W. Jin, B. Xia, Biochem. J. 2009, 418, 337 344.
- [50] M. D. Short, Y. Xie, L. Li, P. B. Cassidy, J. C. Roberts, J. Med. Chem. 2003, 46, 3308 – 3313.
- [51] J. B. Blanco-Canosa, P. E. Dawson, Angew. Chem. Int. Ed. 2008, 47, 6851–6855; Angew. Chem. 2008, 120, 6957–6961.
- [52] S. Klumpp, N. T. Ma, N. Baumer, G. Bechmann, J. Krieglstein, Biochim. Biophys. Acta Proteins Proteomics 2010, 1804, 206– 211.

Received: October 7, 2015 Revised: November 2, 2015 Published online: December 7, 2015

1007