

Cite this: *Chem. Sci.*, 2021, 12, 3194

All publication charges for this article have been paid for by the Royal Society of Chemistry


Received 1st November 2020

Accepted 10th January 2021

DOI: 10.1039/d0sc06001a

rsc.li/chemical-science

# A straightforward methodology to overcome solubility challenges for N-terminal cysteinyl peptide segments used in native chemical ligation†

Skander A. Abboud, El hadji Cisse, Michel Doudeau, H el ene B enedetti and Vincent Aucagne \*

One of the main limitations encountered during the chemical synthesis of proteins through native chemical ligation (NCL) is the limited solubility of some of the peptide segments. The most commonly used solution to overcome this problem is to derivatize the segment with a temporary solubilizing tag. Conveniently, the tag can be introduced on the thioester segment in such a way that it is removed concomitantly with the NCL reaction. We herein describe a generalization of this approach to N-terminal cysteinyl segment counterparts, using a straightforward synthetic approach that can be easily automated from commercially available building blocks, and applied it to a well-known problematic target, SUMO-2.

## Introduction

The advent of the native chemical ligation<sup>1</sup> (NCL) reaction has revolutionized the field of chemical protein synthesis by offering a simple strategy to assemble unprotected peptide segments bearing mutually-reactive C-terminal thioesters and N-terminal cysteines with exquisite chemo- and regioselectivity. Twenty-six years after its discovery, NCL is still the gold standard reaction in the field. The continuous development of many related synthetic methodologies allowed to extend and simplify its applicability,<sup>2</sup> and led to impressive applications in the total synthesis of functional proteins of more than 300 residues.<sup>3</sup>

One of the major current limitations of NCL-based protein synthesis is the low solubility or tendency to aggregate of some of the segments. If this is anticipated when synthesizing a very hydrophobic target such as a transmembrane protein, segments from soluble hydrophilic proteins frequently prove to be problematic, while being often very hard to predict. NCL is usually conducted under denaturing conditions, typically 6 M guanidinium chloride, and is tolerant to the addition of organic solvents<sup>4</sup> and detergents,<sup>5</sup> thus substantially minimizing these solubility/aggregation problems during the reaction itself. Nevertheless, the purification and characterization of the segments prior to NCL regularly remains a severe bottleneck. Many synthetic strategies were developed to overcome this critical problem. For example, modification of the segment through *N*-methylation of backbone amides<sup>6</sup> or fusion with a hydrophilic peptide,<sup>7</sup> can dramatically increase its solubility;

however, these modifications will permanently remain in the synthesized protein. Traceless approaches include the use of acid-labile *N*-2-hydroxy-4-methoxybenzyl (Hmb) groups on backbone amides,<sup>8</sup> or Ser/Thr *O*-acyl isopeptide<sup>3d,e,9</sup> known to inhibit aggregation, but the most widely used strategy is the incorporation of a temporary hydrophilic "solubilizing tag".<sup>10</sup>

Such temporary tags are generally composed of lysines or arginines that bear a cationic charge side chain, often as homo-oligomers, and can be introduced either on a backbone amide,<sup>11</sup> the C-terminus<sup>12</sup> of the segment or the side chain of an Asn,<sup>13</sup> Asp,<sup>13</sup> Cys,<sup>12c,14,15</sup> Gln,<sup>13,16</sup> Glu,<sup>13,17a</sup> Lys<sup>17</sup> or Thr<sup>16</sup> residues.

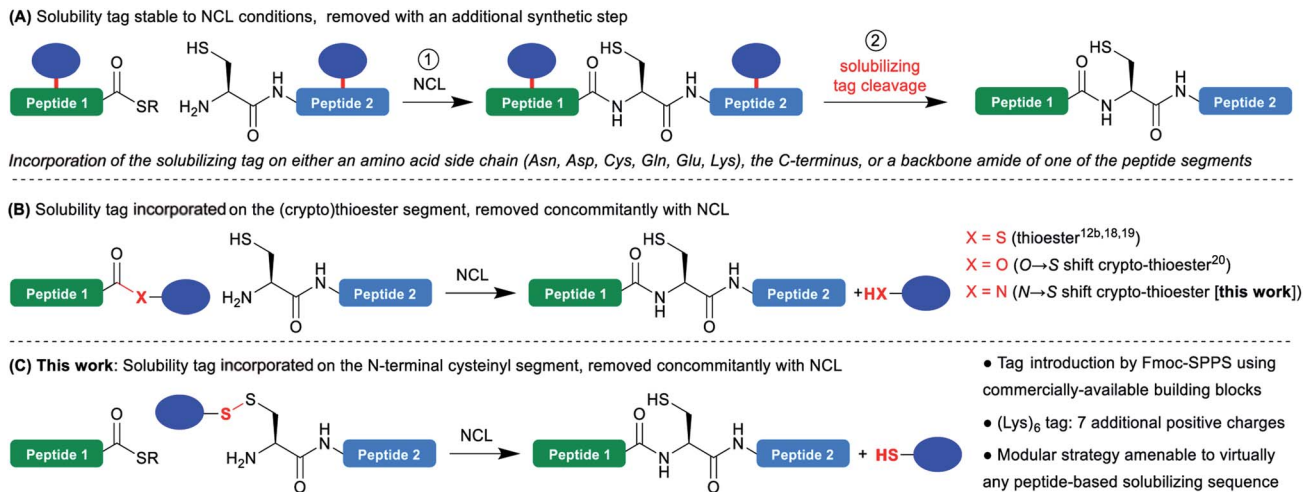
A large variety of linkers used to attach the tag to the peptide segment and designed to be cleaved in an additional step after the NCL (Scheme 1A) have been reported. Cleavage conditions include treatments with acids,<sup>11,14c,d</sup> bases,<sup>12b</sup> pH 4.5 buffer,<sup>17c</sup> sodium nitrite,<sup>12c,13</sup> nucleophiles,<sup>17b,d</sup> transition metal catalysis,<sup>12c,14a,b,15,17a</sup> UV irradiation<sup>16</sup> and autoproteolysis.<sup>12a</sup>

An advantageous alternative is the introduction of the solubilizing tag on the C-terminal thioester moiety (Scheme 1B).<sup>12b,18–20</sup> In this case, the tag is cleaved in the course of the NCL reaction without needing an additional synthetic step, after playing its solubilizing role during the purification, characterization and handling of the problematic segment. This approach was pioneered by Aimoto<sup>18</sup> and Kent<sup>19</sup> for Boc-SPPS-based thioester synthesis, and later extended by Tietze<sup>20</sup> to Fmoc-SPPS using a  $\beta$ -mercaptoester precursor converted *in situ* into a thioester during NCL through an *O*  $\rightarrow$  *S* acyl shift<sup>21</sup> (a so-called peptide crypto-thioester<sup>22</sup>). If this one-pot NCL/tag cleavage approach is clearly not suited to extreme situations where the ligation product remains insoluble or prone to aggregation, it is expected to be widely applicable in the many cases when the low-solubility/aggregation behavior is associated with a single isolated segment of the protein. Indeed, the additional segment ligated to

Centre de Biophysique Mol culaire, CNRS UPR 4301, Rue Charles Sadron, 45071 Orl ans Cedex 2, France. E-mail: aucagne@cnrs-orleans.fr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0sc06001a





**Scheme 1** Synthetic strategies for solubilization of peptide segments used in NCL-based protein synthesis using temporary solubility tags (blue ovals) that are removed either after (A) or concomitantly (B and C) with the NCL reaction.

the problematic one can further play the role of solubility tag, strikingly demonstrated by the synthesis of fragments of transmembrane proteins using these techniques.<sup>18,20</sup> However, the strategy is inherently limited to thioester segments and not suited for N-terminal cysteinyl counterparts.<sup>23</sup>

In a closely related context, Valiyaveetil<sup>24</sup> proposed the idea to introduce a solubility tag linked through a disulfide to the N-terminal Cys of a cysteinyl segment synthesized by Boc-SPPS, the disulfide bond being cleaved in a first additional reduction step prior to NCL. In this case, the tag (Arg–Arg–Arg–Cys–NH<sub>2</sub>) was introduced in solution through air oxidation-mediated formation of a mixed disulfide with the crude segment. This resulted in a very low yield in the tagged segment due to both insolubility of the non-tagged segment and the non-directed formation of the disulfide leading to complex mixtures. Nevertheless, the authors succeeded in the highly demanding semi-synthesis of ion channels through expressed protein ligation<sup>25</sup> (EPL) with a recombinant thioester, demonstrating the feasibility of the approach.

We thought that devising a straightforward methodology for the introduction of such a disulfide-linked solubilizing tag on N-terminal cysteinyl segments, and extending the concept of concomitant NCL/tag cleavage could be extremely valuable and generally applicable in chemical protein synthesis (Scheme 1C). Indeed, disulfides are readily cleaved under standard NCL conditions generally including reducing agents (*e.g.* TCEP) and a large excess of aromatic thiols like 4-mercaptophenylacetic acid (MPAA).<sup>26</sup>

We herein describe a straightforward methodology for the introduction of such a tag through Fmoc-SPPS, which can be easily automated on a standard peptide synthesizer, and exemplified the utility of this method through the synthesis of SUMO-2, a previously reported difficult target.<sup>12c</sup>

## Results and discussion

The disulfide-based linker bridging the N-terminal cysteinyl segment and the solubilizing tag is the cornerstone of the

strategy. Ideally, this linker should be (1) introduced through an automation-friendly solid phase step using commercially available materials, (2) bear a primary aliphatic amine or a suitable precursor for further Fmoc-SPPS elongation of a hydrophilic peptide sequence and (3) be stable to the Fmoc-SPPS conditions, including elongation and TFA-based cleavage.

Considering these requirements, we reasoned that incorporation of the N-terminal cysteinyl residue through the coupling of Boc-Cys(Npys)-OH would be ideal. The Npys group<sup>27</sup> (*S*-3-nitro-2-pyridinesulfonyl) is classically used for the directed formation of a mixed disulfide on an N-terminal cysteine either on solid support<sup>28</sup> or in solution, through reaction with a thiol. The simplest solution for automated synthesis would be a reaction on solid support with an amino-thiol. We selected two commercially available candidates: cysteamine (2-aminoethane-1-thiol, **1**), and 2-amino-1,1-dimethyl-ethane-1-thiol (**2**). Preliminary experiments with a model tripeptide (see ESI p. S5–S11†) showed that disulfide formation was quantitative through simple incubation in NMP for 1 h with an excess (10 equiv.) of **2** as its hydrochloride salt. We also demonstrated that the disulfide was stable to a long piperidine treatment mimicking the repeated Fmoc deprotection conditions needed for the SPPS elongation of a hydrophilic peptide tag. Contrastingly, reaction with **1** was much less clean, and the resulting disulfide was not stable to the piperidine treatment. These results are in accordance with the known higher stability of tertiary thiol disulfides derivatives of cysteine like *S*-*S*tBu towards Fmoc-SPPS conditions as compared to simple non-hindered primary thiol disulfides.<sup>29</sup>

Having in hand a robust method for the solid phase introduction of the linker (referred hereafter as Ades, 2-amino-1,1-dimethylethyl-1-sulfanyl), we applied it to two different long peptides, followed by the automated introduction of a (Lys)<sub>6</sub> hydrophilic tag through repeated couplings of Fmoc-Lys(Boc)-OH under standard conditions (Scheme 2). We started with a 41 amino acids (aa) model sequence devoid of any solubility problem, derived from the human mucin MUC1<sup>30</sup> variable



number tandem repeat (VNTR) region made of a duplicated 20 aa sequence (7, Scheme 3). Gratifyingly, 7 was obtained in excellent yield and purity without needing of any further optimization of the synthetic protocol.

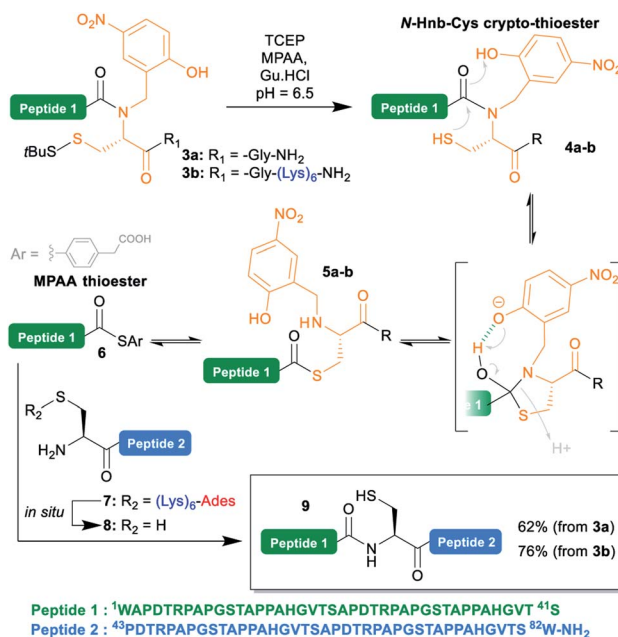
To test the concomitant NCL/tag cleavage, we synthesized the 41 aa crypto-thioester segment counterpart **3a**, also derived from MUC1. This segment is equipped with an *N*-(2-hydroxy-4-nitrobenzyl)cysteine-based device (*N*-Hnb-Cys) capable of forming thioester *in situ* under NCL conditions (Scheme 3).<sup>31</sup> *N*-Hnb-Cys crypto-thioesters are straightforward to synthesize through automated Fmoc-SPPS and show fast ligation kinetics owing to internal catalysis by a judiciously placed phenol group (Scheme 3).<sup>31a</sup>

To our delight, the ligation of **3a** with **7** proceeded very cleanly, the (Lys)<sub>6</sub>-Ades tag being cleaved within seconds under the NCL conditions, giving the target 82 aa polypeptide **9** in a 62% isolated yield (ESI p. S18–S19).

Additionally, we took the opportunity of this work to demonstrate the applicability of the C-terminal thioester solubilization strategy to *N*-Hnb-Cys crypto-thioesters. Quite expectedly, introduction of a (Lys)<sub>6</sub> tag was straightforward, and gave segment **3b** in good yields without needing any optimization. As for **3a**, NCL with **7** proceeded cleanly, and with comparable kinetics (ESI, Fig. S15†), giving **9** in an excellent 76% isolated yield (Scheme 3).

The cysteine residue in **9** was further desulfurized under classical conditions<sup>32</sup> to give an alanine (ESI, p S20–S21) such as in the native MUC1 sequence.

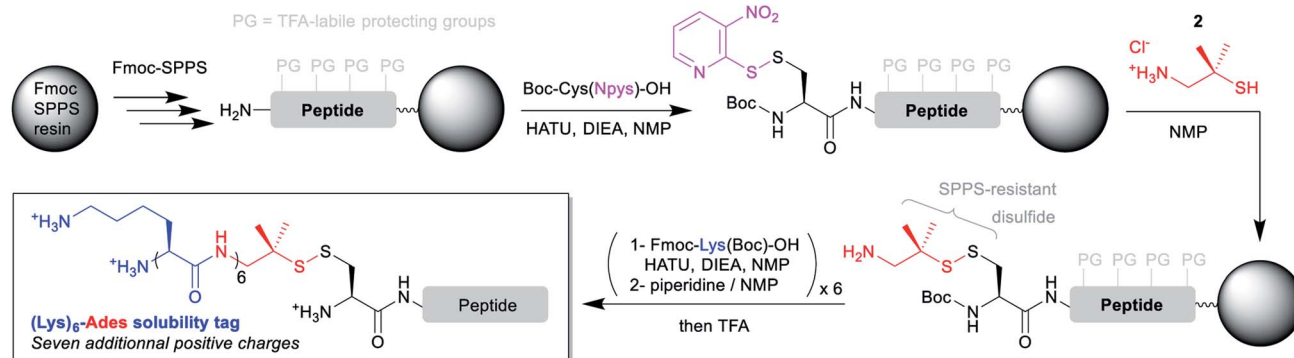
Encouraged by these results, we then implemented this approach for the synthesis of the 93 aa SUMO-2. Small ubiquitin-related modifiers<sup>33</sup> (SUMO) were first discovered in mammals in 1996.<sup>34</sup> To dated, five SUMO isoforms have been identified in humans, SUMO-1, 2, 3, 4 and 5. SUMOylation is a post-translational modification (PTM) consisting in the covalent attachment of the C-terminus of SUMO proteins *via* an isopeptide bond to specific lysine residues in target proteins. An enzymatic cascade controls the attachment, involving activator (E1), conjugating (E2), and sometimes ligase (E3) enzymes. This PTM is reversible, through deSUMOylation by sentrin/SUMO-specific proteases (SENPs).<sup>35</sup> In contrast with ubiquitin,<sup>36,37</sup>



Scheme 3 Native chemical ligation reaction showing the putative (Hnb)Cys self-catalyzed *N* → *S* acyl shift mechanism.

only few examples of the chemical synthesis of SUMO proteins, their dimers and conjugates, have been reported.<sup>5d,12c,38,39</sup> One illustration is the synthesis by Brik<sup>12c</sup> of SUMO-2-diubiquitin hybrid chains. In this work, the authors reported the low solubility, tendency to aggregation and unusual HPLC behavior of an N-terminal cysteinyl SUMO-2[46–93]<sup>40</sup> segment. To circumvent this issue, they developed an elegant C-terminal solubilizing tag in which a 3,4-diaminobenzoic acid (Dbz) linker<sup>41</sup> was employed to attach a poly-Arg tag to the C-terminus of this segment. We thought that it could be interesting to challenge our (Lys)<sub>6</sub>-Ades methodology with this benchmark target.

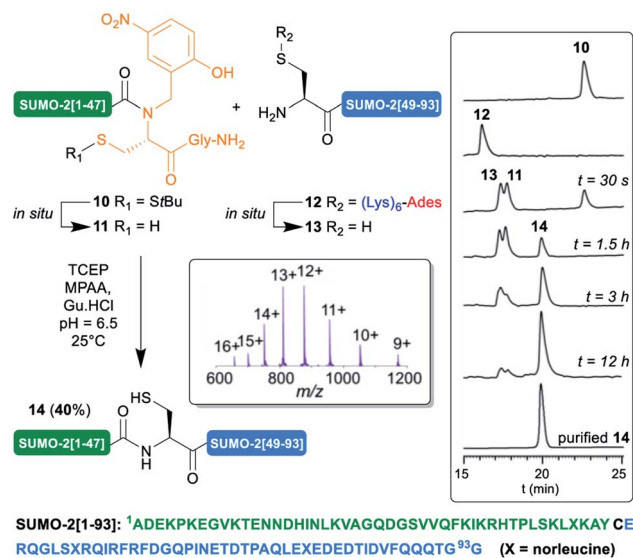
We applied our synthetic protocol to (Lys)<sub>6</sub>-Ades-SUMO-2 [48–93] (**12**),<sup>42</sup> which gave a clean and soluble crude product that was purified by standard HPLC. As anticipated from Brik's report, in sharp contrast with **12**, the non-tagged version exhibited anomalous HPLC behavior (ESI, S11, Fig. S11†) and



Scheme 2 General synthetic strategies for the introduction of (Lys)<sub>6</sub>-Ades solubility tags on N-terminal cysteinyl segments using automated Fmoc-SPPS.







Scheme 4 NCL-based SUMO-2 synthesis using the (Lys)<sub>6</sub>Ades methodology.

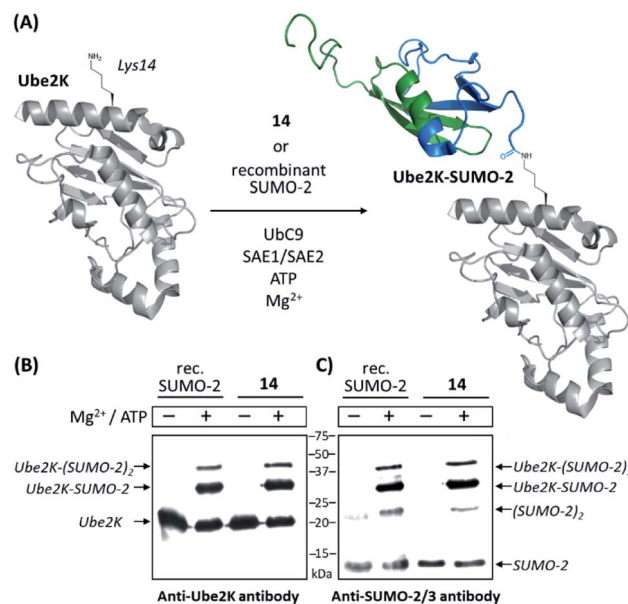


Fig. 1 *In vitro* SUMOylation of Ube2K with **14** and a recombinant SUMO-2 used as a control. (A) Schematic representation of the reaction (cartoons based on NMR and X-ray structures of SUMO-2 and Ube2K, PDB ID: 2BEP and 2N1W, respectively); (B) Western blot analysis using an anti-Ube2K antibody; (C) Western blot analysis using an anti-SUMO-2/3 antibody.

the crude peptide was barely soluble in HPLC solvents: 0.05 mg mL<sup>-1</sup> in 8 : 2 : 0.01 H<sub>2</sub>O/MeCN/TFA, thus further validating the choice of this target for this work. Gratifyingly, NCL of **12** with SUMO-2[1-47] *N*-Hnb-Cys cryptotioester **10** proceeded very cleanly as shown in the analytical HPLC monitoring<sup>43</sup> of the reaction (Scheme 4), giving the expected SUMO-2[1-93] **14** in a good 40% isolated yield. As observed with the MUC1-derived model segments, the HPLC chromatograms show nearly

instantaneous conversion of **12** into non-tagged **13**, concomitantly with the slightly slower conversion of *S*-StBu-protected dormant *N*-Hnb-Cys cryptotioester **10** into an active form **11**.<sup>44</sup>

Note that in accordance with Melnyk's findings,<sup>45</sup> 15% of a side product showing a molecular mass of [M - 18] *m/z* relative to the desired product **14** was observed when performing the reaction at 37 °C for a prolonged time and was attributed to aspartimide formation at one of the Asp-Gly sites. Lowering the temperature to 25 °C nearly abolished this side reaction (ESI p. S22-S25).

Finally, in order to further characterize the synthesized SUMO-2 **14** from a biochemical point of view, it was folded by simple solubilization into a neutral buffer as described.<sup>39b,f,46</sup> Its three-dimensional structure was evaluated by circular dichroism, showing a spectrum essentially identical with the previously reported ones<sup>46</sup> for recombinant and synthetic SUMO-2 (ESI Fig. S21†).

We then compared **14** to a commercially available recombinant version of SUMO-2 for its ability to act as a substrate of the SUMO E2 conjugating enzyme Ubc9 (ref. 47) and a SUMO E1 activating enzyme (namely a heterodimeric complex of SAE1 and SAE2 proteins)<sup>48</sup> for conjugation to the ubiquitin-conjugating enzyme Ube2K (also known as E2-25K and Huntington Interacting Protein 2). This target is reported to be among the best *in vitro* substrates for Ubc9-dependent SUMOylation known thus far, although its SUMOylated lysine residue (Lys14) is not surrounded by a consensus SUMOylation motif.<sup>49</sup> Using anti-SUMO-2/3 and anti-Ube2K antibodies, Western-blot analyses first showed that synthetic SUMO-2 is recognized as efficiently as the recombinant one. The appearance of an intense band recognized by both antibodies and migrating at the expected molecular mass of a SUMO-modified Ube2K demonstrated the successful SUMOylation of the acceptor protein in both cases. Weaker bands likely corresponding to di-SUMOylated Ube2K and di-SUMO-2 were also observed, consistent with the presence of a SUMOylable lysine in SUMO-2 (Lys11). As expected, no reactions occurred in the absence of ATP and Mg<sup>2+</sup>, required cofactors for the E1 enzyme. Altogether, these experiments demonstrate the fully functional nature of synthesized **14** (Fig. 1 and ESI p. S27-S29).

## Conclusions

Poor solubility and aggregation of peptide segments are main bottlenecks for the chemical synthesis of proteins using native chemical ligation. Numerous strategies for the solubilization of problematic segments through temporary modification have been developed such as the introduction of solubilizing tags, but often require complex synthetic strategies, in-house synthesis of building blocks or extra steps to generate the native protein sequence after the NCL-based assembly. In this work, we have introduced a straightforward methodology for the temporary solubilization of *N*-terminal cysteinyl segments, based on the introduction of an oligolysine tag through a disulfide linkage with the *N*-terminal cysteine residue. This (Lys)<sub>6</sub>-Ades tag is easily incorporated in the target segment through automated solid-phase synthesis using commercially



available building blocks, is stable during handling, purification and storage of the segment, while being cleaved within seconds under NCL conditions<sup>50</sup> to generate *in situ* the reactive free cysteine. We exemplified the broad potential of this method through the NCL-based synthesis of a model polypeptide derived from the human mucin MUC1, in addition to a well-known difficult small protein target, SUMO-2. Due to its overall simplicity and efficiency, we believe that this strategy will advantageously complement existing methodologies in the synthesis of other challenging proteins.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We warmly thank Jean-Baptiste Madinier for assistance in peptide synthesis and Dr Cyril Colas for HRMS analyses. The University of Orléans and Doctoral School SSBCV (ED549) are gratefully acknowledged for PhD fellowships for SAA and EC, and ANR (EasyMiniprot, ANR-15-CE07-0022) for financial support.

## Notes and references

- 1 P. Dawson, T. Muir, I. Clark-Lewis and S. Kent, Synthesis of proteins by native chemical ligation, *Science*, 1994, **266**, 776–779.
- 2 V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J.-C. M. Monbaliu and O. Melnyk, Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations, *Chem. Rev.*, 2019, **119**, 7328–7443.
- 3 (a) M. T. Weinstock, M. T. Jacobsen and M. S. Kay, Synthesis and folding of a mirror-image enzyme reveals ambidextrous chaperone activity, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 11679–11684; (b) A. Pech, J. Achenbach, M. Jahnz, S. Schülzchen, F. Jarosch, F. Bordusa and S. Klussmann, A thermostable D-polymerase for mirror-image PCR, *Nucleic Acids Res.*, 2017, **45**, 3997–4005; (c) J. Weidmann, M. Schnölzer, P. E. Dawson and J. D. Hoheisel, Copying Life: Synthesis of an Enzymatically Active Mirror-Image DNA-Ligase Made of D-Amino Acids, *Cell Chem. Biol.*, 2019, **26**, 645–651; (d) Z. Wang, W. Xu, L. Liu and T. F. Zhu, A synthetic molecular system capable of mirror-image genetic replication and transcription, *Nat. Chem.*, 2016, **8**, 698–704; (e) W. Xu, W. Jiang, J. Wang, L. Yu, J. Chen, X. Liu, L. Liu and T. F. Zhu, Total chemical synthesis of a thermostable enzyme capable of polymerase chain reaction, *Cell Discovery*, 2017, **3**, 17008; (f) S. Tang, L.-J. Liang, Y. -Y. Si, S. Gao, J. -X. Wang, J. Liang, Z. Mei, J. -S. Zheng and L. Liu, Practical Chemical Synthesis of Atypical Ubiquitin Chains by Using an Isopeptide-Linked Ub Isomer, *Angew. Chem., Int. Ed.*, 2017, **56**, 13333–13337; (g) H. Sun and A. Brik, The Journey for the Total Chemical Synthesis of a 53 kDa Protein, *Acc. Chem. Res.*, 2019, **52**, 3361–3371.
- 4 (a) G. G. Kochendoerfer, D. Salom, J. D. Lear, R. Wilk-Orescan, S. B. H. Kent and W. F. DeGrado, Total Chemical Synthesis of the Integral Membrane Protein Influenza A Virus M2: Role of Its C-Terminal Domain in Tetramer Assembly, *Biochemistry*, 1999, **38**, 11905–11913; (b) M. Dittmann, J. Sauermaun, R. Seidel, W. Zimmermann and M. Engelhard, Native chemical ligation of hydrophobic peptides in organic solvents, *J. Pept. Sci.*, 2010, **16**, 558–562; (c) F. Shen, S. Tang and L. Liu, Hexafluoro-2-propanol as a potent cosolvent for chemical ligation of membrane proteins, *Sci. China: Chem.*, 2011, **54**, 110–116; (d) M. Dittmann, M. Sadek, R. Seidel and M. Engelhard, Native chemical ligation in dimethylformamide can be performed chemoselectively without racemization: NCL in organic solvents, *J. Pept. Sci.*, 2012, **18**, 312–316.
- 5 (a) F. I. Valiyaveetil, R. MacKinnon and T. W. Muir, Semisynthesis and Folding of the Potassium Channel KcsA, *J. Am. Chem. Soc.*, 2002, **124**, 9113–9120; (b) S. Ingale, T. Buskas and G.-J. Boons, Synthesis of Glyco(lipo)peptides by Liposome-Mediated Native Chemical Ligation, *Org. Lett.*, 2006, **8**, 5785–5788; (c) S. Lahiri, M. Brehs, D. Olschewski and C. F. W. Becker, Total Chemical Synthesis of an Integral Membrane Enzyme: Diacylglycerol Kinase from *Escherichia coli*, *Angew. Chem., Int. Ed.*, 2011, **50**, 3988–3992; (d) J. Bouchenna, M. Sénéchal, H. Drobecq, J. Vicogne and O. Melnyk, Total Chemical Synthesis of All SUMO-2/3 Dimer Combinations, *Bioconjugate Chem.*, 2019, **30**, 2967–2973.
- 6 E. C. B. Johnson and S. B. H. Kent, Studies on the Insolubility of a Transmembrane Peptide from Signal Peptide Peptidase, *J. Am. Chem. Soc.*, 2006, **128**, 7140–7141.
- 7 E. Bianchi, R. Ingenito, R. J. Simon and A. Pessi, Engineering and Chemical Synthesis of a Transmembrane Protein: The HCV Protease Cofactor Protein NS4A, *J. Am. Chem. Soc.*, 1999, **121**, 7698–7699.
- 8 A. M. Levinson, J. H. McGee, A. G. Roberts, G. S. Creech, T. Wang, M. T. Peterson, R. C. Hendrickson, G. L. Verdine and S. J. Danishefsky, Total Chemical Synthesis and Folding of All- l and All- d Variants of Oncogenic KRas(G12V), *J. Am. Chem. Soc.*, 2017, **139**, 7632–7639.
- 9 (a) H. Hojo, H. Katayama, C. Tano, Y. Nakahara, A. Yoneshige, J. Matsuda, Y. Sohma, Y. Kiso and Y. Nakahara, Synthesis of the sphingolipid activator protein, saposin C, using an azido-protected O-acyl isopeptide as an aggregation-disrupting element, *Tetrahedron Lett.*, 2011, **52**, 635–639; (b) Y. Sohma, H. Kitamura, H. Kawashima, H. Hojo, M. Yamashita, K. Akaji and Y. Kiso, Synthesis of an O-acyl isopeptide by using native chemical ligation to efficiently construct a hydrophobic polypeptide, *Tetrahedron Lett.*, 2011, **52**, 7146–7148; (c) H. Kawashima, T. Kuruma, M. Yamashita, Y. Sohma and K. Akaji, Synthesis of an O-acyl isopeptide by using native chemical ligation in an aqueous solvent system, *J. Pept. Sci.*, 2014, **20**, 361–365; (d) R. Desmet, M. Pauzuolis, E. Boll, H. Drobecq, L. Raibaut and O. Melnyk, Synthesis of Unprotected Linear or Cyclic O-



- Acyl Isopeptides in Water Using Bis(2-sulfanylethyl)amido Peptide Ligation, *Org. Lett.*, 2015, **17**, 3354–3357.
- 10 D.-D. Zhao, X.-W. Fan, H. Hao, H.-L. Zhang and Y. Guo, Temporary Solubilizing Tags Method for the Chemical Synthesis of Hydrophobic Proteins, *Curr. Org. Chem.*, 2019, **23**, 2–13.
- 11 (a) J.-S. Zheng, M. Yu, Y.-K. Qi, S. Tang, F. Shen, Z.-P. Wang, L. Xiao, L. Zhang, C.-L. Tian and L. Liu, Expedient Total Synthesis of Small to Medium-Sized Membrane Proteins via Fmoc Chemistry, *J. Am. Chem. Soc.*, 2014, **136**, 3695–3704; (b) J.-S. Zheng, Y. He, C. Zuo, X.-Y. Cai, S. Tang, Z. A. Wang, L.-H. Zhang, C.-L. Tian and L. Liu, Robust Chemical Synthesis of Membrane Proteins through a General Method of Removable Backbone Modification, *J. Am. Chem. Soc.*, 2016, **138**, 3553–3561; (c) S. Tang, C. Zuo, D.-L. Huang, X.-L. Cai, L.-H. Zhang, C.-L. Tian, J.-S. Zheng and L. Liu, Chemical synthesis of membrane proteins by the removable backbone modification method, *Nat. Protoc.*, 2017, **12**, 2554–2569.
- 12 (a) E. C. B. Johnson, E. Malito, Y. Shen, D. Rich, W.-J. Tang and S. B. H. Kent, Modular Total Chemical Synthesis of a Human Immunodeficiency Virus Type 1 Protease, *J. Am. Chem. Soc.*, 2007, **129**, 11480–11490; (b) P. W. R. Harris and M. A. Brimble, Toward the total chemical synthesis of the cancer protein NY-ESO-1, *Biopolymers*, 2010, **94**, 542–550; (c) S. Bondalapati, E. Eid, S. M. Mali, C. Wolberger and A. Brik, Total chemical synthesis of SUMO-2-Lys63-linked diubiquitin hybrid chains assisted by removable solubilizing tags, *Chem. Sci.*, 2017, **8**, 4027–4034.
- 13 S. Tsuda, S. Masuda and T. Yoshiya, Solubilizing Trityl-Type Tag To Synthesize Asx/Glx-Containing Peptides, *ChemBioChem*, 2019, **20**, 2063–2069.
- 14 (a) S. K. Maity, G. Mann, M. Jbara, S. Laps, G. Kamnesky and A. Brik, Palladium-Assisted Removal of a Solubilizing Tag from a Cys Side Chain To Facilitate Peptide and Protein Synthesis, *Org. Lett.*, 2016, **18**, 3026–3029; (b) J. A. Brailsford, J. L. Stockdill, A. J. Axelrod, M. T. Peterson, P. A. Vadola, E. V. Johnston and S. J. Danishefsky, Total chemical synthesis of human thyroid-stimulating hormone (hTSH)  $\beta$ -subunit: application of arginine-tagged acetamidomethyl (AcmR) protecting groups, *Tetrahedron*, 2018, **74**, 1951–1956; (c) S. Tsuda, M. Mochizuki, H. Ishiba, K. Yoshizawa-Kumagaye, H. Nishio, S. Oishi and T. Yoshiya, Easy-to-Attach/Detach Solubilizing-Tag-Aided Chemical Synthesis of an Aggregative Capsid Protein, *Angew. Chem., Int. Ed.*, 2018, **57**, 2105–2109; (d) S. Tsuda, S. Masuda and T. Yoshiya, The versatile use of solubilizing trityl tags for difficult peptide/protein synthesis, *Org. Biomol. Chem.*, 2019, **17**, 1202–1205.
- 15 For a recent example of the incorporation of a tertiary amine as solubilizing tag, see: E. Eid, G. N. Boross, H. Sun, M. Msallam, S. K. Singh and A. Brik, Total Chemical Synthesis of ISGylated-Ubiquitin Hybrid Chain Assisted by Acetamidomethyl Derivatives with Dual Functions, *Bioconjugate Chem.*, 2020, **31**, 889–894.
- 16 Y.-C. Huang, Y.-M. Li, Y. Chen, M. Pan, Y.-T. Li, L. Yu, Q.-X. Guo and L. Liu, Synthesis of Autophagosomal Marker Protein LC3-II under Detergent-Free Conditions, *Angew. Chem., Int. Ed.*, 2013, **52**, 4858–4862.
- 17 (a) Z. Tan, S. Shang and S. J. Danishefsky, Rational development of a strategy for modifying the aggregability of proteins, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4297–4302; (b) M. T. Jacobsen, M. E. Petersen, X. Ye, M. Galibert, G. H. Lorimer, V. Aucagne and M. S. Kay, A Helping Hand to Overcome Solubility Challenges in Chemical Protein Synthesis, *J. Am. Chem. Soc.*, 2016, **138**, 11775–11782; (c) S. Tsuda, H. Nishio and T. Yoshiya, Peptide self-cleavage at a canaline residue: application to a solubilizing tag system for native chemical ligation, *Chem. Commun.*, 2018, **54**, 8861–8864; (d) J. M. Fulcher, M. E. Petersen, R. J. Giesler, Z. S. Cruz, D. M. Eckert, J. N. Francis, E. M. Kawamoto, M. T. Jacobsen and M. J. Kay, Chemical Synthesis of Shiga Toxin Subunit B Using a Next-Generation Traceless “Helping Hand” Solubilizing Tag, *Org. Biomol. Chem.*, 2019, **17**, 10237–10244.
- 18 T. Sato, Y. Saito and S. Aimoto, Synthesis of the C-terminal region of opioid receptor like 1 in an SDS micelle by the native chemical ligation: effect of thiol additive and SDS concentration on ligation efficiency, *J. Pept. Sci.*, 2005, **11**, 410–416.
- 19 E. C. B. Johnson and S. B. H. Kent, Towards the total chemical synthesis of integral membrane proteins: a general method for the synthesis of hydrophobic peptide- $\alpha$ thioester building blocks, *Tetrahedron Lett.*, 2007, **48**, 1795–1799.
- 20 A. C. Baumruck, D. Tietze, L. K. Steinacker and A. A. Tietze, Chemical synthesis of membrane proteins: a model study on the influenza virus B proton channel, *Chem. Sci.*, 2018, **9**, 2365–2375.
- 21 P. Botti, M. Villain, S. Manganiello and H. Gaertner, Native Chemical Ligation through in Situ O to S Acyl Shift, *Org. Lett.*, 2004, **6**, 4861–4864.
- 22 K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto and A. Otaka, N-Sulfanylethylanilide peptide as a crypto-thioester peptide, *ChemBioChem*, 2011, **12**, 1840–1844.
- 23 Note that, besides cationic solubilization tags, concomitant rearrangement of Ser/Thr O-acyl isopeptides during NCL has been exploited, taking advantage of the aggregation-inhibiting properties of the depsipeptide linkage during HPLC purification of segments: see for example ref. 3d, e and 9a.
- 24 K. M. Linn, M. G. Derebe, Y. Jiang and F. I. Valiyaveetil, Semisynthesis of NaK, a Na<sup>+</sup> and K<sup>+</sup> Conducting Ion Channel, *Biochemistry*, 2010, **49**, 4450–4456.
- 25 T. W. Muir, D. Sondhi and P. A. Cole, Expressed protein ligation: a general method for protein engineering, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705–6710.
- 26 E. C. B. Johnson and S. B. H. Kent, Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction, *J. Am. Chem. Soc.*, 2006, **128**, 6640–6646.
- 27 R. Matsueda, S. Higashida, R. J. Ridge and G. R. Matsueda, Activation of conventional S-protecting groups of cysteine





- by conversion into the 3-nitro-2-pyridinesulfonyl (Npys) group, *Chem. Lett.*, 1982, **11**, 921–924.
- 28 W. Tegge, W. Bautsch and R. Frank, Synthesis of cyclic peptides and peptide libraries on a new disulfide linker, *J. Pept. Sci.*, 2007, **13**, 693–699.
- 29 (a) C. Seidel, E. Klauschenz, K. D. Kaufmann and H. Kunzek, Selektiv S-geschützte Cysteinpeptide. Zur Synthese von Cysteinpeptiden unter Verwendung der S-Äthylmercapto-Schutzgruppe, *J. Prakt. Chem.*, 1976, **318**, 229–247; (b) D. Andreu, F. Albericio, N. A. Solé, M. C. Munson, M. Ferrer and G. Barany, Formation of Disulfide Bonds in Synthetic Peptides and Proteins, in *Peptide Synthesis Protocols. Methods in Molecular Biology*, ed. M. W. Pennington and B. M. Dunn, Humana Press, Totowa, NJ, 1994, vol. 35.
- 30 For other examples of use of MUC1 VNTRs as models in protein chemical synthesis, see: (a) V. Aucagne, I. E. Valverde, P. Marceau, M. Galibert, N. Dendane and A. F. Delmas, Towards the Simplification of Protein Synthesis: Iterative Solid-Supported Ligations with Concomitant Purifications, *Angew. Chem., Int. Ed.*, 2012, **51**, 11320–11324; (b) I. E. Decostaire, D. Lelièvre, V. Aucagne and A. F. Delmas, Solid phase oxime ligations for the iterative synthesis of polypeptide conjugates, *Org. Biomol. Chem.*, 2014, **12**, 5536–5543; (c) M. Galibert, V. Piller, F. Piller, V. Aucagne and A. F. Delmas, Combining triazole ligation and enzymatic glycosylation on solid phase simplifies the synthesis of very long glycoprotein analogues, *Chem. Sci.*, 2015, **6**, 3617–3623; (d) S. F. Loibl, Z. Harpaz, R. Zitterbart and O. Seitz, Total chemical synthesis of proteins without HPLC purification, *Chem. Sci.*, 2016, **7**, 6753–6759.
- 31 (a) V. P. Terrier, H. Adihou, M. Arnould, A. F. Delmas and V. Aucagne, A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide cryptothioesters, *Chem. Sci.*, 2016, **7**, 339–345; (b) D. Lelièvre, V. P. Terrier, A. F. Delmas and V. Aucagne, Native Chemical Ligation Strategy to Overcome Side Reactions during Fmoc-Based Synthesis of C-Terminal Cysteine-Containing Peptides, *Org. Lett.*, 2016, **18**, 920–923; (c) G. Martinez, J.-P. Hograindleur, S. Voisin, R. Abi Nahed, T. M. Abd El Aziz, J. Escoffier, J. Bessonnat, C.-M. Fovet, M. De Waard, S. Hennebicq, V. Aucagne, P. F. Ray, E. Schmitt, P. Bulet and C. Arnould, Spermaurin, an La1-like peptide from the venom of the scorpion *Scorpio maurus palmatus*, improves sperm motility and fertilization in different mammalian species, *Mol. Hum. Reprod.*, 2016, **10**, 116–131; (d) V. P. Terrier, A. F. Delmas and V. Aucagne, Efficient synthesis of cysteine-rich cyclic peptides through intramolecular native chemical ligation of N-Hnb-Cys peptide cryptothioesters, *Org. Biomol. Chem.*, 2017, **15**, 316–319; (e) K. Loth, A. Vergnes, C. Barreto, S. N. Voisin, H. Meudal, J. Da Silva, A. Bressan, N. Belmadi, E. Bachère, V. Aucagne, C. Cazevielle, H. Marchandin, R. D. Rosa, P. Bulet, L. Touqui, A. F. Delmas and D. Destoumieux-Garzón, The Ancestral N-Terminal Domain of Big Defensins Drives Bacterially Triggered Assembly into Antimicrobial Nanonets, *mBio*, 2019, **10**, e01821-19; (f) L. De Rosa, R. Di Stasi and L. D. D'Andrea, *Tetrahedron*, 2019, **75**, 894–905; (g) S. A. Abboud and V. Aucagne, An optimized protocol for the synthesis of N-2-hydroxybenzyl-cysteine peptide cryptothioesters, *Org. Biomol. Chem.*, 2020, **18**, 8199–8208.
- 32 Q. Wan and S. J. Danishefsky, Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248–9252.
- 33 Z.-J. Han, Y.-H. Feng, B.-H. Gu, Y.-M. Li and H. Chen, The post-translational modification, SUMOylation, and cancer, *Int. J. Oncol.*, 2018, **52**, 1081–1094.
- 34 M. J. Matunis, E. Coutavas and G. Blobel, A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex, *J. Cell Biol.*, 1996, **135**, 1457–1470.
- 35 C. Guo and J. M. Henley, Wrestling with stress: Roles of protein SUMOylation and deSUMOylation in cell stress response: SUMOylation in Cell Stress, *IUBMB Life*, 2014, **66**, 71–77.
- 36 G. Goldstein, M. Scheid, U. Hammerling, D. H. Schlesinger, H. D. Niall and E. A. Boyse, Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 11–15.
- 37 For recent reviews on the chemical synthesis of ubiquitin derivatives, see for example: (a) M. P. Mulder, K. F. Witting and H. Ovaas, Cracking the Ubiquitin Code: The Ubiquitin Toolbox, *Curr. Issues Mol. Biol.*, 2020, **37**, 1–20; (b) R. Meledin, S. Mali and A. Brik, Pushing the Boundaries of Chemical Protein Synthesis: The Case of Ubiquitin Chains and Polyubiquitinated Peptides and Proteins, *Chem. Rec.*, 2016, **16**, 509–519.
- 38 For a review on the chemical synthesis of SUMO derivatives, see: O. Melnyk and J. Vicogne, Total chemical synthesis of SUMO proteins, *Tetrahedron*, 2016, **57**, 4319–4324.
- 39 (a) E. Boll, H. Drobecq, N. Ollivier, L. Raibaut, R. Desmet, J. Vicogne and O. Melnyk, A novel PEG-based solid support enables the synthesis of >50 amino-acid peptide thioesters and the total synthesis of a functional SUMO-1 peptide conjugate, *Chem. Sci.*, 2014, **5**, 2017–2022; (b) T. G. Wucherpfennig, V. R. Pattabiraman, F. R. P. Limberg, J. Ruiz-Rodríguez and J. W. Bode, Traceless Preparation of C-Terminal  $\alpha$ -Ketoacids for Chemical Protein Synthesis by  $\alpha$ -Ketoacid-Hydroxylamine Ligation: Synthesis of SUMO2/3, *Angew. Chem., Int. Ed.*, 2014, **53**, 12248–12252; (c) E. Boll, H. Drobecq, N. Ollivier, A. Blanpain, L. Raibaut, R. Desmet, J. Vicogne and O. Melnyk, One-pot chemical synthesis of small ubiquitin-like modifier protein-peptide conjugates using bis (2-sulfanylethyl)amido peptide latent thioester surrogates, *Nat. Protoc.*, 2015, **10**, 269–292; (d) H. Drobecq, E. Boll, M. Sénéchal, R. Desmet, J.-M. Saliou, J.-J. Lacapère, A. Mougél, J. Vicogne and O. Melnyk, A Central Cysteine Residue Is Essential for the Thermal Stability and Function of SUMO-1 Protein and SUMO-1



- Peptide–Protein Conjugates, *Bioconjugate Chem.*, 2016, **27**, 1540–1546; (e) M. P. C. Mulder, R. Merckx, K. F. Witting, D. S. Hameed, D. El Atmioui, L. Lelieveld, F. Liebelt, J. Neeffes, I. Berlin, A. C. O. Vertegaal and H. Ovaa, Total Chemical Synthesis of SUMO and SUMO-Based Probes for Profiling the Activity of SUMO-Specific Proteases, *Angew. Chem., Int. Ed.*, 2018, **57**, 8958–8962; (f) J. Bouchenna, M. Sénéchal, H. Drobecq, N. Stankovic-Valentin, J. Vicogne and O. Melnyk, The Role of the Conserved SUMO-2/3 Cysteine Residue on Domain Structure Investigated Using Protein Chemical Synthesis, *Bioconjugate Chem.*, 2019, **30**, 2684–2696; (g) S. Baldauf, A. O. Ogunkoya, G. N. Boross and J. W. Bode, Aspartic Acid Forming  $\alpha$ -Ketoacid–Hydroxylamine (KAHA) Ligations with (*S*)-4,4-Difluoro-5-oxaproline, *J. Org. Chem.*, 2020, **85**, 1352–1364.
- 40 Note that for practical reasons, in this case the native cysteine Cys49 was mutated to a serine, and that ligation was performed at an Ala48  $\rightarrow$  Cys site which was desulfurized after the NCL to restore the native alanine residue.
- 41 J. B. Blanco-Canosa and P. E. Dawson, An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation, *Angew. Chem.*, 2008, **120**, 6957–6961.
- 42 SUMO-2 methionines residues (Met1, Met44, Met55 and Met78) were substituted by isosteric norleucines in **10** and **12**, following the common use in chemical protein synthesis to avoid air-mediated oxidation of the side chain sulfides into sulfoxides.
- 43 Note that the HPLC monitoring of the NCL reaction was delicate, due to the *in situ* formation of non-tagged SUMO-2[48–93] **S11** that required extensive washes of the column between analyses.
- 44 *S*-tert-Butylsulfanyl (*S*-*StBu*) is used to afford the cryptothioester in a dormant state, unable to rearrange into a reactive thioester during handling, purification and storage. Under classical NCL conditions (including TCEP and MPAA), *StBu* is readily cleaved within a few minutes affording active *N*-Hnb-Cys cryptothioester species.
- 45 J. Bouchenna, M. Sénéchal, H. Drobecq, J. Vicogne and O. Melnyk, The Problem of Aspartimide Formation During Protein Chemical Synthesis Using SEA-Mediated Ligation, in *Peptide and Protein Engineering*, ed. O. Iranzo and A. Roque, Springer Protocols Handbooks, Humana, New York, NY, 2019, DOI: 10.1007/978-1-0716-0720-6\_2.
- 46 M. Sénéchal, J. Bouchenna, J. Vicogne and O. Melnyk, Comment on "N-terminal Protein Tail Acts as Aggregation Protective Entropic Bristles: The SUMO Case", *Biomacromolecules*, 2020, **21**, 3480–3482.
- 47 L. M. Lois and C. D. Lima, Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1, *EMBO J.*, 2005, **24**, 439–451.
- 48 B. A. Schulman and J. W. Harper, Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signaling pathways, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 319–331.
- 49 A. Pichler, P. Knipscheer, E. Oberhofer, W. J. van Dijk, R. Körner, J. V. Olsen, S. Jentsch, F. Melchior and T. K. Sixma, SUMO modification of the ubiquitin-conjugating enzyme E2-25K, *Nat. Struct. Mol. Biol.*, 2005, **12**, 264–269.
- 50 Note that we only tested the (Lys)<sub>6</sub>-Ades solubilizing tag under the most classically used conditions for NCL, including a large excess of TCEP and MPAA. In rare cases, NCL reactions are performed under non-reductive conditions, either making use of pre-formed arylthioesters (such as MPAA thioesters), or using azaheterocycles as catalysts instead of an external arylthiol: the strategy herein described would expectedly not be applicable.

