Enzyme-Cleavable Linkers for Protein Chemical Synthesis through Solid-Phase Ligations

Skander A Abboud, Mehdi Amoura, J.B Madinier, Brigitte Renoux, Sébastien Papot, Véronique Piller, Vincent Aucagne

RESEARCH ARTICLE

Enzyme-cleavable linkers for protein chemical synthesis through solid-phase ligations

Skander A. Abboud,[a] Mehdi Amoura,[a] Jean-Baptiste Madinier,[a] Brigitte Renoux,[b] Sébastien Papot,[b] Véronique Piller,[a] and Vincent Aucagne*[a]

[a] Dr Skander A. Abboud, Dr Mehdi Amoura, Jean-Baptiste Madinier, Dr Véronique Piller and Dr Vincent Aucagne
Centre de Biophysique Moléculaire, CNRS UPR 4301
Rue Charles Sadron, 45071 Orléans cedex 2, France
E-mail: vincent.aucagne@cnrs-orleans.fr
[b] Dr Brigitte Renoux and Prof Sébastien Papot
Institut de Chimie des Milieux et des Matériaux de Poitiers, UMR-CNRS 7285
4 rue Michel Brunet, 86073 Poitiers cedex 9, France

Supporting information for this article is given via a link at the end of the document.

Abstract: The total synthesis of long proteins requires the assembly of multiple fragments through successive ligations. The need for intermediate purifications is a strong limitation, as this generally leads to low to very low overall yields. An early-recognized solution to this problem is the assembly of the protein on a solid support, but no mature solid-supported chemical ligation (SPCL) technology has yet emerged. One of the main reasons lies in the difficulty to immobilize a first peptide segment on a SPCL-compatible solid support through a linker that can be cleaved under very mild conditions to release the assembled protein. A wide range of SPCL linkers have been developed but their cleavage requires chemical conditions sometimes incompatible with sensitive protein targets. Here, we describe the exploration of an alternative enzymatic approach in order to trigger cleavage under extremely mild and selective conditions. Optimization of the linker structure and use of a small enzyme able to diffuse into the solid support was a key for the success of the strategy. We demonstrated its utility by the native chemical ligation-based assembly of three peptide segments to afford a 15 kDa polypeptide.

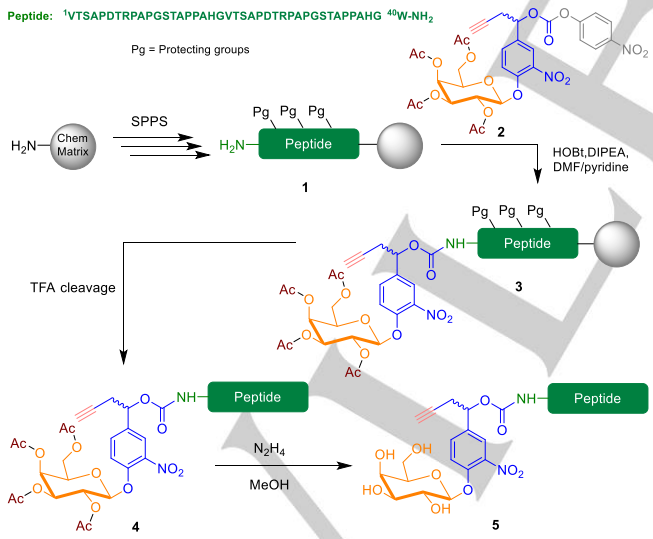
Introduction

In the last decades, the production of proteins by total synthesis became highly complementary to recombinant techniques for investigating biological mechanisms and for drug discovery. Current technologies focus on the modular assembly of unprotected peptide segments in aqueous solutions, through highly chemoselective “chemical ligation” reactions. The ligation approach revolutionized the field some thirty years ago[1] and has gradually been democratized for the synthesis of small proteins. However, the access to more ambitious targets in terms of size remains a challenge[2] since multiple successive ligations are often associated with very low overall yields, due to the laborious intermediate chromatographic purifications. Reminiscent to the solid phase peptide synthesis (SPPS) concept developed by Merrifield[3] for assembling protected amino-acids (aa), solid phase chemical ligation (SPCL)[4] has been proposed to obviate this difficulty. Sequential ligations of unprotected peptide segments on an insoluble support potentially lead to the rapid synthesis of proteins in highest yields, as purifications after each synthetic step are limited to simple washes. Although extremely appealing, SPCL has been so far essentially limited to proof-of-concepts. The choice of the linker is a keystone in SPCL strategies: it should be stable along the protein assembly then be efficiently cleaved to release the synthesized protein into solution without affecting its chemical integrity. An extensive range of SPCL linkers have been developed and introduced on either the C- [4a-c,f,g,i] or N-terminal[4a,d,e,h,l,n] part of the first peptide segment, depending on the directional strategy chosen for the successive ligations: from the C- to the N-terminus of the target protein (C-to-N), or N-to-C. Cleavage conditions include treatments with acids,[4b,c,f,g,n] bases,[4a,d,e,i] and nucleophiles,[4h,l] sometimes incompatible with sensitive targets,[4h,5] and the development of alternative linkers represents a major challenge to expand the applicability of SPCL and thus further push away the limits of chemical protein synthesis, as illustrated very recently by the development of a transition metal-cleavable linker.[6] Within this context, we thought exploiting an enzymatic rather than purely chemical approach, in order to induce cleavage under extremely soft and selective conditions. Indeed, enzymatic transformations often occur at pH 6–8, at mild temperatures and with a high chemo-, regio- and stereoselectivity for the corresponding substrates. Besides considerations on the chemical integrity of the target protein, enzymatic cleavage conditions are also non-denaturing, a potential significant advantage to exploit the “pseudo-dilution” effect[7] for the folding of the immobilized synthesized target, while avoiding intermolecular aggregation.[8] Herein we describe the exploration of a new generation of SPCL linkers, cleavable by β-galactosidases or phosphatases, and demonstrate their utility by the solid supported native chemical ligation (NCL)-based synthesis of a 160 residues polypeptide, the longest sequence ever synthesized through solid-supported NCLs.

Results and Discussion

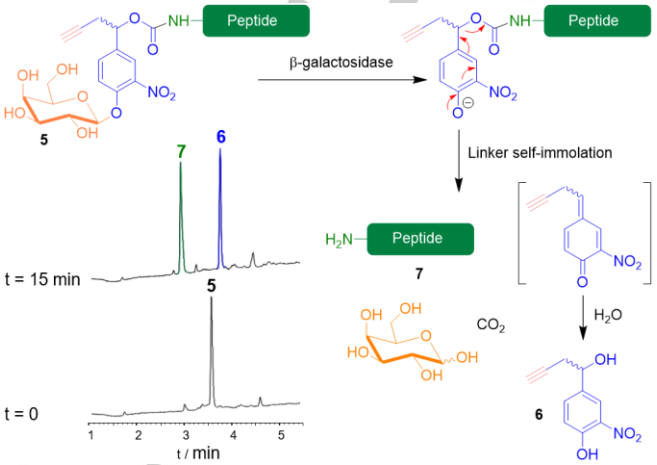
Enzyme-cleavable linkers have been developed for applications to the solid-supported synthesis of peptides,[9] glycopeptides,[9b] oligosaccharides,[10] oligonucleotides,[11] peptide nucleic acids[12] and glycosphingolipids,[13] and small molecules combinatorial libraries.[14] They exploit phosphodiesterases,[9a] lipases,[12b] glycosidases,[10a,11,13] amidases[14a,14c,13] and peptidases,[10b] to release alcohols,[10,14] amines,[11,13,14b-d] or carboxylic acids[9,11,14b] from the solid support after multistep synthesis of the target compounds. Nevertheless, none of these linkers appears as directly applicable to SPCL.

An ideal SPCL linker should allow re-immobilization of the first peptide segment on a solid support different from the one used for SPPS, due to different requirements in terms of physicochemical properties. The SPCL support has to be tolerant to the aqueous environment used for the successive ligations and the enzymatic cleavage, and be compatible with the diffusion of the enzyme. Re-immobilization has other valuable advantages, such as offering the possibility to screen different solid supports, and to allow a purification step after the SPPS in order to provide a clean and pure segment as the starting material for SPCL. We thus privileged an N-terminal linker, well adapted to a re-immobilization strategy, introduced on the N-terminal amino group of the segment after SPPS elongation. The linker precursor should thus be equipped with both an amine-reactive moiety and a functional group adapted for chemoselective immobilization of the unprotected peptide segment, and linker cleavage should release the unmodified N-terminal amine upon enzymatic treatment. Having these requirements in mind, we first evaluated a β-galactosidase-sensitive linker originally developed for targeted drug delivery.[15] Linker precursor **2** is synthesized in a few steps as a diastereomeric mixture,[15a] and bears a per-*O*-acetylated β-galactoside, an alkyne suitable for chemoselective immobilization, and an amine-reactive activated carbonate. **2** was reacted with a solid-supported model peptide **1** which sequence is based on the human mucin MUC1 variable number tandem repeat (VNTR) region, made of a duplicated 20 aa sequence and frequently used as a model for peptide ligation and bioconjugation.[4d,h,g,k,16,17] Resulting peptidyl resin **3** was cleaved from the SPPS solid support, and galactoside deacetylation was carried out in solution through treatment with hydrazine[18,19] to afford compound **5** as a mixture of two diasteroisomers easily separable by HPLC (scheme 1 and SI p S7-S8).



Scheme 1. Introduction of the β-galactosidase-labile linker on a model peptide

We first investigated the enzymatic cleavage of the linker in solution. As expected, treatment of **5** by *E. coli* β-galactosidase resulted in a clean release of peptide **7** in less than 15 minutes, through galactoside hydrolysis followed by self-immolation of the resulting *p*-hydroxybenzyloxy carbamate moiety (scheme 2).[20] Subsequently, **5** was immobilized through copper-catalyzed cycloaddition (CuAAC)[21] on three different azide-functionalized solid supports: polyethylene glycol (PEG)-crosslinked polyacrylamide resin PEGA^{1900} **8a**,[22] and controlled pore glasses CPG1000 **8b** and CPG2000 **8c**, silica-based inorganic matrixes forming porous beads with pores of 1000 and 2000 Å in diameter, respectively.[23]



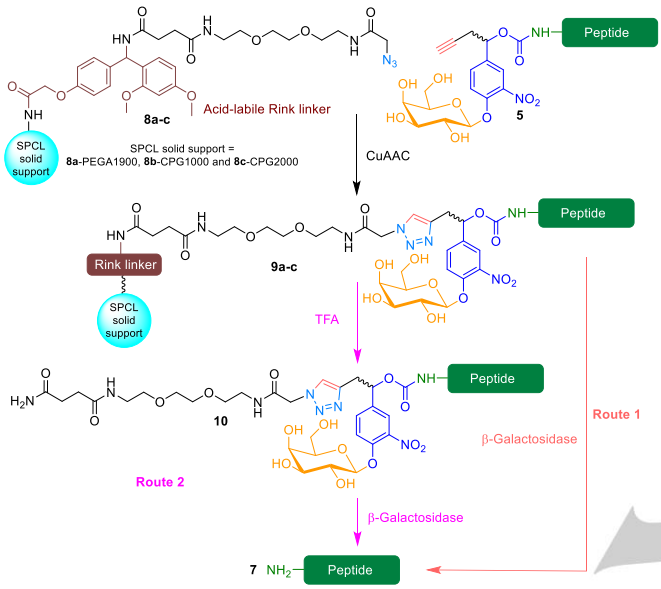
Scheme 2. Solution phase enzymatic cleavage of linker-functionalized model peptide **5**. HPLC monitoring (λ = 280 nm) shown for one diasteromer of **5**, the other one giving similar results

These supports were chosen for their superior efficiency for enzyme-mediated transformations as compared to other water-compatible resins.[4h,14d,24] For analytical purposes, the supports were functionalized with a trifluoroacetic acid (TFA)-labile Rink amide linker, separated from the azide moiety by an oligo(ethylene glycol)-based spacer to ensure accessibility of the enzyme to the substrate[25] (scheme 3).

Very disappointingly, β-galactosidase-mediated linker cleavage of immobilized peptides **9a-c** proceeded considerably slower than for **5** in solution (see SI fig. S8): after more than 5 h reaction using a ten-fold increase in enzyme concentration as compared to solution phase conditions, only 3%, 4% and 16% of peptide **7** was released from PEGA, CPG1000 and CPG2000, respectively. The superior efficacy of CPG towards PEGA is in accordance with Bradley's systematic comparison of different supports with a range of enzyme of increasing sizes,[24b] showing that enzymatic reaction rates are directly related to the ability of enzyme to diffuse inside the beads. This correlates well with the significant differences we observed between CPG1000 and 2000. The relative cleavage kinetics rate in solution and on the latter support was estimated to ~ 60 000 (SI p S59), from the ratio of apparent first order kinetic constants $k_{solution}$ / $k_{CPG2000}$. Such a slow reaction is prohibitive for SPCL applications, and all our attempts to improve cleavage efficiency were unsuccessful: increasing enzyme concentration, using a long PEG spacer (3 kDa), or prior β-galactosidase-mediated "shaving"[26] of the CPG beads in order to install azide groups only at sites easily accessible by the enzyme (see SI p S19-S26) only moderately affected cleavage kinetics. We explored an alternative approach taking advantage of the rapid enzymatic cleavage in solution, consisting in TFA-mediated cleavage of the Rink linker of **9c** followed by β-galactosidase treatment (scheme 3).[27] Although this approach gave clean results and was easy to perform, it requires one

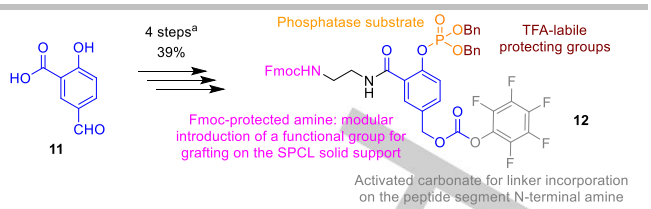
additional step, and uses denaturing conditions that would preclude future potential applications to on-resin folding following the SPCL-based assembly of ambitious targets.[28]

We assumed that the large size of β-galactosidase (a 540 kDa homo-tetramer) hampered its diffusion inside the solid supports. Hence, we developed a second-generation tailor-made linker cleavable by phosphatases, enzymes of much smaller size than β-galactosidases. Linker precursor **11** was easily synthesized in



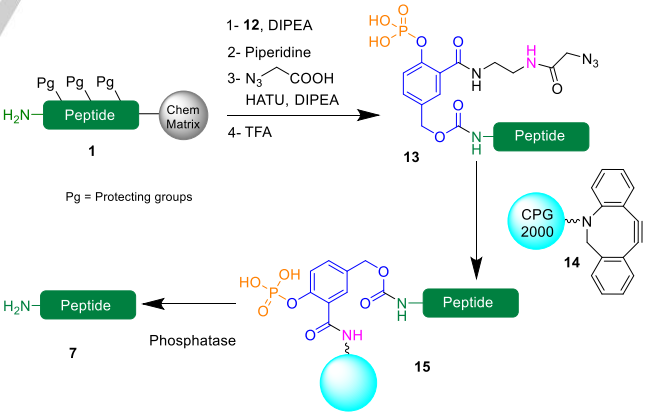
Scheme 3. The two routes explored for the release in solution of peptide **7**.

four steps from commercially-available materials (39% overall yield). (scheme 4, see SI p S29-S40 for the detailed synthesis). **11** was designed to address all the other limitations we identified with the β-galactosidase-labile linker. Indeed, the latter requires an additional step in solution after SPPS to cleave the acetate groups of the galactoside, Cu(I) salts used for CuAAC-based immobilization can lead to the oxidation of peptide segment side-chains if conditions are not well controlled,[29] and presence of a non-controlled chiral center leads to a diastereomeric mixture that complicates peptide segment purification and characterization. Our second-generation linker is achiral, and precursor **12** was designed as a *N*-fluorenylmethyloxycarbonyl (Fmoc) SPPS building block to provide optimal modularity for the further introduction of any functional group suited for chemoselective grafting on a solid support. The phosphate group is protected by TFA-labile *O*-benzyl groups, compatible with standard Fmoc-SPPS protocols.[30] Also note that the *N*-Fmoc amino group is branched through an electron-withdrawing carboxamide group in order to lower the pKa of the phenol generated upon phosphate cleavage and thus promote a fast 1,6-elimination-mediated self-immolation at neutral pH, this process being greatly accelerated by phenol deprotonation.[31]



Scheme 4. Synthesis of the second generation linker precursor. [a]: 1) *N*-Fmoc ethylene diamine, HATU, DIPEA, NMP; 2) dibenzyl phosphite, CCl_4, DIPEA, MeCN, 75% over 2 steps; 3) $NaBH_4$, THF, 97%; 4) *bis*-pentafluorophenyl carbonate, Et_3N, MeCN, 53%.

12 was reacted with solid-supported model peptide **1**, followed by *N*-Fmoc group cleavage and subsequent coupling with azidoacetic acid in order to introduce an azide group suitable for copper-free strain-promoted alkyne/azide cycloaddition (SPAAC).[32] As expected, SPPS resin cleavage cleanly afforded compound **13**. We next screened linker cleavage in solution by four commercially available enzymes: alkaline phosphatase extracted from calf intestine, and the recombinant antarctic phosphatase, shrimp alkaline phosphatase and bacteriophage lambda protein phosphatase, using the respective standard manufacturer recommendations in terms of reaction buffers and enzyme concentrations (SI, pS51). All enzymes proved efficient to mediate a clean release of peptide **1** (see SI, p S51-S53) through phosphoester hydrolysis followed by instantaneous self-immolation of the resulting *p*-hydroxybenzyloxy carbamate, not observed in LC-MS analyses as for the first generation linker. Best results in terms of kinetics were obtained with the shrimp and lambda phosphatases and these two enzymes were further evaluated for solid supported cleavage. Grafting of **13** on the CPG2000-Rink solid support **14** bearing an aza-dibenzocyclooctyne (DBCO) moiety[33] through a chemoselective SPAAC reaction gave solid-supported **15** (scheme 5).[34]

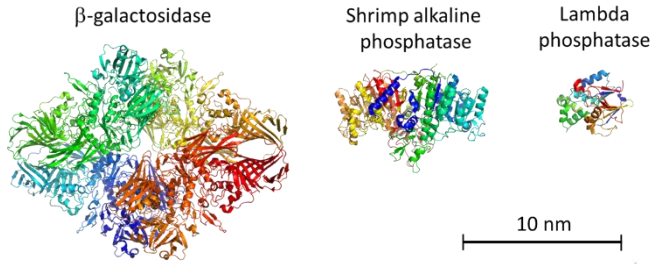


Scheme 5. 2nd generation linker introduction on model peptide and solid supported enzymatic cleavage.

To quantify the relative efficiency of the enzymatic cleavage on solid support as compared to solution phase, the $k_{solution}$ / $k_{CPG2000}$ ratio was evaluated for the two enzymes (table 1, SI p S58 and p S61-S62). As compared with the results obtained for β-galactosidase, the relative ratio was enhanced by a moderate factor of 6 for shrimp phosphatase and by an impressive factor of

300 for lambda phosphatase. These results perfectly correlate with the molecular weight and size of the three enzymes: shrimp phosphatase is a homodimeric 108 kDa protein, whereas lambda phosphatase is a much smaller 25 kDa monomer (fig. 1). Gratifyingly, these results validate our hypothesis on the influence of the size of the enzyme on the kinetics of the solid-supported linker cleavage. Despite the 200-fold slow down compared to the reaction in solution, lambda phosphatase cleavage kinetics are perfectly compatible with our SPCL purpose since the reaction can be completed in a few hours using low enzyme amounts.[35] Importantly, we never detected any peak consistent with the enzyme in LC-MS analyses of the enzymatic cleavages. This suggests a very low molar enzyme amount as compared to the substrate, thus being probably easily removed with a final HPLC "polishing" purification after SPCL assembly of a protein target.

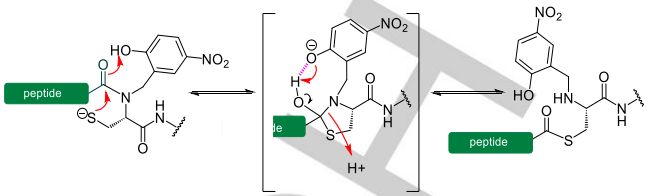
Table 1. Comparison of enzymatic cleavage efficiency.



Entry	Enzyme	MW	aa	$k_{solution}$ / $k_{CPG2000}$
1	β-galactosidase	540 kDa	4092	~ 60 000
2	Shrimp phosphatase	108 kDa	956	~ 10 000
3	Lambda phosphatase	25 kDa	221	~ 200

Figure 1. Comparison of the size of enzymes used in this work. PDB IDs: β-galactosidase: 4V40, shrimp phosphatase: 1SHQ, lambda phosphatase: 1G5B.

Having in hand a robust method for the enzymatic release from CPG2000 solid support, we then implemented this approach for the synthesis of a 160 aa polypeptide derived from MUC1 VNTRs, through two successive solid-supported NCLs. NCL is the reaction between the C-terminal C^{α} thioester function of one unprotected peptide segment and the N-terminal terminal cysteine of another segment. Consequently, a keystone for N-to-C successive NCLs is the use of dormant precursors of thioesters that can be activated on demand, to prevent cyclization and oligomerization of segments. Such precursors include thioacids,[4a] hydrazides,[36] and *N*→*S* acyl shift-based[4e,l] crypto-thioesters.[37] The latter are based on the rearrangement of a β-mercapto amide into a β-amino-thioester under NCL conditions, through intramolecular attack of the thiol on the amide. Conveniently, simple protection of the thiol afford unreactive species, keeping the crypto-thioester in the dormant state until *S*-deprotected.[4e,4l,38] In the present work, we used *N*-(2-hydroxy-4-nitrobenzyl)cysteine-based devices (*N*-Hnb-Cys)[39] that are straightforward to synthesize through automated Fmoc-SPPS and show fast rearrangement and NCL kinetics owing to internal catalysis by a judiciously placed phenol group (Scheme 6).[39a,40]



Scheme 6. Putative *N*-(Hnb)Cys self-catalyzed *N*→*S* acyl shift mechanism.

Segment **16** was equipped with both our second generation linker functionalized by an azide group and an *N*-Hnb-Cys device which thiol is protected by an *S*-*St*Bu disulfide, labile under reductive NCL conditions. This segment was immobilized onto CPG2000 solid support **14** through SPAAC, in a similar fashion to model peptide **13** (scheme 7). The second segment (**18**) bears a N-terminal cysteine and a *N*-Hnb-Cys device where the thiol is protected by a *S*-acetamidomethyl (Acm) group,[41] that is stable under NCL conditions.[38b] NCL between **17** and **18** gave a clean product as judged by an analytical-scale cleavage. The Acm group was then deprotected through treatment with Pd(II) salts[42] to afford an active crypto-thioester (**20**) which was immediately engaged into a second NCL with the C-terminal segment **21**. Finally, enzymatic cleavage with lambda phosphatase released the desired 160 aa polypeptide **23**, recovered in a 60% overall yield as determined by HPLC peak integration-based quantification. Final semi-preparative HPLC purification gave pure **23** in a 13% isolated yield.[43,44]

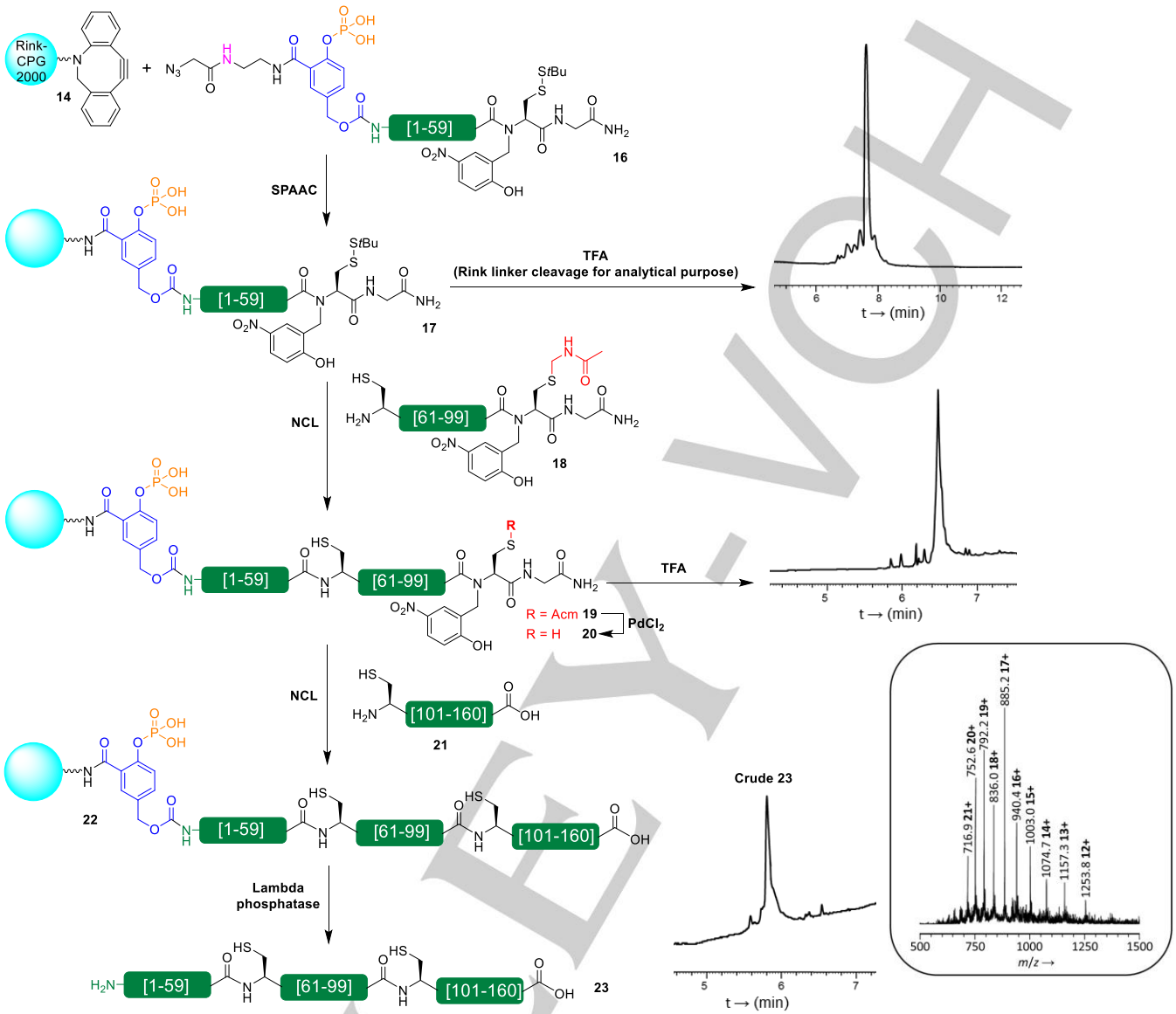
Conclusion

In summary, we introduced a new SPCL strategy based on an enzyme cleavable linker. The major bottleneck identified for the success of the strategy is the difficulty for high molecular weight enzymes to diffuse into the solid support, therefore resulting in extremely slow kinetics. Design of a tailor-made linker cleavable by a small enzyme, lambda phosphatase, gratifyingly solved the problem, and the usefulness of strategy was exemplified for the synthesis of a long polypeptide from three fragments through N-to-C NCL reactions.

Acknowledgements

We thank Dr Cyril Colas for LC/HRMS analyses, Florian Kersaudy for technical assistance as well as Dr Aromal Asokan for proofreading parts of the manuscript. We also warmly thank Mathilde Seinfeld and Laura Lefèvre for their participation to the optimization of the synthesis of different building blocks. The University of Orléans and doctoral school SSBCV (ED549) are gratefully acknowledged for a PhD fellowship for SA, and ANR (EasyMiniprot, ANR-15-CE07-0022) for financial support.

Keywords: SPCL • NCL • Linkers • Enzyme • Solid support



Scheme 7. SPCL-mediated assembly of the 160 aa polypeptide **23**.

References and notes

1. a) M. Schnolzer, S. Kent, *Science* **1992**, *256*, 221; b) P. Dawson, T. Muir, I. Clark-Lewis, S. Kent, *Science* **1994**, *266*, 776.

2. B. Zhang, Y. Li, W. Shi, T. Wang, F. Zhang, L. Liu, *Chem. Res. Chin. Univ.* **2020**, *36*, 733.

3. R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

4. a) L. E. Canne, P. Botti, R. J. Simon, Y. Chen, E. A. Dennis, S. B. H. Kent, *J. Am. Chem. Soc.* **1999**, *121*, 8720; b) A. Brik, E. Keinan, P.E. Dawson, *J. Org. Chem.* **2000**, *65*, 3829; c) E. C. B. Johnson, T. Durek, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2006**, *45*, 3283; *Angew. Chem.* **2006**, *116*, 3361; d) V. Aucagne, I. E. Valverde, P. Marceau, M. Galibert, N. Dendane, A.F. Delmas, *Angew. Chem. Int. Ed.* **2012**, *51*, 11320; *Angew. Chem.* **2012**, *124*, 11482; e) L. Raibaut, H. Adihou, R. Desmet, A. F. Delmas, V. Aucagne, O. Melnyk, *Chem. Sci.* **2013**, *4*, 4061; f) M. Jbara, M. Seenaiah, A. Brik, *Chem. Commun.* **2014**, *50*, 12534; g) I. E. Decostaire, D. Lelièvre, V. Aucagne, A. F. Delmas, *Org. Biomol. Chem.* **2014**, *12*, 5536; h) M. Galibert, V. Piller, F. Piller, V. Aucagne, A. F. Delmas, *Chem. Sci*. **2015**, *6*, 3617; i) R. R. Yu, S. K. Mahto, K. Justus, M. M. Alexander, C. J. Howard, J. J. Ottesen, *Org. Biomol. Chem.* **2016**, *14*, 2603; j) S. F. Loibl, Z. Harpaz, R. Zitterbart, O. Seitz, *Chem. Sci.* **2016**, *7*, 6753; k) R. Zitterbart, O. Seitz, *Angew. Chem. Int. Ed.* **2016**, *55*, 7252; *Angew. Chem.* **2016**, *128*, 7368; l) N. Ollivier, R. Desmet, H. Drobecq, A. Blanpain, E. Boll, B. Leclercq, A. Mougel, J. Vicogne, O. Melnyk, *Chem. Sci.* **2017**, *8*, 5362; m) R. Zitterbart, M. Krumrey, O. Seitz, *J. Pept. Sci.* **2017**, *23*, 539; n) B. Zhou, Faridoon, X. Tian, J. Li, D. Guan, X. Zheng, Y. Guo, W. Huang, *Chin. Chem. Lett.* **2018**, *29*, 1123; l) S. A. Abboud, A. F.

Delmas, V. Aucagne, Solid-phase Chemical Ligation, in *Total Chemical Synthesis of Proteins*, (Eds.: A. Brik, P. Dawson, L. Liu), Wiley-VCH, Weinheim, **2021**, pp. 259–284.

5. R. Zitterbart, N. Berger, O. Reimann, G. T. Noble, S. Lüdtke, D. Sarma, O. Seitz, *Chem. Sci.* **2021**, *12*, 2389.

6. G. B. Vamisetti, G. Satish, P. Sulkshane, G. Mann, M. H. Glickman, A. Brik, *J. Am. Chem. Soc.* **2020**, *142*, 19558.

7. P. Jayalekshmy, S. Mazur, *J. Am. Chem. Soc.* **1976**, *98*, 6710.

8. The appealing approach of post-SPCL solid-supported folding has only been reported once, for the synthesis of EETI-II through oxidative folding. Note that in this case, the disulfide-reticulated miniprotein is not sensitive to (non-reductive) denaturing conditions used for cleavage. See ref. [4c]. Solid-phase NCL followed by spontaneous folding (as judged by functional assays) has also been applied for the generation of small libraries of SH3 domains (61 aa) on microtiter plates, even though no cleavable linker was incorporated to release the synthesized peptides in solution. See references [4a] and [4m].

9. a) D. T. Elmore, D. J. S. Guthrie, A. D. Wallace, S. R. E. Bates, *J. Chem. Soc. Chem. Commun.* **1992**, *14*, 1033; b) B. A. Maltman, M. Bejugam, S. L. Flitsch, *Org. Biomol. Chem.* **2005**, *3*, 2505.

10. a) G. Corrales, A. Fernández-Mayoralas, E. García-Junceda, Y. Rodríguez, *Biocatal. Biotransform.* **2000**, *18*, 271; b) K. Yamada, S. I. Nishimura, *Tetrahedron Lett.* **1995**, *36*, 9493; c) K. Yamada, E. Fujita, S. I. Nishimura, *Carbohydrate Res.* **1997**, *305*, 443.

11. C. Schmitz, M. T. Reetz, *Org. Lett.* **1999**, *1*, 1729.

12. J. Weiler, *Nucleic Acids Res.* **1977**, *25*, 2792.

13. S. I. Nishimura, K. Yamada, *J. Am. Chem. Soc.* **1997**, *119*, 10555.

14. a) G. Böhm, J. Dowden, D. C. Rice, I. Burgess, J. F. Pilard, B. Guilbert, A. Haxton, R. C. Hunter, N. J. Turner, S. L. Flitsch, *Tetrahedron Lett.* **1998**, *39*, 3819; b) B. Sauerbrei, V. Jungmann, H. Waldmann, *Angew. Chem. Int. Ed.* **1998**, *37*, 1143; *Angew. Chem.* **1998**, *110*, 1187; c) U. Grether, H. Waldmann, *Angew. Chem. Int. Ed.* **2000**, *39*, 1629; *Angew. Chem.* **2000**, *112*, 1688; d) U. Grether, H. Waldmann, *Chem. Eur. J.* **2001**, *7*, 959.

15. a) M. Thomas, J. Clarhaut, P. O. Strale, I. Tranoy-Opalinski, J. Roche, S. Papot, *ChemMedChem*, **2011**, *6*, 1006; b) T. Legigan, J. Clarhaut, B. Renoux, I. Tranoy-Opalinski, A. Monvoisin, J. M. Berjeaud, F. Guilhot, S. Papot, *J. Med. Chem.* **2012**, *55*, 4516; c) T. Legigan, J. Clarhaut, I. Tranoy-Opalinski, A. Monvoisin, B. Renoux, M. Thomas, A. Le Pape, S. Lerondel, S. Papot, *Angew. Chem. Int. Ed.* **2012**, *51*, 11606; *Angew. Chem.* **2012**, *124*, 11556; d) B. Renoux, F. Raes, T. Legigan, E. Péraudeau, B. Eddhif, P. Poinot, I. Tranoy-Opalinski, J. Alsarraf, O. Koniev, S. Kolodych, S. Lerondel, A. Le Pape, J. Clarhaut, S. Papot, *Chem. Sci.* **2017**, *8*, 3427; e) R. Châtre, J. Lange, E. Péraudeau, P. Poinot, S. Lerondel, A. Le Pape, J. Clarhaut, B. Renoux, S. Papot, *J. Control. Release*, **2020**, *327*, 19.

16. D. Montoir, M. Amoura, Z. E. A. Ababsa, T. M. Vishwanatha, E. Yen-Pon, V. Robert, M. Beltramo, V. Piller, M. Alami, V. Aucagne, S. Messaoudi, *Chem. Sci.* **2018**, *9*, 8753.

17. An additional Trp residue was introduced at the C-terminus of this model peptide to facilitate HPLC monitoring and quantification of the different steps.

18. H. Kunz, S. Birnbach, *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 360; *Angew. Chem.* **1986**, *98*, 354.

19. Acetate deprotection was at first attempted on solid support using either N_2H_4 and MeONa. However, we could not drive the reaction to completion, leading to a mixture of acylated and deacylated hydroxyl moieties in addition to numerous unidentified co-products. For an example of on resin deacetylation, see: M. Liu, G. Barany, D. Live, *Carbohydrate Res.* **2005**, *340*, 2111.

20. To further optimize the enzymatic hydrolysis, the buffer and the amount of enzyme used were optimized to obtain faster and more economic cleavage conditions (see SI p S11).

21. a) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596; *Angew. Chem.* **2002**, *114*, 2708; b) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.

22. F. I. Auzanneau, M. Meldal, K. Bock, *J. Pept. Sci.* **1995**, *1*, 31.

23. M. J. Damha, P. A. Giannaris, S. V. Zabarylo, *Nucleic Acids Res.* **1990**, *18*, 3813.

24. a) S. Leon, R. Quarrell, G. Lowe, *Bioorg. Med. Chem. Lett*. **1998**, *8*, 2997; b) J. Kress, R. Zanaletti, A. Amour, M. Ladlow, J. G. Frey, M. Bradley, *Chem. Eur. J.* **2002**, *8*, 3769; c) A. Chiva, D. E. Williams, A. B. Tabor, H. C. Hailes, *Tetrahedron Lett.* **2010**, *51*, 2720.

25. J. Deere, R. F. De Oliveira, B. Tomaszewski, S. Millar, A. Lalaouni, L. F. Solares, S. L. Flitsch, P. J. Halling, *Langmuir* **2008**, *24*, 11762.

26. a) J. Vagner, G. Barany, K. S. Lam, V. Krchnak, N. F. Sepetov, J. A. Ostrem, P. Strop, M. Lebl, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8194; b) Z. P. Gates, A. A. Vinogradov, A. J. Quartararo, A. Bandyopadhyay, Z. N. Choo, E. D. Evans, K. H. Halloran, A. J. Mijalis, S. K. Mong, M. D. Simon, E. A. Standley, E. D. Styduhar, S. Z. Tasker, F. Touti, J. M. Weber, J. L. Wilson, T. F. Jamison, B. L. Pentelute, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5298.

27. As an alternative to Rink linker, we systematically evaluated two highly acid-labile linkers classically used in Fmoc-SPPS, but whose susceptibility to acid cleavage under aqueous conditions has never been reported to our knowledge (See SI p S26-S28): a trityl-based linker for carboxylic acids and Sieber's amide linker. The trityl linker was found to be unstable under NCL conditions, thus inappropriate, and Sieber's linker required 10% aqueous TFA for an efficient cleavage.

[28] One could imagine that solid-supported folding could not be achieved for some targets (in particular oligomeric proteins), and that the non-denaturing conditions used for linker cleavage could, in some cases, lead to misfolded species that could potentially hinder the accessibility of the phosphatase to its phosphoester substrate, and thus prevent or slow down cleavage. In such case, denaturing TFA cleavage followed by in solution folding prior to enzymatic linker cleavage in solution would be a valuable option. This is out of the scope of the present article, and will be explored in due course.

29. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879; *Angew. Chem.* **2009**, *121*, 10063.

30. J. W. Perich, E. C. Reynolds, *Synlett*, **1991**, *8*, 577.

31. A. Alouane, R. Labruère, T. Le Saux, I. Aujard, S. Dubruille, F. Schmidt, L. Jullien, *Chem. Eur. J.* **2013**, *19*, 11717.

32. N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046.

33. M. F. Debets, S. S. van Berkel, S. Schoffelen, F. P. J. T. Rutjes, J. C. M. van Hest, F. L. van Delft, *Chem. Commun.* **2010**, *46*, 97.

34. Various conditions were evaluated to carry out the SPAAC reaction before we succeeded in driving it to completion. (see SI p S54).

35. It is important to note that the "enzyme units" (U) value used to characterize the activity of a given enzyme, and herein used as a basic comparison between the enzyme used, is not comparable from one enzyme to another, and does not correlate to the enzyme concentration, which is extremely difficult to evaluate precisely. For

example, commercial lambda phosphatase activity is 400,000 U/mL whereas *E. coli* β-galactosidase is only 500-7,500 U/mL.

36. G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chem. Int. Ed.* **2011**, *50*, 7645; *Angew. Chem.* **2011**, *123*, 7787.

37. K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto, A. Otaka, *ChemBioChem*. **2011**, *12*, 1840.

38. a) L. A. Erlich, K. S. A. Kumar, M. Haj-Yahya, P. E. Dawson, A. Brik, *Org. Biomol. Chem.* **2010**, *8*, 2392; b) K. Sato, A. Shigenaga, K. Kitakaze, K. Sakamoto, D. Tsuji, K. Itoh, A. Otaka, *Angew. Chem. Int. Ed.* **2013**, *52*, 7855; *Angew. Chem.* **2013**, *125*, 8009; c) K. Aihara, K. Yamaoka, N. Naruse, T. Inokuma, A. Shigenaga, A. Otaka, *Org. Lett*. **2016**, *18*, 596.

39. a) V. P. Terrier, H. Adihou, M. Arnould, A. F. Delmas, V. Aucagne. *Chem. Sci.* **2016**, *7*, 339; b) D. Lelièvre, V. P. Terrier, A. F. Delmas, V. Aucagne, *Org. Lett.* **2016**, *18*, 920; 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, C. Arnoult, *Mol. Hum. Reprod.* **2016**, *23*, 116; d) V. P. Terrier, A. F. Delmas, V. Aucagne, *Org. Biomol. Chem.* **2017**, *15*, 316; e) L. De Rosa, R. Di Stasi, L. D. D'Andrea, *Tetrahedron*, **2019**, *75*, 894; f) 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, D. Destoumieux-Garzón, *mBio*, **2019**, *10*, e01821; g) S. A. Abboud, V. Aucagne, *Org. Biomol. Chem.* **2020**, *18*, 8199; h) S. A. Abboud, E. Cisse, M. Doudeau, H. Bénédetti, V. Aucagne, *Chem. Sci.* **2021**, *12*, 3194.

40. S. W. Bi, P. Liu, B. P. Ling, X. A. Yuan, Y. Y. Jiang, *Chinese. Chem. Lett.* **2018**, *29*, 1264.

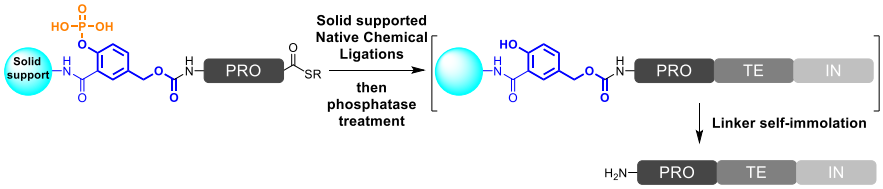
41. D. Veber, J. Milkowski, S. Varga, R. Denkewalter, R. Hirschmann, *J. Am. Chem. Soc.* **1972**, *94*, 5456.

42. S. K. Maity, M. Jbara, S. Laps, A. Brik, *Angew. Chem. Int. Ed.* **2016**, *55*, 8108; *Angew. Chem.* **2016**, *128*, 8240.

43. The overall yield is calculated based on the five synthetic steps (grafting, NCL, Acm deprotection, NCL, and enzymatic cleavage) and the final HPLC purification. The reduction of the isolated yield can be attributed to the small-scale purification, where in loss of product due to non-specific binding can be significant. In addition, we employed a strict signal level-based peak collection method, for maximizing purity over yield.

44. Note that we found important to conduce all the SPCL assembly process (including SPAAC-based immobilization, NCL reactions, Acm deprotection and enzymatic relase) without delay between synthetic steps in order to provide with a reproducible enzymatic release yield, this being attributed to unwanted polypeptide aggregation upon storage as advocated in other SPCL-based protein chemical syntheses. See for example references [4b] and [4i].

Entry for the Table of Contents



We herein introduce a chemo-enzymatic approach for the synthesis of proteins through solid-supported chemical ligation, based on a tailor-made phosphatase-cleavable linker. We show that the size of the enzyme used directly correlates with the efficiency of the release of the synthesized protein. The method was applied to the synthesis of a 15 kDa polypeptide, the longest sequence ever synthesized through solid phase native chemical ligation.

Institute and/or researcher Twitter usernames:
@CBM_UPR4301 / @vvinz