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:

HAL Id: hal-00944814
https://hal.archives-ouvertes.fr/hal-00944814v2
Submitted on 10 Jul 2014

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

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

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 (N3-Esoc linker) with the blocked thioester properties of the SEA 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) 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. Among these reactions, native chemical ligation (NCL) 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.

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 or convergent assembly schemes. While recent advances have led to the development of efficient one-pot three-segment sequential assembly methods, 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. 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. Importantly, the combination of both solid phase strategies could significantly extend the limits of protein chemical synthesis using convergent NCL approaches.

The internal peptide segments used for sequential solid-supported NCLs 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 for C-to-N 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 for C-to-N 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 for C-to-N 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 for C-to-N 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 for C-to-N 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.
phase assembly.\textsuperscript{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).\textsuperscript{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.\textsuperscript{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 $\approx 4$.

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\textsuperscript{28} or on the nitrosation of peptide hydrazides followed by treatment with an exogeneous thiol.\textsuperscript{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-sulfanylyl)amide (SEAon) thioester surrogate (Fig. 2)\textsuperscript{22,29,30} with a recently developed N-terminal anchoring strategy\textsuperscript{31} achieves this goal.

Results and discussion

The SEAon 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 $N/S$-acyl shift mechanism to give a transient thioester (Fig. 2). The ability of the SEAon 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 $N/S$-acyl shift cannot occur in the SEAoff form because the thiol groups are masked within the cyclic

---

**Fig. 2** Principle of the N-to-C elongation cycle based on SEA and NCL chemistries.
disulfide. The SEAeff 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 SEAeff to SEAon by tris(2-carboxyethylphosphine) (TCEP) enables the concatenation of the elongation cycles using mild conditions. The SEAeff 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 SEAeff peptide (Fig. 3).

As an anchoring strategy we chose the N3-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 alkylsulfonylethoxycarbamate core compatible with NCL,34 and an azido group for chemoselective immobilization on a water-compatible matrix, such as polyethylene glycol dimethyl acrylamide copolymer (PEGA1900) 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 N3-Esoc-ONp reagent33 3 featuring a para-nitrophenyl (ONp) activated carbonate (Scheme 1). The N3-Esoc-segment1-SEAeff peptides obtained after deprotection and cleavage in trifluoroacetic acid (TFA) were subsequently oxidized into SEAeff peptides 4a-c with N,N,N’,N’-tetramethylazodicarboxamide (TMAD). Peptides 4a and 4c were efficiently loaded at neutral pH onto alkyne PEGA1900 resin 5a using CuI-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 bicyclononyne38 PEGA1900 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 SEAeff 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

---

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

<table>
<thead>
<tr>
<th>Peptidea</th>
<th>Number of amino acids</th>
<th>Molecular weight</th>
<th>Yieldb (%)</th>
<th>Number of cycles/number of chemical steps</th>
<th>Average yield per step (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>35</td>
<td>3731</td>
<td>21</td>
<td>2/7</td>
<td>80</td>
</tr>
<tr>
<td>10</td>
<td>94</td>
<td>10 401</td>
<td>18</td>
<td>2/7</td>
<td>78</td>
</tr>
<tr>
<td>11</td>
<td>60</td>
<td>6779</td>
<td>22</td>
<td>2/6</td>
<td>78</td>
</tr>
<tr>
<td>12</td>
<td>136</td>
<td>15 662</td>
<td>6.5</td>
<td>3/9</td>
<td>74</td>
</tr>
</tbody>
</table>

a The sequences of model peptides 10–12 are derived from human hepatocyte growth factor. b Starting from N3-Esoc-segment1-SEAeff 4 (HPLC purified).
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<sup>off</sup> group during the following NCL reaction.

Our strategy relies on the inert nature of the SEA<sup>off</sup> internal segments during the NCL reaction. Previous studies<sup>22,40</sup> suggest that it might be the case for Tyr<sup>22</sup> and Gln<sup>40</sup> as C-terminal residues while permitting the use of an aryl thiol additive such as 4-mercaptophenyl acetic acid (MPAA)<sup>41</sup> 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<sub>2</sub> (Table 1), was used as a test case for probing the inertness of the SEA<sup>off</sup> group during the NCL reaction in the worst case, i.e., with a C-terminal Gly residue for both internal segments C(S<sub>T</sub>Bu)HHLEPGG-SEA<sup>off</sup> 7a and C(S<sub>T</sub>Bu)HHLEPAG-SEA<sup>off</sup> 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<sup>off</sup> group during the NCL reaction should lead to the rapid formation of side products. Peptide 9 was assembled starting from N<sub>3</sub>-Esoc-segment-SEA<sup>off</sup> 4a (Scheme 1) and only 1.5 equivalents of segments 7a,b and CILKEPVHGA-NH<sub>2</sub> 8a.‡

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

LC-MS analysis of intermediate SEA<sup>off</sup> 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<sup>off</sup> form even for prolonged reaction times opens the possibility of recovering the internal SEA<sup>off</sup> 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<sup>off</sup> peptide segment 11 by performing the cleavage step on a preparative scale (Fig. 3, Table 1). Note that SEA<sup>off</sup> peptide 11 constitutes the first example of a solid phase synthesis of a large C-terminal peptide thioester surrogate by sequential
NCLs. To illustrate the importance of this approach for mixed solid/liquid phase convergent protein synthesis, SEA\textsuperscript{aff} peptide 11 was subsequently ligated with Cys peptide 8b in solution to yield again peptide 10 (Fig. 3, see also Fig. S16, ESI\textsuperscript{†}) which was found identical by LC-MS to peptide 10 assembled on the solid phase (see Fig. S17, ESI\textsuperscript{†}). This experiment demonstrates the functionality of the large SEA\textsuperscript{aff} 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\textsuperscript{aff} 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\textsubscript{T}–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**

\textsuperscript{†} The cysteine tert-butylsulfenyl protecting groups (SBu), which are introduced temporarily to simplify the synthesis of internal peptide segments 7a–d using SEA polystyrene resin 1, are removed \textit{in situ} during NCL reaction due to the presence of MPAA. See ref. 22.