Chemical Protein Synthesis

DOI: 10.1002/ange.201105837

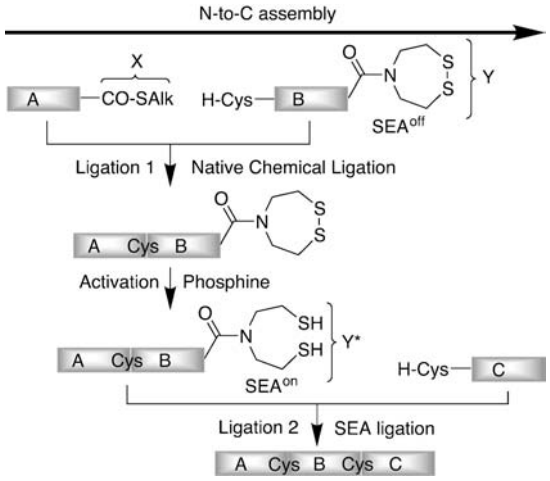
A One-Pot Three-Segment Ligation Strategy for Protein Chemical Synthesis**

*Nathalie Ollivier, Jérôme Vicogne, Aurélie Vallin, Hervé Drobecq, Rémi Desmet, Ouafâa El Mahdi, Bérénice Leclercq, Gautier Goormachtigh, Véronique Fafeur, and Oleg Melnyk**

Protein chemical synthesis by native peptide ligation of unprotected peptide segments is an interesting complement and potential alternative to the use of living systems for producing proteins.[1] Actually, tremendous efforts are focused on the design of one-pot strategies allowing the assembly of three peptide segments.[2–5] The goal is to get rapid access to small proteins (less than 150 amino acid residues), while saving intermediate purification steps and obtaining the products in good yield. Such methods are gaining increasing significance for the study of protein function and appear as a potential option for producing various protein-based therapeutics currently under development.

To date, proteins were mainly assembled by sequential native chemical ligation (NCL)[6] or extended methodologies[7] in the C-to-N direction (for recent achievements, see Refs. [8,9]). NCL involves the chemoselective ligation of a C-terminal peptide thioester, usually an alkylthioester, with an N-terminal cysteine (Cys) peptide. The one-pot sequential C-to-N ligation of three peptide segments designed by Kent et al.[3] is increasingly used for synthesizing proteins.[9,10]

Methods that enable the assembly of peptide segments in the reverse N-to-C direction are rare.[2,5,11] Fundamentally, the combination of N-to-C and C-to-N assembly techniques is at the basis of the convergent total synthesis of proteins.[2] The general principle of the one-pot assembly of three peptide segments in the N-to-C direction is illustrated in Scheme 1. Ligation of peptide segments A-X and H-Cys-B-Y yields segment A-Cys-B-Y (Scheme 1, ligation 1). Group Y must ideally be inert during ligation 1 or at least be significantly less reactive than group X to avoid oligomerization or cyclization of segment B. Activation of group Y into Y* subsequently allows the ligation with the third segment H-Cys-C (Scheme 1, ligation 2). For designing a one-pot process working in the N-to-C direction, this activation must be carried out in situ after ligation 1 by using reagents compatible with ligation 2. Furthermore, the Y* group must enable an efficient ligation with the Cys segment C.



Scheme 1. Total protein synthesis by one-pot assembly of three peptide segments in the N-to-C direction. The first step is a native chemical ligation between thioester segment A and Cys segment B, during which the cyclic disulfide SEA^{off} acts as a blocked thioester group (SEA = bis(2-sulfanylethyl)amido). Activation of SEA^{off} into SEA^{on} by reduction with a phosphine and addition of the third Cys segment C triggers the second ligation step.

To date only few one-pot strategies have been described that work in the N-to-C direction and enable the coupling of three peptide segments.[2,5,12] Fundamentally, these methods, such as kinetically controlled ligation,[2] rely on the differential reactivity of X and Y groups for peptide-bond formation. In other words, the purity of the target polypeptide is highly dependent on the C-terminal residues of A and B segments and more generally on the accessibility of the reactive ends.

Clearly, a strategy in which Y is inert during the first ligation step would bypass these limitations and constitute a critical advance. Herein we show that the combination of NCL and SEA ligation (Scheme 1) permitted design of a solution to this important problem.

Reaction of a peptide featuring a C-terminal bis(2-sulfanylethyl)amido group, called SEA^{on} hereafter (Scheme 1), with a Cys peptide results in the formation of a native peptide bond in water at pH 7.[13] This reaction probably proceeds via

[*] N. Ollivier, Dr. J. Vicogne, A. Vallin, H. Drobecq, R. Desmet, B. Leclercq, Dr. G. Goormachtigh, Dr. V. Fafeur, Dr. O. Melnyk
CNRS UMR 8161, Université Lille Nord de France
Institut Pasteur de Lille
59021 Lille (France)
E-mail: oleg.melnyk@ibl.fr
O. El Mahdi
Université Sidi Mohamed Ben Abdellah (Morocco)

[**] This work was supported by PICS 4862 (CNRS, CNRST), Institut Pasteur de Lille, Université Lille Nord de France, Cancéropôle Nord Ouest, Région Nord pas de Calais, and European Community.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201105837.

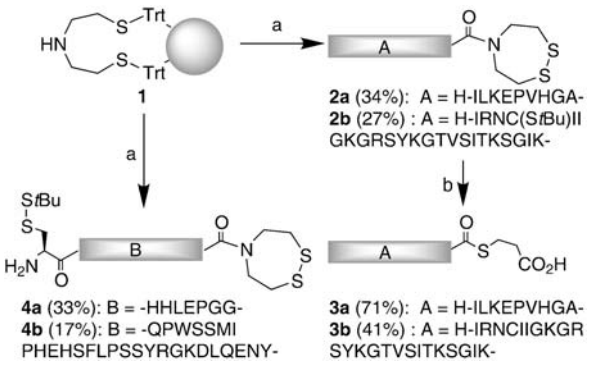
© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

a transient thioester intermediate obtained by intramolecular attack of one SEA^{on} thiol on the C-terminal carbonyl group of the peptide. SEA^{off} designates the corresponding cyclic disulfide derivative, which is unable to ligate, because both thiol groups are blocked within the disulfide bond.

Our study began with the observation that the SEA^{off} group was stable at pH 7 in the presence of (4-carboxymethyl)thiophenol (MPAA).[14] NCL with C-terminal peptide alkylthioesters is usually carried out in the presence of an aromatic thiol such as thiophenol or MPAA.[14,15] The aromatic thiol plays a critical role, because it maintains Cys thiols in a reduced form, activates the alkylthioester by thiol–thioester exchange, and allows the reversal of unproductive thioesters formed by thiol–thioester exchange with internal cysteine residues.

The stability of the SEA^{off} group under typical NCL experimental conditions opened up the possibility to design a novel sequential N-to-C ligation strategy based on NCL and SEA ligation (Scheme 1, X = CO-SAlk, Y = SEA^{off}, Y* = SEA^{on}). The internal segment H-Cys-B-SEA^{off} is a key element in this strategy. It features a free Cys residue on the N terminus and a SEA^{off} group on the C terminus. The first step (Scheme 1, ligation 1) is an MPAA-catalyzed NCL between the alkylthioester group[16] of segment A and the Cys residue of the H-Cys-B-SEA^{off} segment, leading to the formation of A-Cys-B-SEA^{off}. SEA^{off} is then switched on by reduction to allow the second ligation step with Cys segment C. Importantly, this strategy enables the two ligations to be performed in one pot owing to the compatibility of MPAA with SEA ligation conditions and to the possibility of activating the SEA^{off} group in situ with tris(2-carboxyethyl)-phosphine (TCEP) after completion of the first ligation step. Note that this chemistry can be used for assembling more than three peptide segments. For example, A-Cys-B-SEA^{off} can be isolated after ligation 1, then converted into C-terminal peptide alkylthioester A-Cys-B-CO-SAlk.[17] This peptide thioester can be used subsequently as starting material in the one-pot process with H-Cys-C-SEA^{off} and H-Cys-D for assembling four peptide segments. It should be pointed out that the SEA$^{on/off}$ system constitutes the first thioester equivalent that can be locked or unlocked depending on the reducing potential of the ligation mixture.

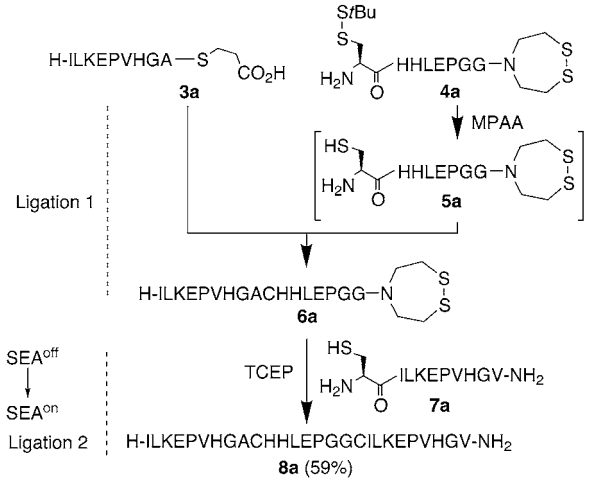
The syntheses of thioester segments A (**3**) and of SEA^{off} segments B (**4**) are shown in Scheme 2. Both segments were synthesized by using Fmoc-based solid-phase peptide synthesis (SPPS)[18] on the bis(2-sulfanylethyl)amino solid support **1**.[13] When present, Cys residues were incorporated as *tert*-butylsulfenyl-protected derivatives during the peptide elongation step to facilitate the subsequent oxidative cyclization of SEA^{on} into SEA^{off}. Deprotection and cleavage from the resin in trifluoroacetic acid furnished the crude SEA^{on} peptides with a *tert*-butylsulfenyl group still attached to Cys-side-chain thiols. SEA^{on} was switched off with iodine. Cys residues protected with *tert*-butylsulfenyl groups remained unaffected during this step, which was carried out in a few seconds prior to the HPLC purification step.[19] This procedure furnished the SEA^{off} peptides **2** and **4** in a straightforward manner.



Scheme 2. Bis(2-sulfanylethyl)amino resin **1** enables the Fmoc-SPPS synthesis of thioester and SEA^{off} peptide segments. a) 1. Fmoc/*t*Bu SPPS; 2. I_2 in water/AcOH. b) 3-Mercaptopropionic acid (5 vol %), TCEP, pH 4, 37 °C. Fmoc = 9-fluorenylmethoxycarbonyl, Trt = triphenylmethyl, I = isoleucine, L = leucine, K = lysine, E = glutamic acid, P = proline, V = valine, H = histidine, G = glycine, A = alanine, R = arginine, N = asparagine, C = cysteine, S = serine, Y = tyrosine, T = threonine, Q = glutamine, W = tryptophan, M = methionine, F = phenylalanine, D = aspartic acid.

Synthesis of thioester segments **3** required further exchange of the SEA^{off} group by 3-mercaptopropionic acid. For this step, SEA^{off} segments **2** were treated with 3-mercaptopropionic acid and TCEP at pH 4, as described elsewhere.[17] Formation of the C-terminal thioester moiety was accompanied by the concomitant deprotection of the Cys residue, when present, to give thioester segments **3** in good yield.

In a preliminary approach, we undertook a proof-of-concept study with model thioester peptide **3a** (segment A), SEA^{off} peptide **4a** (segment B), and Cys peptide amide **7a** (segment C, Scheme 3). To establish the stability of the SEA^{off} group in the mixture used for ligation 1, peptide **4a** was dissolved at 37 °C in a pH 7.5 phosphate buffer containing MPAA (200 mM). LC–MS analysis of the mixture (see the Supporting Information) showed the formation of depro-



Scheme 3. Model one-pot sequential NCL/SEA ligation sequence using peptide segments **3a**, **4a**, and **7a**. Ligation 1: **3a**, **4a**, MPAA 200 mM, Gdn·HCl 4 M, pH 7, 37 °C. SEA^{off}→SEA^{on} and ligation 2: **7a**, TCEP 200 mM, 37 °C. Gdn = guanidinium.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tected peptide H-CHHLEPGG-SEA^{off} **5a** and of the mixed disulfide H-C(MPAA)HHLEPGG-SEA^{off} between the Cys residue of **5a** and MPAA. Notably, incubation of the mixture for up to 48 h showed no significant evolution. In particular, polymerization or cyclization side reactions that could arise from the partial reduction of SEA^{off} into SEA^{on} by MPAA were not observed. We have verified that, when TCEP and MPAA were present in the mixture, cyclization of peptide **5a** occurred by native peptide ligation of the N-terminal Cys residue with the C-terminal SEA^{on} group after in situ reduction of both disulfides. The stability of the SEA^{off} group in the presence of MPAA led us to examine the first ligation step between segments **3a** and **4a** under the same experimental conditions. LC–MS analysis of the reaction mixture showed the successful formation of peptide **6a** featuring a SEA^{off} group at the C terminus, together with 30% of the mixed disulfide between the internal Cys residue of **6a** and MPAA. Disulfide-bond formation between the Cys residue of segment **5a** and MPAA was reversible in the presence of an excess of MPAA and did not disturb the ligation process. The ligation product **6a** and the mixed disulfide accounted for more than 90% of the total absorbance (see the Supporting Information), thereby showing the efficiency of the first ligation step. Here again, the SEA group remained in the off state, and no side-product formation was observed even for extended reaction times.

Next, ligation 2 was started by the addition of TCEP and of Cys segment **7a** to the reaction mixture. We observed clean formation of the target peptide **8a**, which was isolated in good yield (59% overall).

Currently, we are using this one-pot strategy for synthesizing domains of the hepatocyte growth factor (HGF). HGF is the ligand of the MET tyrosine kinase receptor, which is frequently deregulated in human tumors.[20] The role of individual HGF domains for MET activation is still under debate.[21] Protein total synthesis will allow HGF-domain analogues to be produced and thus allow the role of individual domains as well as the mechanism of action of this growth factor to be deciphered. Herein we describe the synthesis of the K1 domain of HGF, which is part of HGF's high-affinity binding site for MET[22] but acts as a weak agonist.[21] K1 is composed of 85 amino acid residues (HGF 125–209) and is stabilized by three intrachain disulfide bridges. K1 was assembled from HGF segments 125–148 (**3b**), 149–176 (**4b**), and 177–209 (**7b**, Figure 1a). Assembly of the K1 domain started by ligation of thioester segment **3b** with SEA^{off} segment **4b** (ligation 1, Lys–Cys (K–C) junction). The reaction led to the successful formation of SEA^{off} segment **6b** after 20 h at 37 °C. TCEP and Cys peptide amide segment **7b** were then added to the reaction mixture, thereby triggering the second ligation step (ligation 2, Tyr–Cys (Y–C) junction) in a one-pot process. This process led to the successful formation of linear K1 HGF domain **8b** after 48 h of reaction. The LC–MS analysis of the reaction mixture highlights the high quality of the crude product (Figure 1b left). The linear K1 HGF domain was purified by reversed-phase HPLC to give homogeneous material (11.5 mg, 39% overall, Figure 1b, right). The linear K1 domain was folded successfully by using the glutathione–glutathione disulfide redox system and dialyzed (Figure 1c).

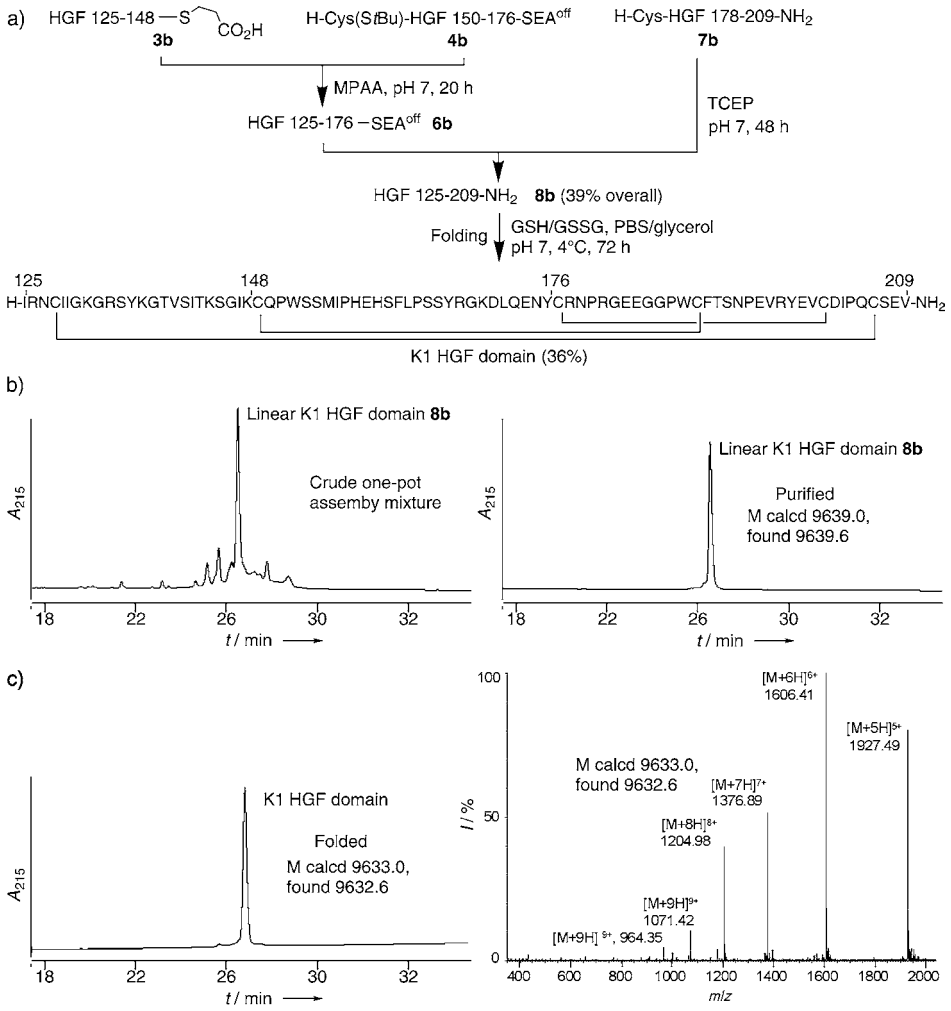


Figure 1. Total synthesis of the K1 HGF domain. a) One-pot assembly and folding of the K1 domain. b) LC–MS chromatograms the crude (left) or HPLC-purified (right) linear K1 domain **8b** with detection by UV/Vis absorption at 215 nm. c) LC–MS chromatogram and spectrum of the folded K1 domain (left: reversed-phase HPLC trace, right: MS data). GSH/GSSG = glutathione redox system. PBS = phosphate-buffered saline.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

The cysteine pairing in the K1 domain is established with certainty based on the crystal structure of NK1, a naturally occurring fragment of HGF.[23] Proteomic analysis of the synthetic K1 domain using enzymatic digestion and subsequent mass-spectrometry analysis of the formed peptide fragments revealed the formation of the native disulfide-bridge pattern during folding (see the Supporting Information).

Finally, the structural integrity and functionality of K1 were confirmed through the study of its biological activity. The treatment of MET-expressing cells with HGF triggers intracellular MET-tyrosine-kinase phosphorylation in a few minutes and subsequent recruitment of adaptors that mediate intracellular signal transduction. In cell culture HGF triggers very pronounced phenotypes and in particular the scattering phenotype on epithelial cells.[20] Therefore, we have evaluated the ability of synthetic K1 to promote MET activation and scattering of Madin–Darby Canine Kidney cells (MDCK, Figure 2), which are considered as the reference model for HGF-induced MET activation, as well as of other epithelial

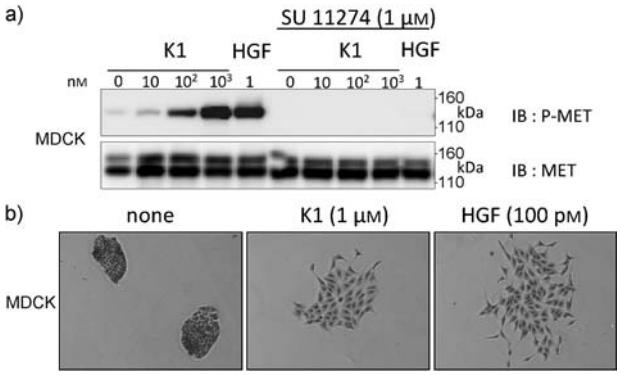


Figure 2. Cellular MET activation by synthetic K1. a) MDCK cells were treated with mature HGF or K1 domain in the absence or presence of MET kinase inhibitor SU11274. Cell lysates were next analyzed by specific total MET or phospho-MET immunoblot (IB). b) MDCK cell islets were incubated in presence of culture media (none), 1 μM K1, or 100 pM HGF (magnification×100).

cell lines (see the Supporting Information). Western blot analysis showed that K1 induced MET phosphorylation in a dose-responsive manner (Figure 2a). In the scattering assay (Figure 2b), K1 induced the disruption of compact islets and cell migration. The specificity of K1-induced MET activation was verified by using the MET-selective inhibitor SU11274,[24] which inhibited both K1-induced MET phosphorylation (Figure 2a) and scattering of MDCK cells (see the Supporting Information).

In summary, the stability of the SEA^{off} group in the conditions used for NCL and the possibility to easily switch on the SEA group with a phosphine permitted the one-pot assembly of three peptide segments in the N-to-C direction. Both ligations were performed under neutral conditions in aqueous solution. This strategy permitted the successful synthesis of biologically active K1 domain from HGF. The total synthesis of K1-domain variants and of other HGF domains is in progress to elucidate the mechanisms underlying HGF–MET signaling.

Received: August 18, 2011
Published online: November 16, 2011

Keywords: growth factors · native chemical ligation · proteins · solid-phase synthesis · synthetic methods

[1] S. B. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338–351.
[2] D. Bang, B. L. Pentelute, S. B. Kent, *Angew. Chem.* **2006**, *118*, 4089–4092; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985–3988.
[3] D. Bang, S. B. Kent, *Angew. Chem.* **2004**, *116*, 2588–2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538.
[4] Z. Tan, S. Shang, S. J. Danishefsky, *Angew. Chem.* **2010**, *122*, 9690–9693; *Angew. Chem. Int. Ed.* **2010**, *49*, 9500–9503.
[5] J. S. Zheng, H. K. Cui, G. M. Fang, W. X. Xi, L. Liu, *ChemBioChem* **2010**, *11*, 511–515.
[6] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
[7] a) C. P. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182–10228; *Angew. Chem. Int. Ed.* **2008**, *47*, 10030–10074; b) H. Rohde, O. Seitz, *Biopolymers* **2010**, *94*, 551–559.
[8] a) K. S. Kumar, L. Spasser, L. A. Erlich, S. N. Bavikar, A. Brik, *Angew. Chem.* **2010**, *122*, 9312–9317; *Angew. Chem. Int. Ed.* **2010**, *49*, 9126–9131; b) K. S. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, *Angew. Chem.* **2011**, *123*, 6261–6265; *Angew. Chem. Int. Ed.* **2011**, *50*, 6137–6141; c) S. Luisier, M. Avital-Shmilovici, M. A. Weiss, S. B. Kent, *Chem. Commun.* **2010**, *46*, 8177–8179; d) S. Lahiri, M. Brehs, D. Olschewski, C. F. Becker, *Angew. Chem.* **2011**, *123*, 4074–4078; *Angew. Chem. Int. Ed.* **2011**, *50*, 3988–3992.
[9] B. L. Pentelute, K. Mandal, Z. P. Gates, M. R. Sawaya, T. O. Yeates, S. B. Kent, *Chem. Commun.* **2010**, *46*, 8174–8176.
[10] K. Mandal, S. B. Kent, *Angew. Chem.* **2011**, *123*, 8179–8183; *Angew. Chem. Int. Ed.* **2011**, *50*, 8029–8033.
[11] a) T. Kawakami, S. Aimoto, *Chem. Lett.* **2007**, *36*, 76–77; L. A. Erlich, K. S. Kumar, M. Haj-Yahya, P. E. Dawson, A. Brik, *Org. Biomol. Chem.* **2010**, *8*, 2392–2396; b) L. E. Canne, P. Botti, R. J. Simon, Y. Chen, E. A. Dennis, S. B. H. Kent, *J. Am. Chem. Soc.* **1999**, *121*, 8720–8727; c) G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chem.* **2011**, *123*, 7787–7791; *Angew. Chem. Int. Ed.* **2011**, *50*, 7645–7649.
[12] K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto, A. Otaka, *ChemBioChem* **2011**, *12*, 1840–1844.
[13] a) N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain, O. Melnyk, *Org. Lett.* **2010**, *12*, 5238–5241. See also b) W. Hou, X. Zhang, F. Li, C. F. Liu, *Org. Lett.* **2011**, *13*, 386–389.
[14] E. C. Johnson, S. B. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
[15] P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
[16] For a recent review on the Fmoc-SPPS of peptide thioesters: F. Mende, O. Seitz, *Angew. Chem.* **2011**, *123*, 1266–1274; *Angew. Chem. Int. Ed.* **2011**, *50*, 1232–1240.
[17] J. Dheur, N. Ollivier, A. Vallin, O. Melnyk, *J. Org. Chem.* **2011**, *76*, 3194–3202.
[18] G. B. Fields, R. L. Noble, *Int. J. Pept. Protein Res.* **1990**, *35*, 161–214.
[19] Conversion of SEA^{on} into SEA^{off} is carried out by addition of an excess of iodine to the crude peptide dissolved in water/acetic acid. After few seconds, dithiothreitol (DTT) is added to decompose the excess of iodine. Cys(S*t*Bu) and SEA^{off} are unaffected, because DTT is unable to reduce disulfides at low pH values. Other residues such as methionine or tryptophan are

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

not affected during the very short exposure to iodine. The product is then purified by HPLC.

[20] L. Trusolino, A. Bertotti, P. M. Comoglio, *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 834 – 848.

[21] O. Holmes, S. Pillozzi, J. A. Deakin, F. Carafoli, L. Kemp, P. J. Butler, M. Lyon, E. Gherardi, *J. Mol. Biol.* **2007**, *367*, 395 – 408.

[22] N. A. Lokker, L. G. Presta, P. J. Godowski, *Protein Eng.* **1994**, *7*, 895 – 903.

[23] M. Ultsch, N. A. Lokker, P. J. Godowski, A. M. de Vos, *Structure* **1998**, *6*, 1383 – 1393.

[24] X. Wang, P. Le, C. Liang, J. Chan, D. Kiewlich, T. Miller, D. Harris, L. Sun, A. Rice, S. Vasile, R. A. Blake, A. R. Howlett, N. Patel, G. McMahon, K. E. Lipson, *Mol. Cancer Ther.* **2003**, *2*, 1085 – 1092.

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim