

A straightforward method to prevent the under-estimated problem of aspartimide formation during chemical ligation-mediated protein synthesis.

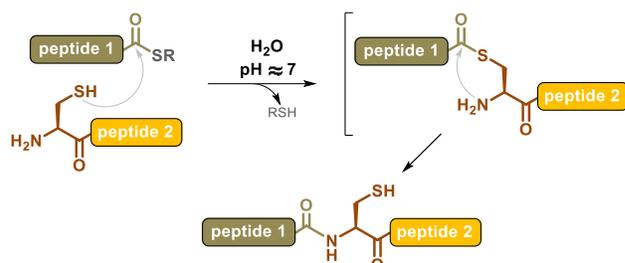
El hadji Cisse and Vincent Aucagne*

Centre de Biophysique Moléculaire, CNRS UPR 4301, 45071 Orléans, France

Email : vincent.aucagne@cnrs-orleans.fr

INTRODUCTION

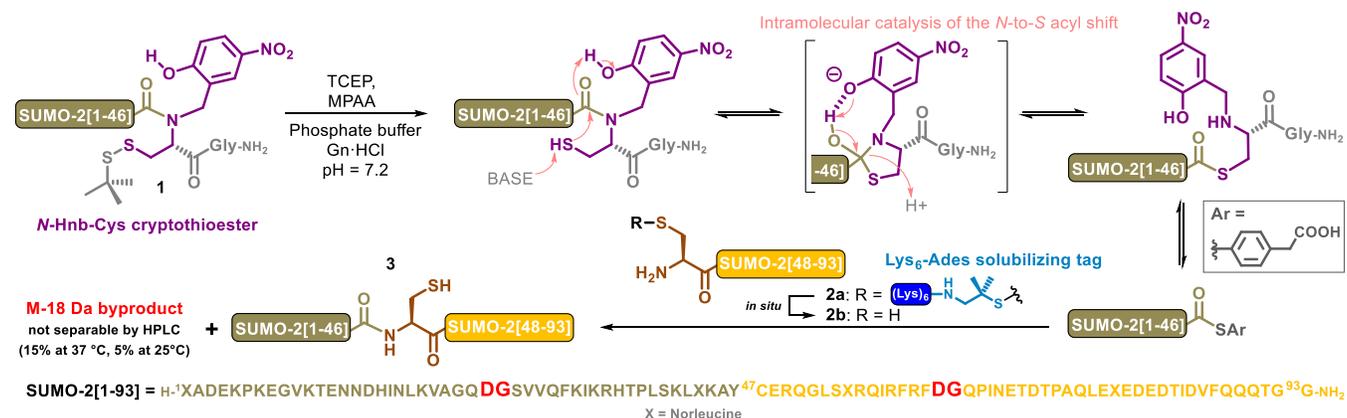
Chemical synthesis offers an effective alternative to access proteins that are difficult or impossible to produce through biotechnological means, such as ones with post-translational modifications (PTM) or “mirror-image” D-proteins, respectively. The invention of the chemical ligation concept¹ in the early 1990s revolutionized this field, introducing proteins into the realm of synthetic organic chemistry. The chemoselective coupling of unprotected peptide segments equipped with mutually reactive groups largely outperformed the limits of solid phase peptide synthesis² (SPPS), increasing the length of polypeptide chains that can be assembled in a controlled fashion from a few dozen amino acids to several hundred, and ultimately leading to correctly folded, high-purity proteins. The key breakthrough in the field was the discovery in 1994, by Kent and co-workers, of native chemical ligation³ (NCL) reaction, consisting of the coupling of a peptide C-terminal thioester segment with a second segment bearing an N-terminal cysteine. NCL is a two-step process based on a reversible trans-thioesterification followed by an intramolecular S-to-N-acyl shift (scheme 1), and can be carried out under dilute and mild aqueous conditions, compatible with a large variety of protein segments. Despite other amide-forming ligation reactions⁴⁻⁷ being actively pursued, NCL remains, almost thirty years after its discovery, the most commonly used reaction, and over a thousand different proteins have been obtained



Scheme 1. Native Chemical Ligation (NCL).

through this method.^{8,9} Recent impressive achievements include the synthesis of a 53 kDa tetra-ubiquitinated α -myoglobin¹⁰ composed of 472 aminoacids (AA), or a 775 AA mirror-image split *Pfu* DNA-polymerase,¹¹ through the ligation of numerous segments.

Not surprisingly, the development of methodologies aimed to broaden the applicability of NCL is an extremely active field of research. Novel strategies include successive ligations, such as one-pot¹² or solid-supported approaches,¹³ methods aimed at simplifying the access to C-terminal thioesters through the widely adopted Fmoc protecting group-based SPPS,^{8b,14} incorporation of temporary solubilizing tags for handling insoluble or aggregation-prone segments,¹⁵ or strategies aimed at overcoming the need for a cysteine residue at a retrosynthetic ligation disconnection.¹⁶ Addressing side reactions occurring during



Scheme 2. Previous work: synthesis of SUMO-2 (3) by NCL, showing the formation of a byproduct with a 18 Da mass loss.

NCL is another key point. Side reactions leading to the loss of reactivity of a segment, for example thioester hydrolysis, can be overcome through application of appropriate purification protocols,¹⁷ while side reactions involving the side chain of one AA residue can prove particularly problematic and greatly compromise the purity of the chemically-synthesized protein. A typical example is oxidation of the methionine side chain, which is often ameliorated through its preventive replacement by a bioisosteric norleucine residue.¹⁸ Another is epimerization at the C-terminal residue, as recently pointed out by Melnyk, Agouridas and co-workers for serine thioesters,¹⁹ which could be an overlooked problem considering the difficulty to discriminate between a single D- or L-AA within a long polypeptide chain through standard HPLC and MS-based analytical methods.

In this context, the possible formation of an aspartimide (Asi) five-membered ring, from an aspartate (Asp)²⁰ or asparagine (Asn)²¹ residue during NCL drew our attention. Base-mediated Asi formation through the attack of a backbone amide to the ester-protected side chain of an Asp has been one of the most prominent long-standing problems in Fmoc-SPPS.^{22,23} It represents a typical example of “sequence-dependant” side reactions, likely to occur in Asp-Gly motif-containing peptides. A large variety of methodologies are now available to circumvent this SPPS side reaction, either based on backbone amide or side-chain protecting group strategies, or alternative conditions for the piperidine-mediated Fmoc group deprotection. Spontaneous Asi formation at physiological pH from either an Asp or Asn residue is involved in the ageing of proteins *in vivo*²⁴ and their stability *in vitro*.²⁵ Additionally, aspartimide have been identified more recently as an evolution-selected PTM.²⁶

Reports of aspartimide formation in the course a chemical ligation-mediated protein assembly have been extremely scarce. An early 2003 work on the NCL-based synthesis of the 166 AA protein H-RAS did mention a compound showing a slightly lower mass than expected ($M-18 \pm 2$ Da) as the sole product.²⁷ This could be attributed to aspartimide formation from one Asp (M-18, loss of a water molecule) or Asn (M-17, loss of ammonia) residue. Asi formation during an NCL reaction has only been suspected four other times to date.^{20,21} In only one case additional evidence, other than mass spectrometry, supported the aspartimide hypothesis: mutation of the Asp residue into a Glu abolished the M-18 byproduct observed during the NCL-mediated cyclization of a small peptide, sunflower trypsin inhibitor-1.^{20a}

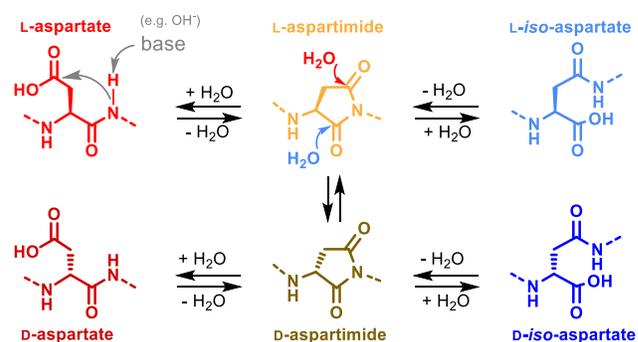
Strikingly, two of the above-cited reports, by Melnyk's lab^{20b,c} and ourselves,^{20d} concern the same protein, human SUMO-2, obtained in both cases through the ligation of two segments with an identical retrosynthetic disconnection at the sole cysteine residue (Cys47). Intrigued by this peculiarity, we initially mistakenly attributed it to a particularly Asi-prone sequence in SUMO-2. The present report details a systematic evaluation on the propensity of Asi formation under NCL conditions: an overlooked and probably frequent side reaction whose occurrence remains uncharted because it forms related byproducts that are not distinguishable from the target synthetic proteins by means of standard analytical methods. We propose a potentially generally-applicable methodology dedicated to completely abolish the formation of Asi and related byproducts during

NCL or other ligation reactions, and apply it to the aspartimide-free synthesis of SUMO-2 and derivatives.

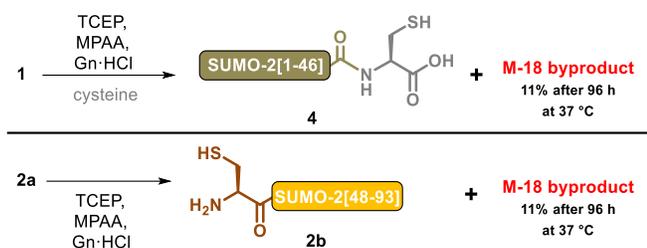
RESULTS AND DISCUSSION

SUMO-2. Discovered in 1995,²⁸ the Small Ubiquitin-like Modifier (SUMO) protein family belong to the ubiquitin-like category,²⁹ a class of proteins that exhibit weak sequence similarity but have a three-dimensional structure strongly analogous to ubiquitin. To date five different paralogs, referred to as SUMO-1 to 5, have been identified in humans.³⁰ SUMOylation is a PTM in which the C-terminus of the SUMO is attached to the lysine side chain of a target protein *via* an isopeptide bond. Dysregulation of SUMOylation is associated with many diseases.^{30,31} However, the physiological and pathological roles of SUMOylation remains to be fully elucidated, and the development of dedicated chemical biological tools is essential to improve our understanding of the underlying mechanisms.

Early hints on aspartimide (Asi) formation during previous chemical synthesis of SUMO-2. In our previous study, we synthesized SUMO-2 through the ligation of segments **1** and **2a** (Scheme 2).^{20d} The SUMO-2[1-46] segment **1** was equipped with the crypto-thioester device *N*-(2-hydroxy-4-nitrobenzyl)cysteine (*N*-Hnb-Cys), which has the ability to form a thioester *in situ* under NCL condition. This occurs through a spontaneous *N*-to-*S* acyl shift which is considerably accelerated by an intramolecular catalysis mediated by the phenol group.^{32,33} By itself, the SUMO-2[47-93] segment **2b** is poorly soluble and shows a propensity for aggregation and an unusual HPLC behaviour,³⁴ hampering its handling and purification. To solve this problem, we developed a dedicated strategy based on the introduction of a temporary solubilizing tag. The (Lys)₆-Ades (Ades = 2-amino-1,1-dimethyl-1-sulfanyl) is linked to the N-terminal cysteine by a disulfide linkage. The resulting segment **2a** incorporates six additional hydrophilic lysine residues, which dramatically enhanced its aqueous solubility. NCL is usually conducted under reducing conditions to prevent formation of disulfides, typically using large excesses of *tris*-carboxyethylphosphine (TCEP) and 4-mercaptophenylacetic acid (MPAA),³⁵ which also acts as a catalyst by forming *in situ* a highly reactive aryl thioester through a *trans*-thioesterification. Under such conditions, the (Lys)₆-Ades tag of **2a** is cleaved within few minutes, yielding an NCL-reactive segment (**2b**) which remains soluble: the ligation buffer contains 6 M guanidine hydrochloride (Gn-HCl), a denaturing medium extensively used for NCL to improve peptide solubility and avoid the partial folding of segments that could



Scheme 3. Aspartimide and related byproducts.



Scheme 4. Individual incubation of the two SUMO-2 segments under standard NCL condition: 100 mM MPAA, 50 mM TCEP, 6 M Gu-HCl, 200 mM phosphate buffer, pH = 7.2.

bury the reactive sites. Overnight incubation at 37 °C led to the obtainment of the desired protein 3. However, we also observed the formation of $\approx 15\%$ of a byproduct showing an 18 Da mass decrease that could not be separated by HPLC, and could be attributed to aspartimide formation from an Asp residue. This was consistent with Melnyk's observations when synthesizing SUMO-2 using a *bis*-sulfanylethyl (SEA) *N*-to-*S* shift-based crypto-thioester,³⁶ suggesting that formation of this byproduct is inherent to the target protein, and not to the method used to perform the NCL. In accordance with Melnyk's findings,^{20b,c} the amount of M-18 byproduct could be decreased to $\approx 5\%$ by conducting the reaction at room temperature for 12 h.

Characterization and rationalization of Asi byproducts on SUMO-2 -based models. In the present work, we sought to characterize the putative aspartimide byproduct and to decipher the reasons for its formation. Besides environmental effects^{25a,37} (buffer, temperature etc.), numerous studies have highlighted the strong impact of the intrinsic properties of the peptide or protein. The local conformation, and the three-dimensional structure of the protein have a substantial influence on its propensity for Asi formation.³⁸ The sequence surrounding the Asn/Asp is recognized to play a prominent role, and the Asp-Gly (DG) and Asn-Gly (NG) dipeptide motifs represent, by far, the main hotspots.^{25,39} The occurrence of DG and NG motifs has been estimated to an average 1.2 and 1.5 units per protein, respectively.^{8b} Given that over 1000 different proteins have been synthesized using NCL,^{8b,9} it is reasonable to assume that several hundred amongst them contain these motifs. However, to the best of our knowledge, only 5 articles have mentioned or suspected aspartimide formation during ligation reactions. This could mean that only a limited number of very specific peptide sequences are likely to form aspartimides in the context of NCL. Considering that, among these five examples, two articles concern human SUMO-2,^{20c,d} it seemed likely that this protein bears such a sequence with an exacerbated propensity for aspartimide formation. SUMO-2 contains a total of 7 aspartate residues, including two in a DG motif (Asp27 and Asp64). We hypothesized that one of these two residues is likely to be involved, in either the context of the individual segments or in the full-length polypeptide chain.

We first investigated whether SUMO-2 had an intrinsic capacity to form an aspartimide, even under non-denaturing conditions. We were surprised to see a decrease in the amount of the M-18 byproduct after incubating a sample of the protein in phosphate buffer saline (PBS, pH 7.4) at 37 °C

for 2 days, (Supporting Information, p. S6). Aspartimide is known to be susceptible to nucleophilic attack by water leading to the formation of Asp and *iso*Asp (Scheme 3),^{25a} the latter being often preferred (average ratio 1:3). In addition to hydrolysis, Asi can also epimerize due to increased acidity of the C α hydrogen.⁴⁰ This leads to the formation of D-Asi, which will in turn be able to generate D-Asp and D-*iso*Asp (Scheme 3).

The decrease we observed in the amount of the M-18 byproduct can thus likely be attributed to hydrolysis, which would therefore be faster than Asi formation if any occurs under the incubation conditions. This observation did not rule out our hypothesis of a particularly Asi formation-sensitive Asp residue in SUMO-2, and we therefore conducted experiments to determine the culpable residue. Taking advantage that the two DG motifs of SUMO-2 are located

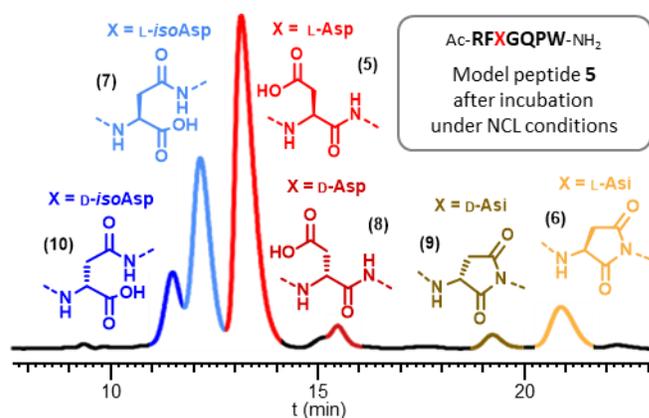
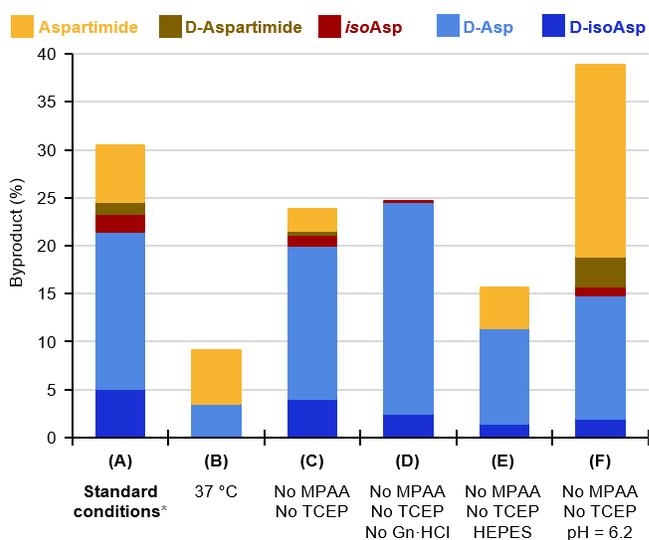


Figure 1. Typical RP-HPLC monitoring of the formation of aspartimide and related byproducts from the DG motif-containing model peptide 5 upon incubation under NCL conditions (featured chromatogram: standard NCL condition, at 50 °C for 192 h).



*: 100 mM MPAA, 50 mM TCEP, 6 M Gn-HCl, 0.2 M phosphate, pH = 7.2, 50 °C

Figure 2. Different conditions screened to try to identify which parameters were responsible for the formation of Asi and related byproducts during NCL using peptide 5 as a model substrate. All incubations were performed for 196 h.

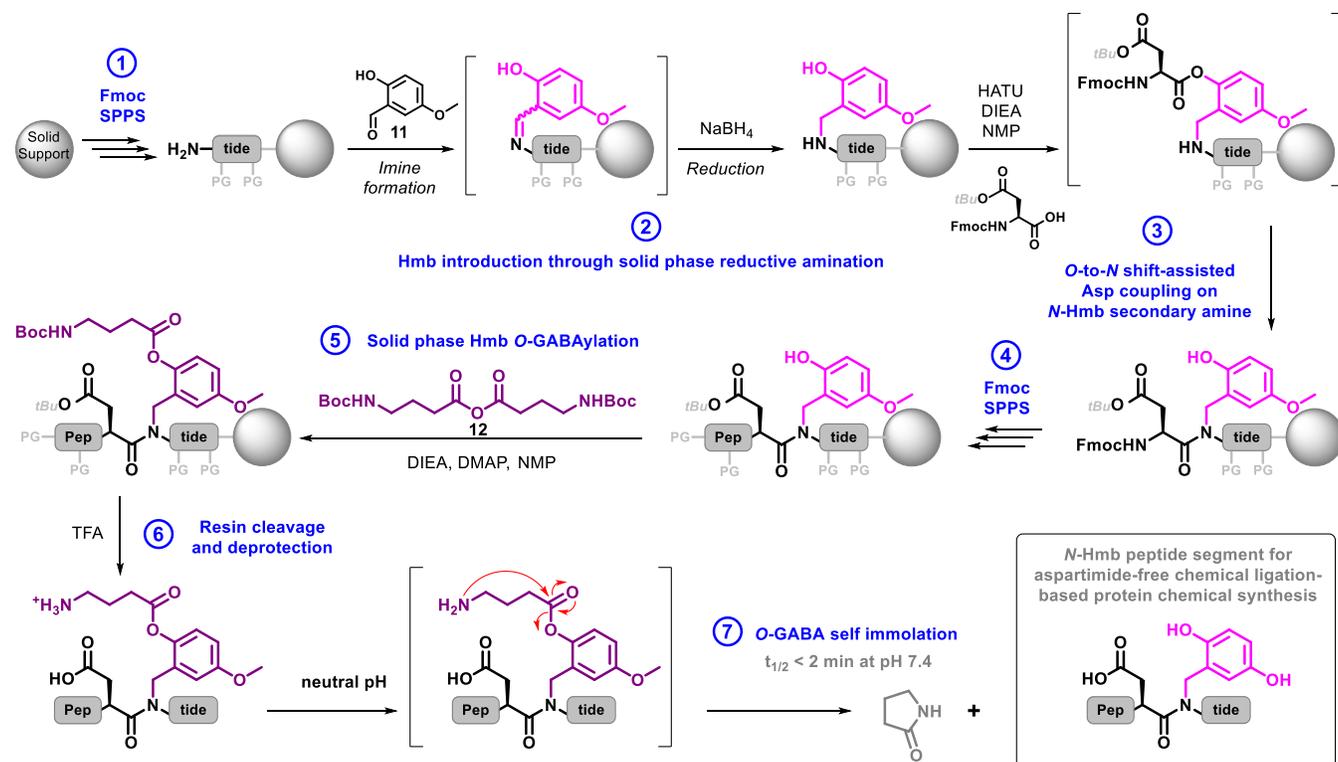
in each of the separate segments, we separately incubated **1** and **2a** under NCL conditions (Scheme 4). Contrary to our expectations, we found that the two segments both led to the formation of M-18 byproducts, and with comparable kinetics: approximately 11% after 96 h for each of the two segments (Supporting Information, p. S6-S7). If the hypothesis that only a few specific sequences, amongst ones containing a DG or NG motif, are genuinely predisposed to aspartimide formation under NCL conditions is valid, it would seem very improbable that SUMO-2 contains two such sequences. The M-18 Da mass observed during the analysis of the full-length NCL product **3** is thus likely a mixture of the two aspartimide formed from both Asp residues at positions 27 and 64.

Taken collectively, our results suggest that the formation of aspartimide, assumed to be the M-18 Da products, could be a widespread and underestimated side reaction frequently occurring during NCL. Given that the separation by standard HPLC means of the target protein from Asi, but also from all the related byproducts (e.g. *iso*Asp, D-Asp, D-*iso*Asp, and D-Asi), is not expected in the case of a long polypeptide chain, and that the mass of three of the byproducts is identical to the target, this raises an important concern that needs to be addressed.

To scrutinize the formation of Asi during NCL, we concentrated on the short model heptapeptide **5** encompassing the SUMO-2 [61-67] sequence with a central DG motif, and incorporating an extra C-terminal tryptophan residue for UV spectrophotometry quantification purposes. In addition to **5** we synthesized five analytical standards (**6-10**) corresponding to Asi formation (**6**), as well as related byproducts

that could be potentially be formed from **6** under NCL conditions. Peptides **7**, **8** and **10** (*L-iso*Asp, *D-iso*Asp and D-Asp, respectively) were obtained by Fmoc-SPPS using commercially available building blocks: Fmoc-Asp-OtBu, Fmoc-D-Asp(OtBu)-OH and Fmoc-D-Asp(OtBu)-OH, respectively. During the synthesis of *L/D iso*Asp-containing peptides **7** and **10**, we were able to obtain *L/D* aspartimide-containing peptides **6** and **9** respectively, formed during the SPPS (Supporting Information, p. S8-S17). We prepared an equimolar mixture of peptides **5-10** and optimized HPLC conditions for their analytical separation (Fig. S15), feasible for such small segments in strong contrast from what is expected with much longer segments typically employed for NCL-based chemical protein synthesis.

Having such a powerful analytical tool in hands, we first examined the stability of the aspartimide peptide **6**. Confirming the results observed with full length SUMO-2, incubation of **6** in PBS for 48 h at 37°C led to > 98% conversion into a mixture of peptides **5** and **7-10** (table. S1 p S18), with the *iso*Asp peptide **7** being the major compound (67%), having the same molecular mass as the parent Asp peptide **5**. Similarly, incubation under standard NCL conditions (6 M Gn-HCl, 100 mM MPAA, 50 mM TCEP and 200 mM phosphate, pH = 7.2) for 16 h at 37 °C also led to the formation of significant amounts of the same products, albeit at lower rates, including 12% of epimerized forms (**8-10**) and 16% of *iso*Asp **7**. This clearly indicates that, if formed during NCL, an aspartimide can further be converted into other byproducts, including D- and *iso*-forms undistinguishable by simple MS and HPLC analysis.



Scheme 5. General strategy to synthesize Hmb-containing peptide segments. Steps 1-5 are fully automated on a standard peptide synthesizer. Step 6 makes use of standard TFA cleavage cocktails. Step 7 consists in simple dissolution in PBS buffer.

We then turned to monitoring the formation of Asi and related byproducts upon incubation of the native Asp peptide **5** under various conditions, starting with a standard NCL buffer, at 50 °C for several days (Figure 1). These NCL conditions were deliberately boosted in terms of temperature and time in order to facilitate the quantification of all the expected byproducts. During this first experiment, we clearly observed the formation of a peak (c.a. 9%) that was co-eluting with the Asi standard (**6**), confirming that the product with an 18 Da mass loss formed during NCL was an aspartimide. Importantly, we were also able to observe the formation of *iso*-Asp and epimerization products in considerable amounts, Asi only representing 20% of the total amount of byproducts. The incubation of an NG motif-containing peptide analogue under similar conditions also led to aspartimide and related byproducts formation, in comparable amounts (Fig. S17).

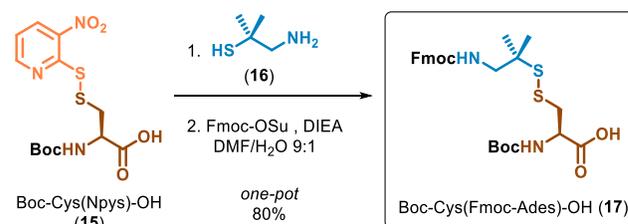
Following these observations, we tried to identify which parameters were responsible for the formation of Asi during NCL in order to help prevent this in future ligation reactions. To do this, we screened a set of different conditions, either by omitting some of the NCL components (TCEP, MPAA, Gn-HCl), or by changing the pH, the temperature or the nature of the buffer (Figure 2 and Supporting Information, p. S18-S28). Although none of the above variations did suppress the formation of Asi and related byproducts, each set of conditions gave rise to different outcomes. For example, we observed that the presence of phosphate as a buffer significantly increased the extent of epimerization as well as Asi formation compared to HEPES (Figure 2, entries C and E and Fig. S18); or again, that guanidinium chloride exerted a protective effect against the hydrolysis of Asi. Consistent with earlier studies on small peptides models of protein degradation,^[16] lowering the pH proved to have a strong effect: if the total amount of byproducts and their relative proportions do not significantly vary in the 6.7-7.7 pH range, a much higher rate of aspartimide was observed at pH 6.2 (entries C and F and Fig. S25-27).

Not surprisingly,^{25,20c,d} the parameter that most affected the occurrence of side reactions is temperature. Lowering it from 50 °C to 37 °C (entries A and B) or 20 °C reduced the total amount of byproducts by nearly threefold, and more than twentyfold, respectively, but still did not completely suppress the side reactions.

In this comprehensive analytical study, we clearly demonstrated that aspartimide is formed under NCL conditions. Our results suggest that this could be the source of side-reactions potentially occurring frequently during NCL-based chemical protein synthesis. Such byproducts are likely to have been somewhat underestimated because of the furtive features of Asi and its low stability, resulting in byproducts with a mass equal to that of the target protein, and indistinguishable using standard HPLC/MS methods.⁴¹ Considering that the formation of aspartimide and its related byproducts can significantly alter the 3D structure and the activity of a protein, it is of prime importance to devise methods to prevent it. In our attempt to identify the NCL parameters promoting aspartimide formation, we established that the reaction can occur spontaneously and is probably mainly catalyzed by water rather than other NCL buffer components, as observed in previous biochemical studies and confirmed

by DFT simulations.⁴² The formation of aspartimide and related byproducts can be limited by adopting "good NCL practices", which involve restricting the ligation temperature and reaction times, as well as replacing the commonly used phosphate buffer with HEPES. However, the efficiency of such precautions is expected to vary considerably depending on the Asi formation-propensity of the sequence of the target protein, which remains difficult to evaluate, even though DG- and NG-motifs containing sequences can be likely considered as the most problematic given the sum of NCL-unrelated studies. Even more concerning, the amounts of Asi and its byproducts are expected to grow with the length of the target protein, as a result of the number of NCL reactions to be implemented and potential aspartimide hotspots (Supporting Information p S29-S32).

A general synthetic strategy to suppress the formation of aspartimide. Therefore, we wished to develop a straightforward, efficient and generally-applicable strategy to totally suppress Asi formation at targeted sites during NCL. Inspired by strategies developed for SPPS, we decided to incorporate a protecting group on the backbone amide nitrogen atom involved in Asi-forming cyclization, thus making the side reaction impossible to occur.²² Our criteria for an ideal protecting group were that it should be: chemically-compatible with Fmoc-SPPS and NCL, easily incorporated into the peptide using inexpensive reagents and easily removable under mild conditions. We opted for the 2-hydroxy-4-methoxybenzyl (Hmb) moiety as a removable backbone modification, initially developed to counter peptide aggregation during SPPS,⁴³ and later used for the same purpose during peptide purification⁴⁴ and NCL.⁴⁵ Importantly, *N*-Hmb protection has also been reported as a highly effective method to suppress aspartimide formation during Fmoc-SPPS.⁴⁶ It can be incorporated in a peptide either using commercially available Hmb-containing building blocks, or through solid-phase reductive amination. Hmb is readily removed by trifluoroacetic acid (TFA) treatments used for peptide resin cleavages after Fmoc-SPPS, but in our case, we need to retain the Hmb protection while carrying out the NCL. Conveniently, its sensitivity to TFA can be completely switched off by acylating the 2-hydroxy group,^{44,45} typically with an acetyl group. Acyl group removal restores the sensitivity towards TFA, mild conditions for most peptides. Deacetylation is carried out either by a treatment with hydrazine,^{44,45a} or a very large excess of thiolate,^{45b} which are basic conditions potentially detrimental to functional groups, and susceptible to promote formation of Asi and related byproducts. Considering this drawback, we opted for



Scheme 6. Synthesis from commercially-available materials of the Fmoc SPPS building block Boc-Cys(Fmoc-Ades)-OH (**17**) used for the peptide synthesizer-automated introduction of a disulfide-based (Lys)₆-Ades solubilizing tag on the N-terminal cysteines.

a self-immolative phenol ester that would be removed under mild neutral aqueous conditions through intramolecular aminolysis. Such a strategy has been initially introduced for the protection of tyrosine phenol by the *N*-methyl-*N*-[2-(methyl-amino)-ethyl]-carbamoyl (Nmec) group,⁴⁷ later applied to an acid sensitivity-switchable peptide solubilizing tag,⁴⁸ and an *S*-Hmb cysteine side chain protecting group.⁴⁹ However, due to the modest electrophilicity of carbamate and low nucleophilicity of the secondary amine, the removal of Nmec is relatively slow at neutral pH. We decided to set our sights on the gamma-aminobutyric acid ester (GABA) group, for which a fast self-immolation (Scheme 5, step 7) has been highlighted in other contexts.⁵⁰

The synthesis of Hmb-equipped segments required the optimization of two key steps, (1) Hmb incorporation and (2) its solid-phase GABAylation (Scheme 5, steps 2 and 5, and Supporting Information, p. S33-S39). We opted for a reductive amination approach to incorporate the Hmb from the inexpensive 2-hydroxy-4-methoxybenzaldehyde (HmbA, **11**).⁵¹ The resulting secondary amine benefits from the assistance of the 2-hydroxy group of Hmb to facilitate the coupling of the next residue (Scheme 5, steps 3),⁵² in our case Asp(OtBu) or Asn(Trt).

We first optimized conditions for Hmb incorporation using the SUMO-2[47-93] segment as a model.⁵³ We found that

a key point to ensure a quantitative reaction was the repetition of a long treatment with excess HmbA (2 treatments of 2 hours each with 10 equiv.) in a DMF/MeOH/AcOH 44.5:44.5:1 mixture to form the imine, before washing away the excess aldehyde and performing its reduction with sodium borohydride in DMF (Supporting Information, p. S31). Gratifyingly, these conditions were readily transposable to other sequences, including SUMO-2[1-46], and were easy to automatize on a peptide synthesizer.

After Hmb incorporation and peptide elongation,⁵⁴ GABAylation can be conveniently carried out using *N*-Boc GABA anhydride (**12**) as an acylating agent readily synthesized on a multi-gram scale (Supporting Information, p. S34). As optimal conditions for this reaction we found that using an excess of *i*-Pr₂NET and **12**, in the presence of catalytic amounts of DMAP (1 mol%) in DMF, yielded very robust results (Supporting Information, p. S36).

The Boc group of Boc-GABA-Hmb is cleaved during subsequent TFA-mediated resin cleavage, resulting in a GABA-Hmb-protected peptide, which is stable in its ammonium form, for example when using RP-HPLC solvents containing 0.1% TFA.⁵⁵ Conversely, a rapid and clean self-immolation of the GABA moiety was observed upon incubation in a neutral buffer

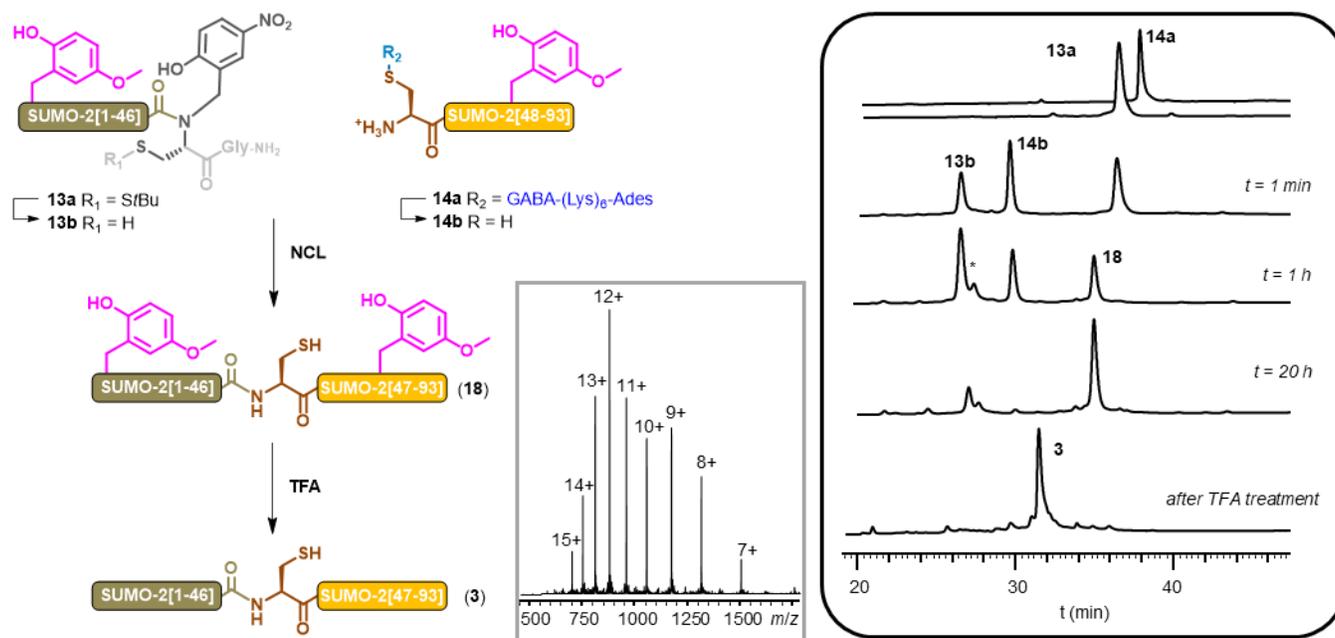


Figure 3. NCL-based aspartimide-free synthesis of SUMO-2 (NCL conditions: 200 mM phosphate, 6 M Gu-HCl, 100 mM MPAA, 50 mM TCEP, 37 °C). Grey box insert: MS of purified **3**. *: HPLC peak corresponding to the MPAA thioester derivative of **13b**.

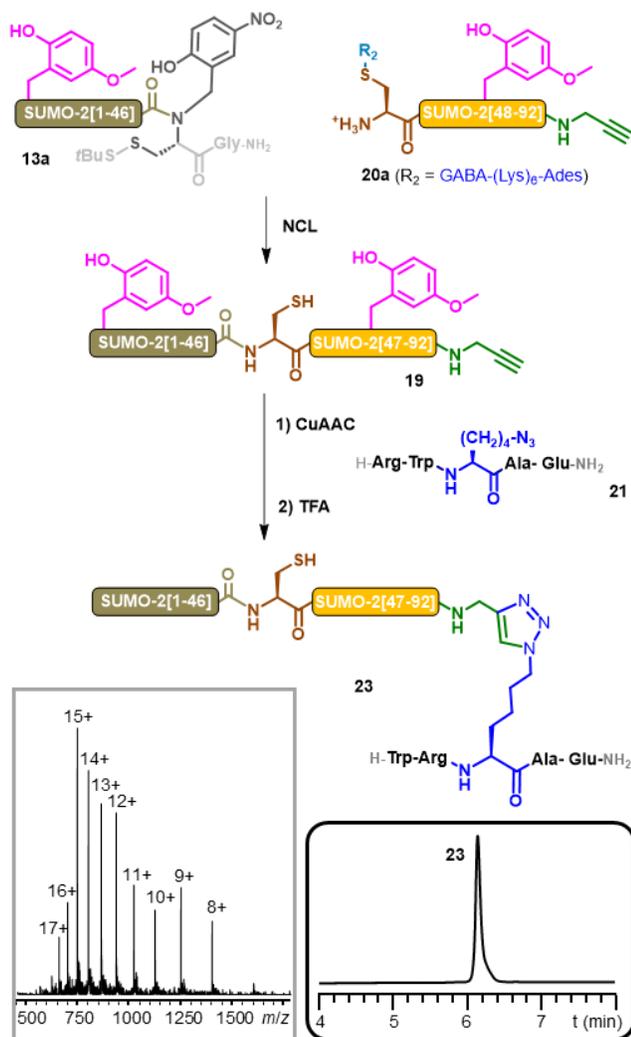


Figure 4. Aspartimide-free synthesis of SUMOylated peptide mimics **23** through a combination of native chemical ligation and peptidomimetic triazole ligation. Grey and black box inserts: MS and RP-HPLC analyses of purified **23**.

min, and complete removal within half an hour. The reaction demonstrated typical first order kinetics ($k = 0.09 \text{ s}^{-1}$), and was several orders of magnitude faster than the hydrolysis of an acetyl group, which did not even reach 10% completion after 6 days incubation under the same conditions (Supporting Information, p. S38-S39).

Having devised robust conditions for Hmb incorporation, the next step was to demonstrate the effectiveness of our strategy against Asi formation during NCL, and we chose to apply it to the synthesis of the problematic SUMO-2 protein.

The incorporation of Hmb was performed at the DG motifs of each segment. As in our previous report,^{20d} SUMO-2[1-46] (**13a**) was synthesized as a C-terminal *N*-Hmb-Cys cryptothioester⁵⁶ and SUMO-2[47-93] segment (**14a**) was equipped with a (Lys)₆-Ades solubilizing tag. The hexalysine of **14a** was N-terminally modified with a GABA group during the peptide synthesizer-mediated automated process because we thought it unnecessary to keep the N-terminal amine of the tag for the solubilizing purposes, which

would have required the coupling of the significantly more expensive Boc-Lys(Boc)-OH instead of the Fmoc-Lys(Boc)-OH standard Fmoc-SPPS building block.

We took the occasion of this work to improve conditions for the installation of the tag. Our initial strategy was based on the coupling of Boc-Cys(Npys)-OH (**15**), followed by reaction with excess 2-amino-1,1-dimethyl-ethane-1-thiol (**16**, ADET) to form the disulfide-based linker. A key feature of the Ades disulfide linker is the tertiary carbon bearing a sulfur atom within ADET, which makes the cysteine disulfide residue resistant to beta-elimination during piperidine-mediated Fmoc cleavage, inspired from the *S*-*t*Bu Cys protecting group.^{57,58}

However, we later found that our published protocol sometimes gave unreliable results, perhaps because of the instability of Boc-Cys(Npys)-OH, either during its coupling or when standing in solution prior to delivery to the peptidyl resin by the automatic peptide synthesizer, as evidenced by the observation of peptides not incorporating the Lys-Ades solubilizing tag. Additionally, this strategy requires a large excess of costly Boc-Cys(Npys)-OH and ADET. To overcome these limitations we decided to introduce the linker using a building-block, Boc-Cys(Fmoc-Ades)-OH (**22**). This compound is obtained in excellent yields through the reaction of stoichiometric amounts of Boc-Cys(Npys)-OH and ADET followed by one-pot protection of the Ades amino group using Fmoc-OSu (Scheme 6 and Supporting Information, p. S44-S46). This building block can be readily coupled as the N-terminal Cys group of a peptide segment, followed by Fmoc-SPPS-based elongation to incorporate the hexalysine solubilizing sequence.

The crude **13a** and **14a** segments were analyzed by HPLC after GABA removal under neutral conditions. Fairly clean chromatograms indicated satisfactory Hmb incorporation, GABAylation and self-immolation of the GABA (Supporting Information, p. S43-S48).

Native chemical ligation of the two purified segments proceeded smoothly and led to the expected *bis*-Hmb SUMO-2 (**18**) after overnight incubation at 37 °C (Figure 3). After a dialysis step to remove salts, TCEP and MPAA (82% recovery yield), the Hmb groups were removed by a TFA treatment followed by precipitation in diethyl ether, gratifyingly leading to a very clean crude SUMO-2 protein (**3**). A final HPLC purification afforded the pure compound in a satisfactory 21% yield. Most importantly, no trace of M-18 Da by-product could be detected in the MS analysis, not only demonstrating the effectiveness of our Hmb-based strategy, but also unequivocally confirming that the byproducts observed during our first synthesis of SUMO-2 were indeed due to aspartimide formation at the DG motifs.

Enthused by these promising outcomes, we further challenged our methodology through the synthesis of a SUMOylated peptide. The preparation of such compounds has been previously reported. However, relying exclusively on NCL necessitates multiple steps of ligation and deprotection of terminal reactive groups.^{34,59} We opted to synthesize a SUMOylated peptide by combining NCL with another orthogonal ligation method, copper(I)-catalyzed azide/alkyne

cycloaddition (CuAAC).⁶⁰ The 1,4-disubstituted 1,2,3-triazole formed in this reaction is considered as an excellent mimic of an amide bond.⁶¹

Bis-Hmb-SUMO-2-alkyne **19** was obtained through NCL using a similar synthetic strategy as for *bis*-Hmb-SUMO-2 **18** (Supporting Information, p. S54-56), from the NCL of crypto-thioester **13a** and SUMO-2[47-92]-propargylamide segment **20a** (Figure 4).

Functionalization with an alkyne moiety was achieved by introducing a propargylamide in place of ⁹³Gly in the [47-93] segment, using a backbone amide (BAL) linker⁶² (Supporting Information, p. S49).

The sequence of peptide **21** was inspired by the consensus motif of SUMOylation,⁶³ ΨKxD/E (with Ψ = any hydrophobic residue, and x = any amino acid) and was obtained by incorporating azidonorleucine instead of lysine through Fmoc-SPPS (Supporting Information, p. S57).

The CuAAC ligation between **19** and **21** proceeded cleanly in a mixture of pH 8 HEPES buffer and hexafluoroisopropanol⁶⁴ in presence of a catalytic system composed of CuSO₄, the Cu(I) ligand THPTA, and sodium ascorbate as a copper(II) ion reductant.⁶⁵ Incubation at 37 °C during 4 h resulted in the complete conversion of **19** into the corresponding *bis*-Hmb SUMOylated peptide **22**. The reaction was purified *via* HPLC before performing a TFA treatment to remove the Hmb groups. A final purification step provided the desired SUMOylated peptide mimic **23** with an overall 28% yield, with no trace of M-18 Da byproduct being detected by MS. This final result highlights the applicability of our Hmb-method to chemical protein synthesis beyond the scope of NCL, which represents a simple and efficient strategy to obtain SUMOylated peptide mimics.

CONCLUSION

In summary, we demonstrate in this work that formation of aspartimide and related byproducts occurs during NCL and could be much more frequent than one might expect. We show that these side-reactions can be limited by restricting the ligation temperature and reaction times, as well as replacing the commonly used phosphate buffer with HEPES. However, the efficiency of such “good practices” is expected to vary considerably depending on the sequence of the target protein. Moreover, the amount of byproducts is expected to grow with its length, as a result of the number of potential aspartimide hotspots and NCL reactions implemented. For example, if only 5% of aspartimide and related byproducts are formed during a first ligation involving one Asi-sensitive segment, 23% would be obtained after 4 further NCL (table S6).⁶⁶ Applying similar simple maths to proteins containing several Asi hotspots leads to alarming conclusions: If one would wish to apply similar strategies to oligo-SUMO chains which have proved instrumental in the deciphering of the ubiquitin code,⁶⁷ and considering two DG motif / Asi hotspot per SUMO, formation of as low as 1% byproduct per sensitive Asp during one ligation would translate to 31% contamination in a tetraSUMO obtained through 7 successive NCLs (table S8). 5% byproduct per hotspot would lead to 39% contamination in a diSUMO obtained through only 3 ligations (table S10).

Facing these results, we propose a simple and potentially generally applicable methodology based on an *O*-GABAylated Hmb protecting group. This removable backbone modification is easily incorporated into the problematic site by automated solid phase reductive amination using the inexpensive 2-hydroxy-4-methoxybenzaldehyde. The *O*-acylation of Hmb by the GABA group is a key point in our strategy, as it switches off lability towards TFA-based peptidyl resin ending Fmoc-SPPS. The rapid self-immolation of the GABA ester under neutral conditions yields Hmb-equipped segments that are compatible with NCL, completely preventing formation of aspartimide at the incorporation site. The Hmb can be cleanly and efficiently removed from the synthesized protein through a final simple TFA treatment. The efficiency of this strategy was demonstrated through the aspartimide-free synthesis of the SUMO-2 protein and an analogue equipped with a C-terminal alkyne, as well as a mimic of a SUMOylated peptide, further demonstrating the compatibility of the methodology with copper-catalyzed azide/alkyne cycloaddition, another chemoselective ligation reaction frequently used in peptide chemistry.

ACKNOWLEDGMENT

We thank Jean-Baptiste Madinier for his help in peptide synthesis, and Dr Cyril Colas for performing LC/HRMS analyses. We also thank the NMR sub-platform of the MO2VING facility (Orléans, France) and its technical manager Hervé Meudal for the NMR analyses. We warmly thank Dr Carlo Pifferi and Dr Ross Ballantine for proofreading of the manuscript, and Dr Skander Abboud for preliminary exploration of the synthesis of SUMO-2 alkyne and its CuAAC conjugation to a model peptide. The University of Orléans and doctoral school SSBCV (ED549) are gratefully acknowledged for a PhD fellowship for EhC, and the Centre de Biophysique Moléculaire for the financial support granted to acquire the LC-MS and preparative HPLC used for characterization and purification.

REFERENCES

- (1) Schnölzer, M.; Kent, S. B. H. Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease. *Science* **1992**, *256* (5054), 221–225. <https://doi.org/10.1126/science.1566069>.
- (2) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154. <https://doi.org/10.1021/ja00897a025>.
- (3) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266* (5186), 776–779. <https://doi.org/10.1126/science.7973629>.
- (4) a) Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. A “Traceless” Staudinger Ligation for the Chemoselective Synthesis of Amide Bonds. *Org. Lett.* **2000**, *2* (14), 2141–2143. <https://doi.org/10.1021/ol006054v>; b) Nilsson, B. L.;

- Kiessling, L. L.; Raines, R. T. Staudinger Ligation: A Peptide from a Thioester and Azide. *Org. Lett.* **2000**, *2* (13), 1939–1941. <https://doi.org/10.1021/ol0060174>.
- (5) Bode, J. W. Chemical Protein Synthesis with the α -Ketoacid–Hydroxylamine Ligation. *Acc. Chem. Res.* **2017**, *50* (9), 2104–2115. <https://doi.org/10.1021/acs.accounts.7b00277>.
- (6) Liu, H.; Li, X. Serine/Threonine Ligation: Origin, Mechanistic Aspects, and Applications. *Acc. Chem. Res.* **2018**, *51* (7), 1643–1655. <https://doi.org/10.1021/acs.accounts.8b00151>.
- (7) *Alternative strategies to chemical ligations are also actively pursued, such as automated flow SPPS:* Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of Proteins by Automated Flow Chemistry. *Science* **2020**, *368* (6494), 980–987. <https://doi.org/10.1126/science.abb2491>.
- (8) **a)** Kulkarni, S. S.; Sayers, J.; Premdjee, B.; Payne, R. J. Rapid and Efficient Protein Synthesis through Expansion of the Native Chemical Ligation Concept. *Nat. Rev. Chem.* **2018**, *2* (4), 0122. <https://doi.org/10.1038/s41570-018-0122>; **b)** Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119* (12), 7328–7443. <https://doi.org/10.1021/acs.chemrev.8b00712>.
- (9) Agouridas, V.; El Mahdi, O.; Cargoët, M.; Melnyk, O. A Statistical View of Protein Chemical Synthesis Using NCL and Extended Methodologies. *Bioorg. Med. Chem.* **2017**, *25* (18), 4938–4945. <https://doi.org/10.1016/j.bmc.2017.05.050>.
- (10) Sun, H.; Brik, A. The Journey for the Total Chemical Synthesis of a 53 KDa Protein. *Acc. Chem. Res.* **2019**, *52* (12), 3361–3371. <https://doi.org/10.1021/acs.accounts.9b00372>.
- (11) Fan, C.; Deng, Q.; Zhu, T. F. Bioorthogonal Information Storage in L-DNA with a High-Fidelity Mirror-Image Pfu DNA Polymerase. *Nat. Biotechnol.* **2021**, *39* (12), 1548–1555. <https://doi.org/10.1038/s41587-021-00969-6>.
- (12) *See for example:* Kamo, N.; Hayashi, G.; Okamoto, A. Triple Function of 4-Mercaptophenylacetic Acid Promotes One-Pot Multiple Peptide Ligation. *Angew. Chem. Int. Ed.* **2018**, *57* (50), 16533–16537. <https://doi.org/10.1002/anie.201809765>.
- (13) *See for example:* Abboud, S. A.; Amoura, M.; Madinier, J.; Renoux, B.; Papot, S.; Piller, V.; Aucagne, V. Enzyme-Cleavable Linkers for Protein Chemical Synthesis through Solid-Phase Ligations. *Angew. Chem. Int. Ed.* **2021**, *60* (34), 18612–18618. <https://doi.org/10.1002/anie.202103768>.
- (14) *See for example: a)* Fang, G.; Li, Y.; Shen, F.; Huang, Y.; Li, J.; Lin, Y.; Cui, H.; Liu, L. Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angew. Chem. Int. Ed.* **2011**, *50* (33), 7645–7649. <https://doi.org/10.1002/anie.201100996>; **b)** Blanco-Canales, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. Chemical Protein Synthesis Using a Second-Generation *N*-Acylurea Linker for the Preparation of Peptide–Thioester Precursors. *J. Am. Chem. Soc.* **2015**, *137* (22), 7197–7209. <https://doi.org/10.1021/jacs.5b03504>.
- (15) Zhao, D.-D.; Fan, X.-W.; Hao, H.; Zhang, H.-L.; Guo, Y. Temporary Solubilizing Tags Method for the Chemical Synthesis of Hydrophobic Proteins. *Curr. Org. Chem.* **2019**, *23* (1), 2–13. <https://doi.org/10.2174/1385272822666181211121758>.
- (16) **a)** Jin, K.; Li, X. Advances in Native Chemical Ligation–Desulfurization: A Powerful Strategy for Peptide and Protein Synthesis. *Chem. Eur. J.* **2018**, *24* (66), 17397–17404. <https://doi.org/10.1002/chem.201802067>; **b)** Trunschke, S.; Piemontese, E.; Fuchs, O.; Abboud, S.; Seitz, O. Enhancing Auxiliary-Mediated Native Chemical Ligation at Challenging Junctions with Pyridine Scaffolds. *Chem. Eur. J.* **2022**, *28* (68), e202202065. <https://doi.org/10.1002/chem.202202065> and references cited therein.
- (17) Loibl, S. F.; Harpaz, Z.; Zitterbart, R.; Seitz, O. Total Chemical Synthesis of Proteins without HPLC Purification. *Chem. Sci.* **2016**, *7* (11), 6753–6759. <https://doi.org/10.1039/C6SC01883A>.
- (18) *See for example:* Xu, W.; Jiang, W.; Wang, J.; Yu, L.; Chen, J.; Liu, X.; Liu, L.; Zhu, T. F. Total Chemical Synthesis of a Thermostable Enzyme Capable of Polymerase Chain Reaction. *Cell Discovery* **2017**, *3* (1), 17008. <https://doi.org/10.1038/celldisc.2017.8>.
- (19) Snella, B.; Diemer, V.; Drobecq, H.; Agouridas, V.; Melnyk, O. Native Chemical Ligation at Serine Revisited. *Org. Lett.* **2018**, *20* (23), 7616–7619. <https://doi.org/10.1021/acs.orglett.8b03355>.
- (20) **a)** Cowper, B.; Shariff, L.; Chen, W.; Gibson, S. M.; Di, W.-L.; Macmillan, D. Expanding the Scope of $N \rightarrow S$ Acyl Transfer in Native Peptide Sequences. *Org. Biomol. Chem.* **2015**, *13* (27), 7469–7476. <https://doi.org/10.1039/C5OB01029B>; **b)** Bouchenna, J.; Sénéchal, M.; Drobecq, H.; Vicogne, J.; Melnyk, O. The Problem of Aspartimide Formation During Protein Chemical Synthesis Using SEA-Mediated Ligation. In *Peptide and Protein Engineering*; Iranzo, O., Roque, A. C., Eds.; Springer Protocols Handbooks; Springer US: New York, NY, **2020**; pp 13–28. https://doi.org/10.1007/978-1-0716-0720-6_2; **c)** Bouchenna, J.; Sénéchal, M.; Drobecq, H.; Stankovic-Valentin, N.; Vicogne, J.; Melnyk, O. The Role of the Conserved SUMO-2/3 Cysteine Residue on Domain Structure Investigated Using Protein Chemical Synthesis. *Bioconjugate Chem.* **2019**, *30* (10), 2684–2696. <https://doi.org/10.1021/acs.bioconjugchem.9b00598>; **d)** Abboud, S. A.; Cisse, E. h.; Doudeau, M.; Bénédetti, H.; Aucagne, V. A Straightforward Methodology to

Overcome Solubility Challenges for N-Terminal Cysteinylyl Peptide Segments Used in Native Chemical Ligation. *Chem. Sci.* **2021**, *12* (9), 3194–3201. <https://doi.org/10.1039/D0SC06001A>.

(21) Lelièvre, D.; Terrier, V. P.; Delmas, A. F.; Aucagne, V. Native Chemical Ligation Strategy to Overcome Side Reactions during Fmoc-Based Synthesis of C-Terminal Cysteine-Containing Peptides. *Org. Lett.* **2016**, *18* (5), 920–923. <https://doi.org/10.1021/acs.orglett.5b03612>.

(22) For a review on aspartimide formation during SPPS, see: Subirós-Funosas, R.; El-Faham, A.; Albericio, F. Aspartimide Formation in Peptide Chemistry: Occurrence, Prevention Strategies and the Role of N-Hydroxylamines. *Tetrahedron* **2011**, *67* (45), 8595–8606. <https://doi.org/10.1016/j.tet.2011.08.046>.

(23) For a recent strategy for the prevention of aspartimide formation during SPPS, based on the masking of the aspartate side chain as a sulfurylide, see: **a)** Neumann, K.; Farnung, J.; Baldauf, S.; Bode, J. W. Prevention of Aspartimide Formation during Peptide Synthesis Using Cyanosulfurylides as Carboxylic Acid-Protecting Groups. *Nat. Commun.* **2020**, *11* (1), 982. <https://doi.org/10.1038/s41467-020-14755-6>; **b)** Neumann, K.; Vujinovic, A.; Kamara, S.; Zwicky, A.; Baldauf, S.; Bode, J. W. Synthesis of Multi-Module Low Density Lipoprotein Receptor Class A Domains with Acid Labile Cyanopyridiniumylides (CyPY) as Aspartic Acid Masking Groups. *RSC Chem. Biol.* **2023**, *4* (4), 292–299. <https://doi.org/10.1039/D2CB00234E>.

(24) **a)** Clarke, S. Aging as War between Chemical and Biochemical Processes: Protein Methylation and the Recognition of Age-Damaged Proteins for Repair. *Ageing Res. Rev.* **2003**, *2* (3), 263–285. [https://doi.org/10.1016/S1568-1637\(03\)00011-4](https://doi.org/10.1016/S1568-1637(03)00011-4); **b)** Robinson, N. E.; Robinson, A. B. Molecular Clocks. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (3), 944–949. <https://doi.org/10.1073/pnas.98.3.944>.

(25) **a)** Geiger, T.; Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides. Succinimide-Linked Reactions That Contribute to Protein Degradation. *J. Biol. Chem.* **1987**, *262* (2), 785–794. [https://doi.org/10.1016/S0021-9258\(19\)75855-4](https://doi.org/10.1016/S0021-9258(19)75855-4); **b)** Stephenson, R. C.; Clarke, S. Succinimide Formation from Aspartyl and Asparaginyl Peptides as a Model for the Spontaneous Degradation of Proteins. *J. Biol. Chem.* **1989**, *264* (11), 6164–6170. [https://doi.org/10.1016/S0021-9258\(18\)83327-0](https://doi.org/10.1016/S0021-9258(18)83327-0); **c)** Wakankar, A. A.; Borchardt, R. T. Formulation Considerations for Proteins Susceptible to Asparagine Deamidation and Aspartate Isomerization. *J. Pharm. Sci.* **2006**, *95* (11), 2321–2336. <https://doi.org/10.1002/jps.20740>.

(26) **a)** Elashal, H. E.; Koos, J. D.; Cheung-Lee, W. L.; Choi, B.; Cao, L.; Richardson, M. A.; White, H. L.; Link, A. J. Biosynthesis and Characterization of Fuscimiditide, an Aspartimidylated Graspptide. *Nat. Chem.* **2022**, *14* (11), 1325–1334. <https://doi.org/10.1038/s41557-022-01022-y>; **b)** Cao, L.; Do, T.; Zhu, A.; Duan, J.; Alam, N.; Link, A. J. Genome Mining and

Discovery of Imiditides, a Family of RiPPs with a Class-Defining Aspartimide Modification. *J. Am. Chem. Soc.* **2023**, *145* (34), 18834–18845. <https://doi.org/10.1021/jacs.3c03991>; **c)** Choi, B.; Acuña, A.; Koos, J. D.; Link, A. J. Large-Scale Bioinformatic Study of Graspimiditides and Structural Characterization of Albusimiditide. *ACS Chem. Biol.* **2023**, *18* (11), 2394–2404. <https://doi.org/10.1021/acscchembio.3c00365>; **d)** Cao, L.; Elashal, H. E.; Link, A. J. Kinetics of Aspartimide Formation and Hydrolysis in Lasso Peptide Lihuanodin. *Biochemistry* **2023**, *62* (3), 695–699. <https://doi.org/10.1021/acs.biochem.2c00707>.

(27) Becker, C. F. W.; Hunter, C. L.; Seidel, R.; Kent, S. B. H.; Goody, R. S.; Engelhard, M. Total Chemical Synthesis of a Functional Interacting Protein Pair: The Protooncogene H-Ras and the Ras-Binding Domain of Its Effector c-Raf1. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (9), 5075–5080. <https://doi.org/10.1073/pnas.0831227100>.

(28) Meluh, P. B.; Koshland, D. Evidence That the MIF2 Gene of *Saccharomyces Cerevisiae* Encodes a Centromere Protein with Homology to the Mammalian Centromere Protein CENP-C. *Mol. Biol. Cell* **1995**, *6* (7), 793–807. <https://doi.org/10.1091/mbc.6.7.793>.

(29) Cappadocia, L.; Lima, C. D. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chem. Rev.* **2018**, *118* (3), 889–918. <https://doi.org/10.1021/acs.chemrev.6b00737>.

(30) Celen, A. B.; Sahin, U. Sumoylation on Its 25th Anniversary: Mechanisms, Pathology, and Emerging Concepts. *FEBS J.* **2020**, *287* (15), 3110–3140. <https://doi.org/10.1111/febs.15319>.

(31) **a)** Mandel, N.; Agarwal, N. Role of SUMOylation in Neurodegenerative Diseases. *Cells* **2022**, *11* (21), 3395. <https://doi.org/10.3390/cells11213395>; **b)** Sarge, K. D.; Park-Sarge, O.-K. Sumoylation and Human Disease Pathogenesis. *Trends Biochem. Sci.* **2009**, *34* (4), 200–205. <https://doi.org/10.1016/j.tibs.2009.01.004>.

(32) Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas, A. F.; Aucagne, V. A Straightforward Method for Automated Fmoc-Based Synthesis of Bio-Inspired Peptide Crypto-Thioesters. *Chem. Sci.* **2016**, *7* (1), 339–345. <https://doi.org/10.1039/C5SC02630J>.

(33) Bi, S.; Liu, P.; Ling, B.; Yuan, X.; Jiang, Y. Mechanism of N-to-S Acyl Transfer of N-(2-Hydroxybenzyl) Cysteine Derivatives and Origin of Phenol Acceleration Effect. *Chin. Chem. Lett.* **2018**, *29* (8), 1264–1268. <https://doi.org/10.1016/j.ccllet.2017.11.043>.

(34) Bondalapati, S.; Eid, E.; Mali, S. M.; Wolberger, C.; Brik, A. Total Chemical Synthesis of SUMO-2-Lys63-Linked Diubiquitin Hybrid Chains Assisted by Removable Solubilizing Tags. *Chem. Sci.* **2017**, *8* (5), 4027–4034. <https://doi.org/10.1039/C7SC00488E>.

- (35) Johnson, E. C. B.; Kent, S. B. H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. *J. Am. Chem. Soc.* **2006**, *128* (20), 6640–6646. <https://doi.org/10.1021/ja058344i>.
- (36) **a)** Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-Sulfanyylethyl)Amino Native Peptide Ligation. *Org. Lett.* **2010**, *12* (22), 5238–5241. <https://doi.org/10.1021/ol102273u>; **b)** Hou, W.; Zhang, X.; Li, F.; Liu, C.-F. Peptidyl *N*, *N*-Bis(2-Mercaptoethyl)-Amides as Thioester Precursors for Native Chemical Ligation. *Org. Lett.* **2011**, *13* (3), 386–389. <https://doi.org/10.1021/ol102735k>.
- (37) Lewis, U. J.; Cheever, E. V.; Hopkins, W. C. Kinetic Study of the Deamidation of Growth Hormone and Prolactin. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1970**, *214* (3), 498–508. [https://doi.org/10.1016/0005-2795\(70\)90310-7](https://doi.org/10.1016/0005-2795(70)90310-7).
- (38) **a)** Radkiewicz, J. L.; Zipse, H.; Clarke, S.; Houk, K. N. Neighboring Side Chain Effects on Asparaginyl and Aspartyl Degradation: An Ab Initio Study of the Relationship between Peptide Conformation and Backbone NH Acidity. *J. Am. Chem. Soc.* **2001**, *123* (15), 3499–3506. <https://doi.org/10.1021/ja0026814>; **b)** Irudayanathan, F. J.; Zarzar, J.; Lin, J.; Izadi, S. Deciphering Deamidation and Isomerization in Therapeutic Proteins: Effect of Neighboring Residue. *mAbs* **2022**, *14* (1), 2143006. <https://doi.org/10.1080/19420862.2022.2143006>.
- (39) **a)** McKerrow, J. H.; Robinson, A. B. Primary Sequence Dependence of the Deamidation of Rabbit Muscle Aldolase. *Science* **1974**, *183* (4120), 85–85. <https://doi.org/10.1126/science.183.4120.85>; **b)** Lura, R.; Schirch, V. Role of Peptide Conformation in the Rate and Mechanism of Deamidation of Asparaginyl Residues. *Biochemistry* **1988**, *27* (20), 7671–7677. <https://doi.org/10.1021/bi00420a015>; **c)** Patel, K.; Borchardt, R. T. Chemical Pathways of Peptide Degradation. III. Effect of Primary Sequence on the Pathways of Deamidation of Asparaginyl Residues in Hexapeptides. *Pharm. Res.* **1990**, *07* (8), 787–793. <https://doi.org/10.1023/A:1015999012852>.
- (40) **a)** Radkiewicz, J. L.; Zipse, H.; Clarke, S.; Houk, K. N. Accelerated Racemization of Aspartic Acid and Asparagine Residues via Succinimide Intermediates: An Ab Initio Theoretical Exploration of Mechanism. *J. Am. Chem. Soc.* **1996**, *118* (38), 9148–9155. <https://doi.org/10.1021/ja953505b>; **b)** Takahashi, O.; Kirikoshi, R.; Manabe, N. Racemization of the Succinimide Intermediate Formed in Proteins and Peptides: A Computational Study of the Mechanism Catalyzed by Dihydrogen Phosphate Ion. *Int. J. Mol. Sci.* **2016**, *17* (10), 1698. <https://doi.org/10.3390/ijms17101698>.
- (41) *For a review on analytical tools for the detection aspartimide and related byproducts:* Ying, Y.; Li, H. Recent Progress in the Analysis of Protein Deamidation Using Mass Spectrometry. *Methods* **2022**, *200*, 42–57. <https://doi.org/10.1016/j.ymeth.2020.06.009>.
- (42) **a)** Takahashi, O. Two-Water-Assisted Racemization of the Succinimide Intermediate Formed in Proteins. A Computational Model Study. *Health (N. Y.)* **2013**, *05* (12), 2018–2021. <https://doi.org/10.4236/health.2013.512273>; **b)** Takahashi, O.; Kobayashi, K.; Oda, A. Modeling the Enolization of Succinimide Derivatives, a Key Step of Racemization of Aspartic Acid Residues: Importance of a Two-H₂O Mechanism. *Chem. Biodivers.* **2010**, *7* (6), 1349–1356. <https://doi.org/10.1002/cbdv.200900296>; **c)** Takahashi, O.; Kirikoshi, R. Intramolecular Cyclization of Aspartic Acid Residues Assisted by Three Water Molecules: A Density Functional Theory Study. *Comput. Sci. Discovery* **2014**, *7* (1), 015005. <https://doi.org/10.1088/1749-4699/7/1/015005>; **d)** Takahashi, O.; Kirikoshi, R.; Manabe, N. Roles of Intramolecular and Intermolecular Hydrogen Bonding in a Three-Water-Assisted Mechanism of Succinimide Formation from Aspartic Acid Residues. *Molecules* **2014**, *19* (8), 11440–11452. <https://doi.org/10.3390/molecules190811440>; **e)** Nakayoshi, T.; Kato, K.; Fukuyoshi, S.; Takahashi, H.; Takahashi, O.; Kurimoto, E.; Oda, A. Computational Studies on Nonenzymatic Succinimide-Formation Mechanisms of the Aspartic Acid Residues Catalyzed by Two Water Molecules. *Biochim. Biophys. Acta, Proteins Proteomics* **2020**, *1868* (9), 140459. <https://doi.org/10.1016/j.bbapap.2020.140459>.
- (43) Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C. A Reversible Protecting Group for the Amide Bond in Peptides. Use in the Synthesis of 'Difficult Sequences.' *J. Chem. Soc., Chem. Commun.* **1993**, *4*, 369–372. <https://doi.org/10.1039/C39930000369>.
- (44) Quibell, M.; Turnell, W. G.; Johnson, T. Preparation and Purification of .Beta.-Amyloid (1-43) via Soluble, Amide Backbone Protected Intermediates. *J. Org. Chem.* **1994**, *59* (7), 1745–1750. <https://doi.org/10.1021/jo00086a025>.
- (45) **a)** Abdel-Aal, A.-B. M.; Raz, R.; Papageorgiou, G.; Offer, J. Synthesis of Amide Backbone-Modified Peptides, Chapter 15. In *Peptide Synthesis*; Hussein, W. M., Skwarczynski, M., Toth, I., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2020; Vol. 2103, pp 225–237. https://doi.org/10.1007/978-1-0716-0227-0_15; **b)** Levinson, A. M.; McGee, J. H.; Roberts, A. G.; Creech, G. S.; Wang, T.; Peterson, M. T.; Hendrickson, R. C.; Verdine, G. L.; Danishefsky, S. J. Total Chemical Synthesis and Folding of All- l and All- d Variants of Oncogenic KRas(G12V). *J. Am. Chem. Soc.* **2017**, *139* (22), 7632–7639. <https://doi.org/10.1021/jacs.7b02988>.
- (46) **a)** Quibell, M.; Owen, D.; Packman, L. C.; Johnson, T. Suppression of Piperidine-Mediated Side Product Formation for Asp(OBut)-Containing Peptides by the Use of N-(2-Hydroxy-4-Methoxybenzyl) (Hmb) Backbone Amide Protection. *J. Chem. Soc. Chem. Commun.* **1994**, *20*, 2343. <https://doi.org/10.1039/c39940002343>; **b)** Mergler, M.; Dick, F.; Sax, B.; Weiler, P.; Vorherr, T. The Aspartimide Problem in Fmoc-Based SPPS. Part I. *J. Pept. Sci.* **2003**, *9* (1), 36–46. <https://doi.org/10.1002/psc.430>.

- (47) Wahlström, K.; Planstedt, O.; Undén, A. A Carbamoyl-Protective Group for Tyrosine That Facilitates Purification of Hydrophobic Synthetic Peptides. *Tetrahedron Lett.* **2008**, *49* (23), 3779–3781. <https://doi.org/10.1016/j.tetlet.2008.04.014>.
- (48) Zheng, J.-S.; Yu, M.; Qi, Y.-K.; Tang, S.; Shen, F.; Wang, Z.-P.; Xiao, L.; Zhang, L.; Tian, C.-L.; Liu, L. Expedient Total Synthesis of Small to Medium-Sized Membrane Proteins via Fmoc Chemistry. *J. Am. Chem. Soc.* **2014**, *136* (9), 3695–3704. <https://doi.org/10.1021/ja500222u>.
- (49) a) Qi, Y.-K.; Tang, S.; Huang, Y.-C.; Pan, M.; Zheng, J.-S.; Liu, L. Hmb^{off/on} as a Switchable Thiol Protecting Group for Native Chemical Ligation. *Org. Biomol. Chem.* **2016**, *14* (18), 4194–4198. <https://doi.org/10.1039/C6OB00450D>; b) Zoukimian, C.; Bérout, R.; Boturyn, D. 2-Hydroxy-4-Methoxybenzyl as a Thiol-Protecting Group for Directed-Disulfide Bond Formation. *Org. Lett.* **2022**, *24* (18), 3407–3410. <https://doi.org/10.1021/acs.orglett.2c01190>.
- (50) a) DeWit, M. A.; Gillies, E. R. Design, Synthesis, and Cyclization of 4-Aminobutyric Acid Derivatives: Potential Candidates as Self-Immolative Spacers. *Org. Biomol. Chem.* **2011**, *9* (6), 1846. <https://doi.org/10.1039/c0ob00890g>; b) Huang, D.; Montigny, C.; Zheng, Y.; Beswick, V.; Li, Y.; Cao, X.; Barbot, T.; Jaxel, C.; Liang, J.; Xue, M.; Tian, C.; Jamin, N.; Zheng, J. Chemical Synthesis of Native S-Palmitoylated Membrane Proteins through Removable-Backbone-Modification-Assisted Ser/Thr Ligation. *Angew. Chem. Int. Ed.* **2020**, *59* (13), 5178–5184. <https://doi.org/10.1002/anie.201914836>.
- (51) a) Abdel-Aal, A.-B. M.; Raz, R.; Papageorgiou, G.; Offer, J. Synthesis of Amide Backbone-Modified Peptides, Chapter 15. In *Peptide Synthesis*; Hussein, W. M., Skwarczynski, M., Toth, I., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2020; Vol. 2103, pp 225–237. https://doi.org/10.1007/978-1-0716-0227-0_15; b) Abdel-Aal, A. M.; Papageorgiou, G.; Raz, R.; Quibell, M.; Burlina, F.; Offer, J. A Backbone Amide Protecting Group for Overcoming Difficult Sequences and Suppressing Aspartimide Formation. *J. Pept. Sci.* **2016**, *22* (5), 360–367. <https://doi.org/10.1002/psc.2877>.
- (52) Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C. A Reversible Protecting Group for the Amide Bond in Peptides. Use in the Synthesis of ‘Difficult Sequences.’ *J. Chem. Soc., Chem. Commun.* **1993**, *4*, 369–372. <https://doi.org/10.1039/C39930000369>.
- (53) Optimization of the conditions for Hmb incorporation conditions were first undertaken using a short peptide with the same sequence as **5** (Supporting Information, p. S33). The optimized conditions were readily applicable to a NG motif-containing derivative, where the Asp(OtBu) residue was replaced by an Asn(Trt) (peptide **S9**, Supporting Information, p. S40). Disappointingly, these conditions did not enable quantitative incorporation into longer peptide segments. Consequently, we carried out a second optimization tailored for longer peptides, using the SUMO-2[47-93] segment, the key point being the repetition of a long treatment with excess HmbA (Supporting Information, p. S34).
- (54) The unprotected phenol group of the Hmb moiety is partially acylated during each AA coupling SPPS step and subsequent acetic anhydride-based capping of potentially unreacted amine groups. The ester formed is readily cleaved during the following piperidine treatment used to deprotect the N-Fmoc group. Therefore, the last AA coupled, N-terminal of the sequence, required to be protected with a TFA cleavable group (N-Boc), and its coupling is followed by a piperidine treatment to ensure recovering free phenol-containing Hmb for subsequent Boc-GABAylation. This precaution was not necessary for the incorporation of the Lys₆ solubilizing sequence cleaved during the NCL, which was thus terminated with a hydrophilic GABA group, bearing a positive charge under acidic HPLC conditions.
- (55) The free phenol group of the N-Hmb group of the crypto-thioester device is also O-GABAylated during the synthetic process. Contrasting with GABA-Hmb, we observed slow self-immolation of the GABA-Hmb under the acidic conditions used for HPLC purification (0.1% TFA, pH ≈ 2), and thus preferred to perform a neutral pH treatment to self-immolate all GABA esters prior to segment purification and subsequent NCL, a precaution that should be unnecessary for peptide segments not containing an Hmb.
- (56) Abboud, S. A.; Aucagne, V. An Optimized Protocol for the Synthesis of N -2-Hydroxybenzyl-Cysteine Peptide Crypto-Thioesters. *Org. Biomol. Chem.* **2020**, *18* (40), 8199–8208. <https://doi.org/10.1039/D0OB01737J>.
- (57) a) Seidel, C.; Klauschen, E.; Kaufmann, K. D.; Kunzek, H. Selektiv S-geschützte Cysteinpeptide. Zur Synthese von Cysteinpeptiden unter Verwendung der S-Äthylmercapto-Schutzgruppe. *J. Prakt. Chem.* **1976**, *318*, 229–247. <https://doi.org/10.1002/prac.19763180208>; b) Andreu, D.; Albericio, F.; Solé, N. A.; Munson, M. C.; Ferrer, M.; Barany, G. Formation of Disulfide Bonds in Synthetic Peptides and Proteins. In: Pennington, M.W., Dunn, B.M. (eds) *Peptide Synthesis Protocols. Methods in Molecular Biology* **1994** vol 35. Humana Press, Totowa, NJ. <https://doi.org/10.1385/0-89603-273-6:91>.
- (58) The year after our report on the Lys₆-Ades solubilizing tag, a very similar disulfide-based reductive conditions-sensitive solubilizing tag was reported by Li and co-workers, but based on a primary thiol Cys-disulfide. No mentions are made in this report of potential beta-elimination byproducts, thus contradicting the literature cited in reference 57 and our own observations. This suggests that our tertiary thiol-based Cys disulfide could be an unnecessary precaution, at least for the peptide sequences and the Fmoc-SPPS conditions described therein: Liu, J.; Wei, T.; Tan, Y.; Liu, H.; Li, X. Enabling chemical protein (semi)synthesis via reducible solubilizing tags (RSTs). *Chem. Sci.* **2022**, *13* (5): 1367–1374. <https://doi.org/10.1039/d1sc06387a>.

- (59) **a)** Bouchenna, J.; Sénéchal, M.; Drobecq, H.; Vicogne, J.; Melnyk, O. Total Chemical Synthesis of All SUMO-2/3 Dimer Combinations. *Bioconjugate Chem.* **2019**, *30* (11), 2967–2973. <https://doi.org/10.1021/acs.bioconjchem.9b00661>; **b)** Boll, E.; Drobecq, H.; Ollivier, N.; Raibaut, L.; Desmet, R.; Vicogne, J.; Melnyk, O. 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* (5), 2017–2022. <https://doi.org/10.1039/C3SC53509F>; **c)** Boll, E.; Drobecq, H.; Ollivier, N.; Blanpain, A.; Raibaut, L.; Desmet, R.; Vicogne, J.; Melnyk, O. 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* (2), 269–292. <https://doi.org/10.1038/nprot.2015.013>; **d)** Drobecq, H.; Boll, E.; Sénéchal, M.; Desmet, R.; Saliou, J.-M.; Lacapère, J.-J.; Mougel, A.; Vicogne, J.; Melnyk, O. 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* (6), 1540–1546. <https://doi.org/10.1021/acs.bioconjchem.6b00211>.
- (60) **a)** Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **2002**, *41* (14), 2596–2599. [https://doi.org/10.1002/1521-3773\(20020715\)41:14<2596::AID-ANIE2596>3.0.CO;2-4](https://doi.org/10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4); **b)** Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, *67* (9), 3057–3064. <https://doi.org/10.1021/jo011148j>.
- (61) **a)** Valverde, I. E.; Lecaillon, F.; Lalmanach, G.; Aucagne, V.; Delmas, A. F. Synthesis of a Biologically Active Triazole-Containing Analogue of Cystatin A Through Successive Peptidomimetic Alkyne–Azide Ligations. *Angew. Chem. Int. Ed.* **2012**, *51* (3), 718–722. <https://doi.org/10.1002/anie.201107222>; **b)** Aucagne, V.; Valverde, I. E.; Marceau, P.; Galibert, M.; Dendane, N.; Delmas, A. F. Towards the Simplification of Protein Synthesis: Iterative Solid-Supported Ligations with Concomitant Purifications. *Angew. Chem. Int. Ed.* **2012**, *51* (45), 11320–11324. <https://doi.org/10.1002/anie.201206428>; **c)** Galibert, M.; Piller, V.; Piller, F.; Aucagne, V.; Delmas, A. F. Combining Triazole Ligation and Enzymatic Glycosylation on Solid Phase Simplifies the Synthesis of Very Long Glycoprotein Analogues. *Chem. Sci.* **2015**, *6* (6), 3617–3623. <https://doi.org/10.1039/C5SC00773A>; **d)** Beltramo, M.; Robert, V.; Galibert, M.; Madinier, J.-B.; Marceau, P.; Dardente, H.; Decourt, C.; De Roux, N.; Lomet, D.; Delmas, A. F.; Caraty, A.; Aucagne, V. Rational Design of Triazololipopeptides Analogs of Kisspeptin Inducing a Long-Lasting Increase of Gonadotropins. *J. Med. Chem.* **2015**, *58* (8), 3459–3470. <https://doi.org/10.1021/jm5019675>.
- (62) Jensen, K. J.; Alsina, J.; Songster, M. F.; Vágner, J.; Albericio, F.; Barany, G. Backbone Amide Linker (BAL) Strategy for Solid-Phase Synthesis of C-Terminal-Modified and Cyclic Peptides. *J. Am. Chem. Soc.* **1998**, *120* (22), 5441–5452. <https://doi.org/10.1021/ja974116f>.
- (63) Hendriks, I. A.; D’Souza, R. C. J.; Yang, B.; Verlaan-de Vries, M.; Mann, M.; Vertegaal, A. C. O. Uncovering Global SUMOylation Signaling Networks in a Site-Specific Manner. *Nat. Struct. Mol. Biol.* **2014**, *21* (10), 927–936. <https://doi.org/10.1038/nsmb.2890>.
- (64) Peschke, F.; Taladriz-Sender, A.; Andrews, M. J.; Watson, A. J. B.; Burley, G. A. Glutathione Mediates Control of Dual Differential Bio-orthogonal Labelling of Biomolecules. *Angew. Chem. Int. Ed.* **2023**, e202313063. <https://doi.org/10.1002/anie.202313063>.
- (65) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angew. Chem.* **2009**, *121* (52), 10063–10067. <https://doi.org/10.1002/ange.200905087>.
- (66) Assuming the same rate of byproduct formation during each NCL, and only one Asi-sensitive residue in the whole sequence.
- (67) Kriegesmann, J.; Brik, A. Synthesis of Ubiquitinated Proteins for Biochemical and Functional Analysis. *Chem. Sci.* **2023**, *14* (37), 10025–10040. <https://doi.org/10.1039/D3SC03664B>.