HAL
open science

Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments

Laurent Raibaut, Hélène Adihou, Rémi Desmet, Agnès F. Delmas, Vincent Aucagne, Oleg Melnyk

To cite this version:

Laurent Raibaut, Hélène Adihou, Rémi Desmet, Agnès F. Delmas, Vincent Aucagne, et al.. Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments. Chemical Science, 2013, 4, pp.4061-4066. 10.1039/c3sc51824h . hal-00944814v1

HAL Id: hal-00944814

https://hal.science/hal-00944814v1

Submitted on 19 Feb 2014 (v1), last revised 10 Jul 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chemical Science

RSC Publishing

EDGE ARTICLE

View Article Online
View Journal | View Issue

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

Cite this: *Chem. Sci.*, 2013, **4**, 4061

Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments†

Laurent Raibaut,[a] Hélène Adihou,[b] Rémi Desmet,[a] Agnès F. Delmas,[b] Vincent Aucagne*[b] and Oleg Melnyk*[a]

Received 28th June 2013
Accepted 22nd July 2013

DOI: 10.1039/c3sc51824h

www.rsc.org/chemicalscience

Up to now, the advantages of solid phase protein synthesis have been largely under-utilized due to the difficulty of designing a simple and efficient elongation cycle enabling the concatenation of unprotected peptide segments. The combination of selective N-terminal anchoring (N_3-Esoc linker) with the blocked thioester properties of the SEA^{off} group enabled the solid phase concatenation of unprotected peptide segments by N-to-C sequential formation of native peptide bonds. The strategy was applied to the synthesis of a 60 amino acid-long latent peptide thioester or to the assembly of five peptide segments to give a 15 kDa polypeptide.

Introduction

Modern protein chemical synthesis is made possible by the combination of several essential chemical tools. Solid phase peptide synthesis (SPPS)[1] enables the synthesis of peptide segments by the iterative coupling of protected amino acids to a solid support. Other major tools are chemoselective peptide bond forming reactions which enable the coupling of unprotected peptide segments in water.[2–7] Among these reactions, native chemical ligation (NCL)[2,8] which is based on the coupling of a C-terminal peptide thioester with an N-terminal cysteinyl peptide, is undoubtedly the most popular reaction for protein total synthesis.[9]

The ligation of two peptide segments potentially gives access to polypeptides composed of up to ~100 amino acid residues, considering the current limits of SPPS. The synthesis of larger polypeptides usually requires the chemical ligation of three or more peptide segments either using iterative[10,11] or convergent[12–15] assembly schemes.[13,16–20] While recent advances have led to the development of efficient one-pot three-segment sequential assembly methods,[12,15,21–23] the ligation of more than three segments usually involves time-consuming intermediate isolation or purification steps, accompanied by significant material losses.

Not surprisingly, the advantages of solid phase synthesis were early considered as a potential solution to the limitations encountered in solution.[24–27] Elongation in the C-to-N direction can facilitate the synthesis of large N-terminal cysteinyl peptides, and elongation in the N-to-C direction potentially gives access to large C-terminal peptide thioester surrogates (Fig. 1). Importantly, the combination of both solid phase strategies could significantly extend the limits of protein chemical synthesis using convergent NCL approaches.[12,13,20]

The internal peptide segments used for sequential solid-supported NCLs (Fig. 1) feature both an N-terminal cysteine and a C-terminal thioester group, one of which must be protected during the ligation to avoid potential cyclization or oligomerization. This protection must be compatible with the method used to attach the starting peptide segment onto a water-compatible solid-support.

Several protecting groups (PG) are available for the temporary protection of N-terminal cysteines (Fig. 1) for C-to-N solid

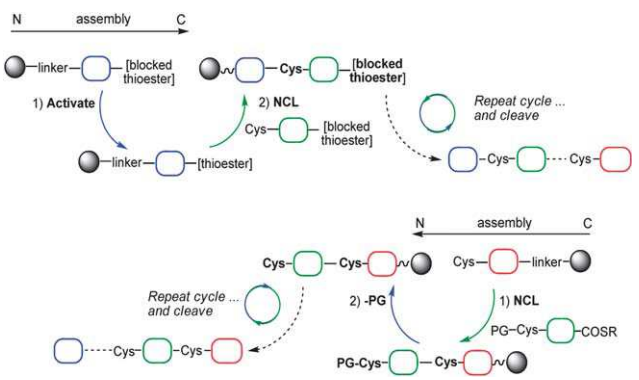


Fig. 1 General principle of solid phase sequential NCL processes in the C-to-N or N-to-C direction (PG = protecting group).

[a]*CNRS UMR 8161, Institut Pasteur de Lille, 59021, Lille, France. E-mail: oleg.melnyk@ibl.fr*

[b]*Centre de Biophysique Moléculaire, CNRS UPR 4301, Rue Charles Sadron, 45071 Orléans cedex 2, France. E-mail: aucagne@cnrs-orleans.fr*

† Electronic supplementary information (ESI) available: Experimental procedures and characterization for all compounds. See DOI: 10.1039/c3sc51824h

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

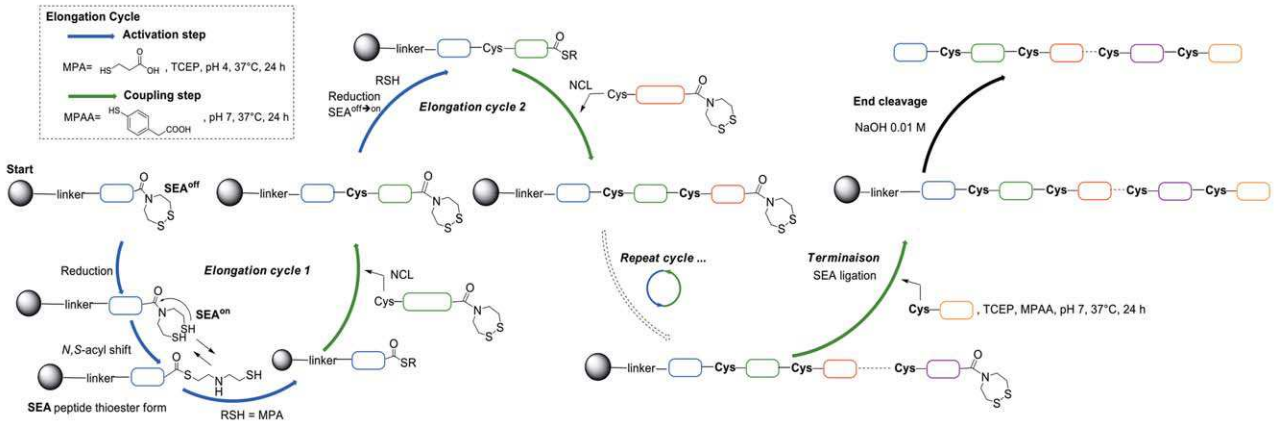


Fig. 2 Principle of the N-to-C elongation cycle based on SEA and NCL chemistries.

phase assembly.[24–26] On the other hand, the design of an elongation cycle in the N-to-C direction is far more challenging. It requires a less common blocked C-terminal thioester precursor, which must be *unreactive* during the NCL reaction, and then selectively activated before the next ligation (Fig. 1).[25] The activation and ligation steps must be high yielding to allow the elongation cycle to be repeated several times. Ideally, the thioester precursor should be stable under the conditions used for the release from the resin to facilitate the monitoring of the ligation steps or to yield large C-terminal peptide thioester precursors.

The unique example of an N-to-C solid phase sequential NCL process was reported by Kent and co-workers in 1999, who exploited the thiocarboxylate group as a masked thioester surrogate for the assembly of three peptide segments.[25] The partial cyclization of the internal segment was mentioned by the authors due to the residual reactivity of the thiocarboxylate group for the N-terminal Cys residue during the NCL reaction. The thiocarboxylate group was subsequently converted into an alkylthioester by selective alkylation with bromoacetic acid at pH $\sim$ 4.

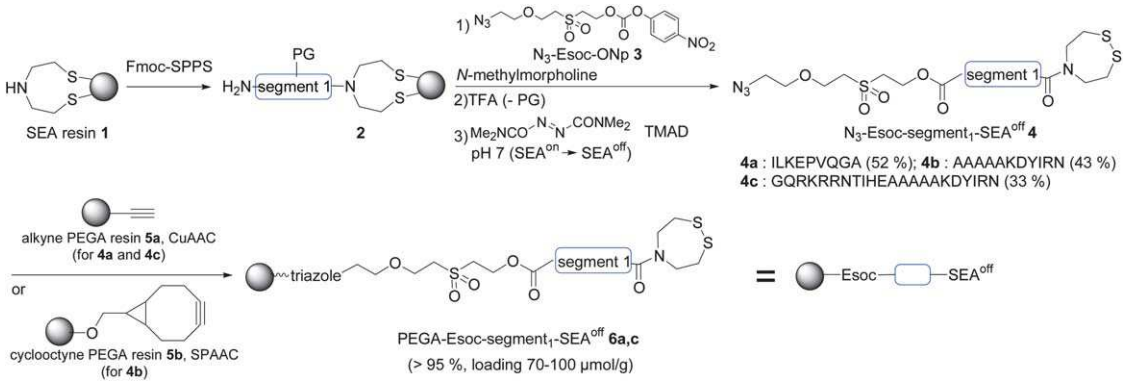
More recently, a few C-terminal peptide thioesters precursors have been designed based on the rearrangement of peptide esters or amides featuring an internal thiol nucleophile[28] or on the nitrosation of peptide hydrazides followed by treatment with an exogeneous thiol.[17] However, the compatibility of these methods for the N-to-C solid phase assembly of large peptides remains to be established.

We sought to develop a simple and efficient N-to-C solid phase assembly strategy, considering the unmet needs in this area and the high potential impact in the field of protein total synthesis. We show here that the combination of a bis(2-sulfanylethyl)amide (SEA^{on}) thioester surrogate (Fig. 2)[22,29,30] with a recently developed N-terminal anchoring strategy[31] achieves this goal.

Results and discussion

The SEA^{on} group is an amide which features two thiol groups in a 1,5 relationship to the carbonyl moiety. It rearranges spontaneously in water by an $N \rightarrow S$-acyl shift mechanism to give a transient thioester (Fig. 2). The ability of the SEA^{on} transient thioester to react with thiols enables the installation of a C-terminal thioester group by reaction with an exogeneous thiol RSH (activation step, Fig. 2) and the coupling of the last Cys segment using SEA ligation (termination step, Fig. 2).

In contrast, the $N \rightarrow S$-acyl shift cannot occur in the SEA^{off} form because the thiol groups are masked within the cyclic



Scheme 1 Synthesis and immobilization of the first segments.

 This journal is © The Royal Society of Chemistry 2013

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

Table 1 Polypeptides synthesized by N-to-C solid phase iterative native peptide ligation

Peptide[a]	Number of amino acids	Molecular weight	Yield[b] (%)	Number of cycles/number of chemical steps	Average yield per step (%)
9	35	3731	21	2/7	80
10	94	10 401	18	2/7	78
11	60	6779	22	2/6	78
12	136	15 662	6.5	3/9	74

[a] The sequences of model peptides **10–12** are derived from human hepatocyte growth factor. [b] Starting from N_3-Esoc-segment$_1$-SEA^{off} **4** (HPLC purified).

disulfide. The SEA^{off} group was used as a blocked thioester surrogate during the NCL step of the N-to-C elongation cycle (Fig. 2). Importantly, the reduction of SEA^{off} to SEA^{on} by tris(2-carboxyethylphosphine) (TCEP) enables the concatenation of the elongation cycles using mild conditions. The SEA^{off} group is also resistant to various nucleophiles and bases.[32] We exploited this property to set up a quality control for the elongation cycle (Fig. 2) and to isolate a large C-terminal SEA^{off} peptide (Fig. 3).

As an anchoring strategy we chose the N_3-Esoc linker (2-[2-(2-azido-ethoxy)ethyl sulfonyl]-2-ethoxycarbonyl) developed for capture and release purification and peptidomimetic ligation of unprotected peptides (Scheme 1).[31,33] This linker features a base-labile alkylsulfonylethyloxycarbamate core compatible with NCL,[34] and an azido group for chemoselective immobilization on a water-compatible matrix, such as polyethylene glycol dimethyl acrylamide copolymer (PEGA^{1900}) employed in this work.

First, the peptide chain was assembled on SEA polystyrene resin **1** (ref. 29) using standard Fmoc-SPPS and further derivatized at its N-terminus by N_3-Esoc-ONp reagent[31] **3** featuring a *para*-nitrophenyl (ONp) activated carbonate (Scheme 1). The N_3-Esoc-segment$_1$-SEA^{on} peptides obtained after deprotection and cleavage in trifluoroacetic acid (TFA) were subsequently oxidized into SEA^{off} peptides **4a–c** with *N,N,N′,N′*-tetramethylazodicarboxamide (TMAD). Peptides **4a** and **4c** were efficiently loaded at neutral pH onto alkyne PEGA^{1900} resin **5a** using Cu^{I}-catalyzed azide/alkyne cycloaddition (CuAAC) ligation,[35,36] resulting in the formation of a stable 1,4-disubstituted 1,2,3-triazole linkage. In contrast, peptide **4b** was soluble only at acidic pH which precluded the use of CuAAC for the immobilization step due to a significant decrease in reaction rate. Alternatively, the strain-promoted azide/alkyne cycloaddition (SPAAC)[37] of peptide **4b** with the bicyclononyne[38] PEGA^{1900} resin **5b** at pH ∼ 2 successfully furnished the supported peptide **6b**.

The first step of the N-to-C elongation cycle, *i.e.*, the activation step, converts the C-terminal SEA^{off} group into a thioester by treating peptidyl resin **6** with 3-mercaptopropionic acid and TCEP at pH 4.[39] The rate of this reaction depends on the nature

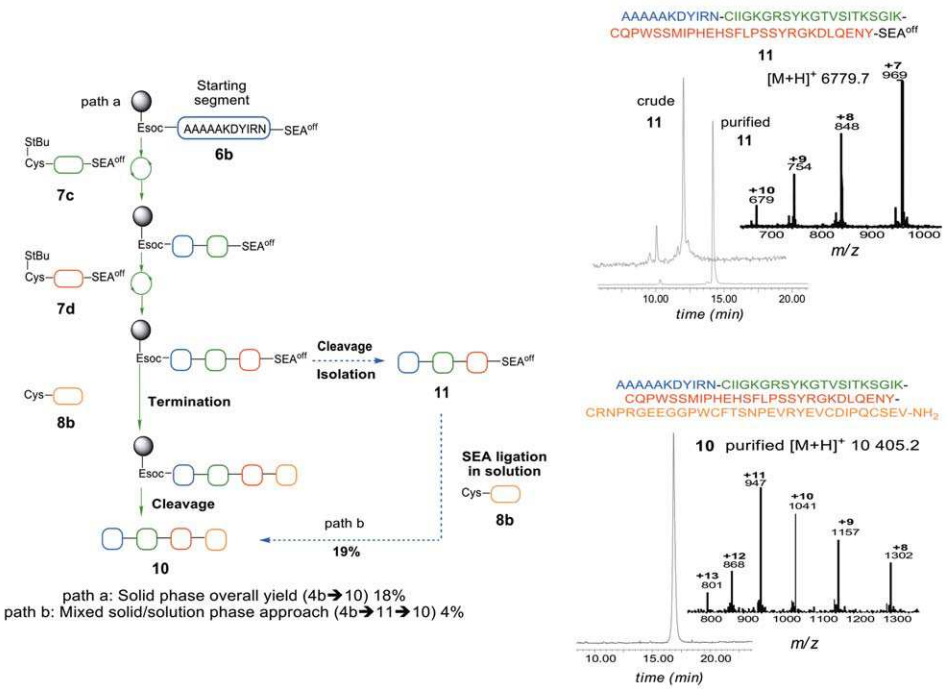


Fig. 3 N-to-C assembly of peptide **10** and of SEA^{off} peptide **11**.

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

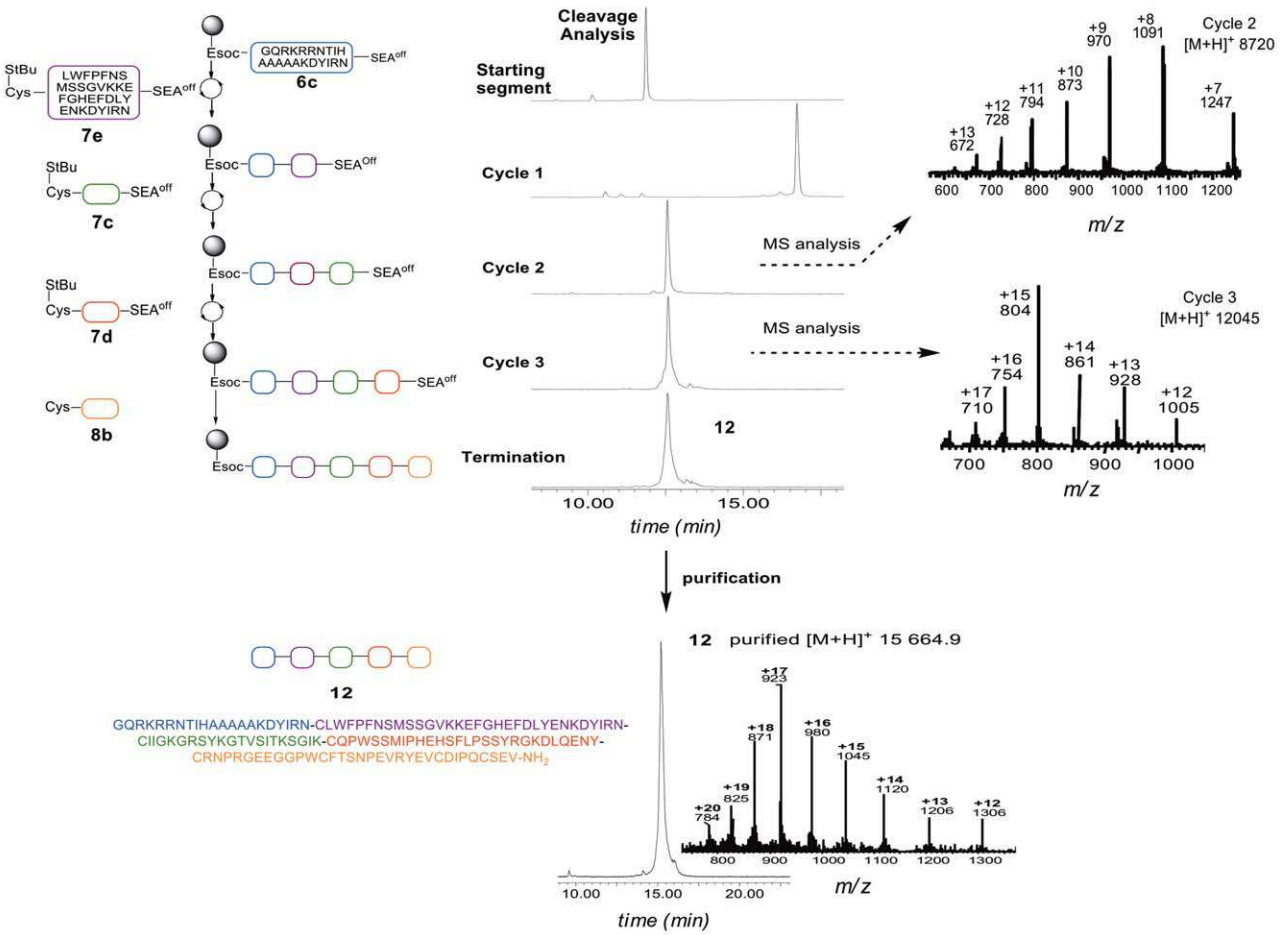


Fig. 4 N-to-C assembly of peptide **12** using five peptide segments.

of the C-terminal residue. An activation time of 24 h was satisfactory for all the activations performed in this study (Table 1). Following activation, it was crucial to remove excess TCEP by extensive washing to avoid partial activation of the SEA^{off} group during the following NCL reaction.

Our strategy relies on the inert nature of the SEA^{off} internal segments during the NCL reaction. Previous studies[22,40] suggest that it might be the case for Tyr[22] and Gln[40] as C-terminal residues while permitting the use of an aryl thiol additive such as 4-mercaptophenyl acetic acid (MPAA)[41] to catalyze the NCL reaction, thereby allowing the coupling step to be carried out in the best conditions in terms of kinetics. The synthesis of the 35 amino acid peptide ILKEPVQGA-CHHLEPGG-CHHLEPAG-CILKEPVHGA-NH_2 **9** (Table 1), was used as a test case for probing the inertness of the SEA^{off} group during the NCL reaction in the worst case, *i.e.*, with a C-terminal Gly residue for both internal segments C(S*t*Bu)HHLEPGG-SEA^{off} **7a** and C(S*t*Bu)HHLEPAG-SEA^{off} **7b**. Indeed, the Gly residue is one of the most reactive amino acid residues in SEA native peptide ligation, meaning that partial activation of the SEA^{off} group during the NCL reaction should lead to the rapid formation of side products. Peptide **9** was assembled starting from N_3-Esoc-segment-SEA^{off} **4a** (Scheme 1) and only 1.5 equivalents of segments **7a,b** and CILKEPVHGA-NH_2 **8a**.‡

LC-MS analysis of intermediate SEA^{off} peptide segments and of target polypeptide **9** after basic cleavage from the resin indicated insertion of only one copy of internal segments **7a** and **7b** in the peptide chain (Fig. S5 and 7, ESI†). Furthermore, LC-MS analysis of the beads supernatant during the NCL reaction showed that potential cyclization or oligomerization side products were below detectable levels (Fig. S6, ESI†). Note that the presence of the SEA group exclusively in the SEA^{off} form even for prolonged reaction times opens the possibility of recovering the internal SEA^{off} peptide segment if used in large excess. This peculiarity was not exploited in this work because the NCL steps proceeded efficiently by using only a slight excess (1.5 equiv.) of the segments. Overall, the assembly of model peptide **9** proved highly efficient, being isolated in a 21% yield after seven chemical steps, meaning an average yield of 80% per step (Table 1).

Gratifyingly, the four segment assembly process proved equally efficient for assembling the 94 amino acid peptide **10** starting from solid support **6b** (Fig. 3, Table 1). Most importantly, we were able to isolate also the intermediate 60 amino acid SEA^{off} peptide segment **11** by performing the cleavage step on a preparative scale (Fig. 3, Table 1). Note that SEA^{off} peptide **11** constitutes the first example of a solid phase synthesis of a large C-terminal peptide thioester surrogate by sequential

 This journal is © The Royal Society of Chemistry 2013

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

NCLs. To illustrate the importance of this approach for mixed solid/liquid phase convergent protein synthesis, SEA^{off} peptide **11** was subsequently ligated with Cys peptide **8b** in solution to yield again peptide **10** (Fig. 3, see also Fig. S16, ESI†) which was found identical by LC-MS to peptide **10** assembled on the solid phase (see Fig. S17, ESI†). This experiment demonstrates the functionality of the large SEA^{off} peptide **11** and the potential utility of the solid phase assembly method for providing access to larger proteins using convergent approaches. Note that the overall yield for peptide **10** produced by the solid phase approach is significantly higher than that for the mixed solid/liquid phase one. This result illustrates the interest of the solid phase approach for simplifying the isolation step and improving yield.

The assembly of five peptide segments starting from solid-supported peptide **6c** was also very efficient to produce the 136 amino acid peptide **12** (Fig. 4, Table 1). The progress of the solid phase assembly was visualized by HPLC and MS analyses of the cleaved intermediate SEA^{off} peptide segments, and the high purity of the crude target peptide **12**, which was successfully purified and characterized. This example illustrates the robustness of the iterative elongation cycle and the potential of this methodology for the straightforward synthesis of large peptides.

Conclusions

Up to now, the advantages of solid phase ligation have been largely under-utilized particularly in the N-to-C direction due to the difficulty of designing a simple and efficient elongation cycle. Here we have demonstrated that the combination of the N_3-Esoc N-terminal linker and SEA chemistry is a solution to this long-lasting problem. This method enables the solid phase N-to-C sequential NCL of up to five peptide segments for the synthesis of a 15 kDa polypeptide as well as the isolation of a 60 amino acid thioester surrogate. We believe that this method will help to significantly extend the size limit of the polypeptides that can be produced by chemical synthesis.

Acknowledgements

This work was supported by ANR grants SEA SPPS (12-EMMA-0006-01) and SynProt (JCJC08-0138). LR and HA thank the Région Nord Pas-de-Calais and the Région Centre councils for a PhD fellowship.

Notes and references

‡ The cysteine *tert*-butylsulfenyl protecting groups (S*t*Bu), which are introduced temporarily to simplify the synthesis of internal peptide segments **7a–d** using SEA polystyrene resin **1**, are removed *in situ* during NCL reaction due to the presence of MPAA. See ref. 22.

1 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
2 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776–779.
3 J. W. Bode, R. M. Fox and K. D. Baucom, *Angew. Chem., Int. Ed.*, 2006, **45**, 1248–1252.
4 E. Saxon, J. I. Armstrong and C. R. Bertozzi, *Org. Lett.*, 2000, **2**, 2141–2143.
5 B. L. Nilsson, L. L. Kiessling and R. T. Raines, *Org. Lett.*, 2000, **2**, 1939–1941.
6 V. R. Pattabiraman and J. W. Bode, *Nature*, 2011, **480**, 471–479.
7 Y. Zhang, C. Xu, H. Y. Lam, C. L. Lee and X. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 6657–6662.
8 S. B. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338–351.
9 C. P. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030–10074.
10 F. Wintermann and S. Engelbrecht, *Angew. Chem., Int. Ed.*, 2013, **52**, 1309–1313.
11 A. O. Ogunkoya, V. R. Pattabiraman and J. W. Bode, *Angew. Chem., Int. Ed.*, 2012, **51**, 9693–9697.
12 D. Bang, B. L. Pentelute and S. B. Kent, *Angew. Chem., Int. Ed.*, 2006, **45**, 3985–3988.
13 G.-M. Fang, J.-X. Wang and L. Liu, *Angew. Chem., Int. Ed.*, 2012, **51**, 10347–10350.
14 T. Durek, V. Y. Torbeev and S. B. Kent, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 4846–4851.
15 K. Sato, A. Shigenaga, K. Kitakaze, K. Sakamoto, D. Tsuji, K. Itoh and A. Otaka, *Angew. Chem., Int. Ed.*, 2013, **52**, 7855–7859.
16 K. S. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon and A. Brik, *Angew. Chem., Int. Ed.*, 2011, **50**, 6137–6141.
17 G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui and L. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 7645–7649.
18 S. Lahiri, M. Brehs, D. Olschewski and C. F. Becker, *Angew. Chem., Int. Ed.*, 2011, **50**, 3988–3992.
19 C. Piontek, D. Varon Silva, C. Heinlein, C. Pohner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1941–1945.
20 P. Wang, S. Dong, J. A. Brailsford, K. Iyer, S. D. Townsend, Q. Zhang, R. C. Hendrickson, J. Shieh, M. A. Moore and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2012, **51**, 11576–11584.
21 D. Bang and S. B. Kent, *Angew. Chem., Int. Ed.*, 2004, **43**, 2534–2538.
22 N. Ollivier, J. Vicogne, A. Vallin, H. Drobecq, R. Desmet, O. El-Mahdi, B. Leclercq, G. Goormachtigh, V. Fafeur and O. Melnyk, *Angew. Chem., Int. Ed.*, 2012, **51**, 209–213.
23 L. Raibaut, N. Ollivier and O. Melnyk, *Chem. Soc. Rev.*, 2012, **41**, 7001–7015.
24 A. Brik, E. Keinan and P. E. Dawson, *J. Org. Chem.*, 2000, **65**, 3829–3835.
25 L. E. Canne, P. Botti, R. J. Simon, Y. Chen, E. A. Dennis and S. B. H. Kent, *J. Am. Chem. Soc.*, 1999, **121**, 8720–8727.
26 E. C. Johnson, T. Durek and S. B. Kent, *Angew. Chem., Int. Ed.*, 2006, **45**, 3283–3287.
27 G. J. Cotton and T. W. Muir, *Chem. Biol.*, 2000, **7**, 253–261.
28 J.-S. Zheng, S. Tang, Y.-C. Huang and L. Liu, *Acc. Chem. Res.*, 2013, DOI: 10.1021/ar400012w.
29 N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain and O. Melnyk, *Org. Lett.*, 2010, **12**, 5238–5241.

Published on 24 July 2013. Downloaded by Aarhus University Library on 15/10/2013 17:34:04.

30 W. Hou, X. Zhang, F. Li and C. F. Liu, *Org. Lett.*, 2011, **13**, 386–389.
31 V. Aucagne, I. E. Valverde, P. Marceau, M. Galibert, N. Dendane and A. F. Delmas, *Angew. Chem., Int. Ed.*, 2012, **51**, 11320–11324.
32 E. Boll, J. Dheur, H. Drobecq and O. Melnyk, *Org. Lett.*, 2012, **14**, 2222–2225.
33 I. E. Valverde, F. Lecaille, G. Lalmanach, V. Aucagne and A. F. Delmas, *Angew. Chem., Int. Ed.*, 2012, **51**, 718–722.
34 J. A. Camarero, G. J. Cotton, A. Adeva and T. W. Muir, *J. Pept. Res.*, 1998, **51**, 303–316.
35 C. W. Tornoe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.
36 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
37 N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046–15047.
38 J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl and F. L. van Delft, *Angew. Chem., Int. Ed.*, 2010, **49**, 9422–9425.
39 J. Dheur, N. Ollivier, A. Vallin and O. Melnyk, *J. Org. Chem.*, 2011, **76**, 3194–3202.
40 L. Raibaut, J. Vicogne, B. Leclercq, H. Drobecq, R. Desmet and O. Melnyk, *Bioorg. Med. Chem.*, 2013, **21**, 3486–3494.
41 E. C. Johnson and S. B. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640–6646.

 This journal is © The Royal Society of Chemistry 2013