UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

Centre de Biophysique Moléculaire, Orléans

THÈSE présentée par :

Alba CASAS MORA

soutenue le : **12 décembre 2017**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : Chimie

A *catch*-and-*release* purification method to simplify the synthesis of cysteine-rich peptides

THÈSE co-dirigée par :

Vincent AUCAGNE	Chargé de recherche CNRS, CBM, Orléans
Agnès DELMAS	Directeur de recherche CNRS, CBM, Orléans

RAPPORTEURS :

Laurent DEVEL	Chercheur, CEA-Saclay
Sophie FAURE	Chargé de recherche CNRS, ICCF Clermont-Ferrand

JURY :

Josef HAMACEK	Président du Jury, Professeur, Université d´Orléans
Vincent AUCAGNE	Chargé de recherche CNRS, CBM, Orléans
Agnès DELMAS	Directeur de recherche CNRS, CBM, Orléans
Laurent DEVEL	Chercheur, CEA-Saclay
Sophie FAURE	Chargée de recherche CNRS, ICCF Clermont-Ferrand
Pascal KESSLER	Chercheur, CEA-Saclay

“Los científicos dicen que estamos hechos de átomos,

pero yo creo que estamos hechos de historias”

Eduardo Galeano

A mis padres,

que me han enseñado a mirar hacia el horizonte.

En memoria de Josefa y Carlitos.

Acknowledgements

Le travail de thèse présenté dans ce manuscrit a été réalisé au Centre de Biophysique Moléculaire (CBM), unité propre du Centre National de la Recherche Scientifiques (CNRS).
Je voudrais remercier le CNRS et la Région Centre pour avoir financé mes travaux de thèse.
Je souhaiterais également remercier la directrice du laboratoire, le docteur Eva Jakab Toth pour m'avoir permis d'effectuer cette thèse au CBM.

Je tiens à remercier les Drs. Sophie Faure et Laurent Devel d´avoir accepté d´être les rapporteurs de ce manuscrit. Je remercie également le Dr.Pascal Kessler et le Pr. Josef Hamacek d´avoir accepté de faire partie du jury de la soutenance de ma thèse.

Je voudrais remercier très chaleureusement le Dr. Vincent Aucagne, mon co-directeur de thèse pour m'avoir transmis plein de connaissances ainsi que l´importance du travail rigoureux et bien fait. Je le remercie pour sa grande aide à l´heure de la rédaction et surtout pour ne pas avoir perdu la patience pendant ces trois dernières années. Je tiens à remercier le Dr. Agnès Delmas, ma co-directrice de thèse, pour m´avoir fait participer à ce projet de recherche stimulant et de m'avoir fait bénéficier de sa grande expérience scientifique.

Je voudrais remercier le Dr. Véronique Piller pour son soutien scientifique (et moral !) pendant la rédaction de ce manuscrit, ainsi que pour sa précieuse amitié. J´adresse un grand merci à Dominique Lelièvre pour sa complicité pendant toute la thèse, sa sympathie (grâce à lui je connais la danse de la « sardine ») et pour ses très bons conseils peptidiques. Je souhaite aussi remercier le Dr. Mehdi Amoura pour sa bonne humeur et sa capacité pour générer une très bonne ambiance de travail. Je tiens à remercier les membres présents et passés de l´équipe que j´ai eu le plaisir de côtoyer : Philipe Marceau, Vishwanath, Guillaume Jousset et les Drs. Aline Mongis, Mathieu Galibert, Nathan Stanley et Friedrich Piller. Je ne peux pas oublier de remercier le Synthétiseur de peptides et l´HPLC qui ont fait partie de mes pensées à chaque instant de ces années, ainsi que toutes les colonnes que j´ai vu passer dans mes mains (R.I.P).

Un grand merci à Cyril Colas pour l'acquisition des spectres de masse haute résolution et à Guillaume Gabant pour l'acquisition des spectres de masse. Je voudrais aussi remercier Hervé Meudal pour l'acquisition de spectres RMN.

Je veux remercier en particulier Yannick Berteaux, Justo Torres et Martial Bruneau pour leurs disponibilités à m´aider. Heureusement que vous êtes ici pour faire fonctionner le CBM !

Je n´oublie pas le reste des membres du CBM que j´ai eu le plaisir de côtoyer pendant ces trois années.

Maintenant, je m´adresse à mes collègues de bataille, Dounia, Maamar, Sophie et Patrick, avec qui j´ai partagé toute cette folie de thèses. Il n'y a pas assez de mots pour décrire tout

ce qu´on a vécu ensemble (bêtises incluses). Un merci spécial à mi amiga Dounia, tu as été présente à chaque instant. Impossible oublier le début de cette amitié que a poursuivre avec plein des aventure surréalistes. Ton amitié est sans doute mon meilleur cadeau de thèse.

Merci à tous mes amies que j´ai eu le plaisir de rencontrer à Orléans (Anne, Marcello, Debo, Eline, Celia, Sarah…) avec qui j´ai partagé de très bons moments et qui ont fait de cette expérience une véritable aventure. On se voit à la Galice pour boire quelques « Estrellas » et danser au rythme galicien de « SES ».

A vosotros, mi gente, me siento en el deber de agradeceros vuestro apoyo de una forma más familiar.

Papá, Mamá; no se me ocurre mejor manera de agradeceros todo lo que habéis hecho por mí que habiendo terminado esta tesis. Solo vosotros sabéis el sacrificio tan enorme que ello ha supuesto, y no solo para mí, sino también para vosotros. Esos botecillos de pisto y el ingenio de papá me han salvado en más de una ocasión. Gracias por hacer todo más simple y por vuestro apoyo y amor incondicional. Os quiero

A Prado y Chus, que desde siempre han compartido mis logros y alegrías, gracias. A mi hermano, pionero en estos viajes migratorios a Francia, gracias por creer en mí. A mi abuelo Julio, por nuestra complicidad y por esa capacidad tan bonita que tiene de hacerme sentir especial y en general, a mi familia, gracias a todos por vuestro apoyo, y buen humor.

Por último, A Carba, por ser o meu mellor compañeiro durante estos últimos anos. Un anaquiño de esta tesis pertenececheche, graciñas.

A todos los que un día compartieron está pasión conmigo, gracias.

Alba

Abbreviations

AA: amino acid

Acm: acetamidomethyl

Boc: *tert*-butyloxycarbonyl

Bzl: benzyl

CAPS: *N*-cyclohexyl-3-aminopropanesulfonic acid

CPG: controlled pore glass

CuAAC: copper(I)-catalyzed azide alkyne cycloaddition

DBU: 1,8-diazabicyclo[5.4.0]undec-7-ene

DCC: *N*,*N*´-dicyclohexylcarbodiimide

DIC: *N*,*N*´-diisopropylcarbodiimide

DCM: dichloromethane

DMAP: 4-(dimethylamino)pyridine

DMF: *N*,*N*-dimethylformamide

DMSO: dimethyl sulfoxide

DODT: 3,6-dioxa-1,8-octanedithiol

DTT: DL-dithiothreitol

EDCI: *N*-ethyl-*N*'-(3-dimethylaminopropyl)carbodiimide

EDT: ethanedithiol

EDTA: ethylenediaminetetraacetic acid

EPL: expressed protein ligation

Equiv: equivalent

ESI: electronspray ionisation

DIEA: *N*,*N*-diisopropylethylamine

Fmoc: 9-fluorenylmethoxycarbonyl

Fmoc-PAL-OH: 5-[3,5-dimethoxy-4-(9-fluorenylmethoxycarbonylamino)phenoxy]pentanoic acid

Fmoc-TTDS-OH: *N*1-(9-fluorenylmethoxycarbonyl)-1,13-diamino-4,7,10-trioxatridecan-succinamic acid

Fmoc-Rink-OH: 4'-{(*R*,S)-α-[1-(9-fluorenylmethoxycarbonylamino]-2,4-dimethoxybenzyl}-phenoxyacetic acid

Abbreviations

GSH : glutathione

GSSG: oxidized glutathione (disulfide)

HATU: 1-[*bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

HBTU: 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate

HCTU: 2-(6-chloro-1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate

HEPES: 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid

HOBt: 3*H*-[1,2,3]-triazolo[4,5-*b*]pyridin-3-ol

HPLC: high pressure liquid chromatography

HRMS: high resolution mass spectrometry

***i*Bu**: isobutyl

***i*Pr**: isopropyl

LCMS: high pressure liquid chromatography coupled with mass spectrometry

MALDI: matrix-assisted laser desorption ionization

MesNa: sodium 2-mercaptoethanesulfonate

MPAA: 4-mercaptophenylacetic acid

MS: mass spectrometry

n.d: non determined

NCL: native chemical ligation

NMP: 1-methyl-2-pyrrolidone

NMR: nuclear magnetic resonance

PEG: poly(ethyleneglycol)

PEGA: acrylamide-PEG copolymer

PyAOP: 7-azabenzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate

PyBOP: benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

SPAAC : strain-promoted alkyne-azide cycloaddition

SPCL: solid phase chemical ligation

SPPS: solid phase peptide synthesis

***t*Bu**: *tert*-butyl

TCEP: *tris*(2-carboxyethyl)phosphine

TFA: trifluoroacetic acid

Abbreviations

TFE: 2,2,2-trifluoroethanol

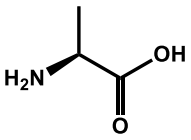
TIS: triisopropylsilane

TOF: time of flight

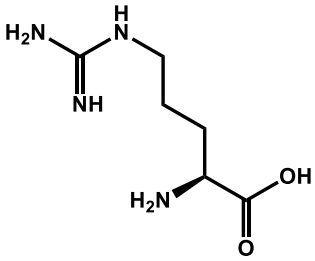
TRIS: *tris*(hydroxymethyl)aminomethane

UV: ultra-violet

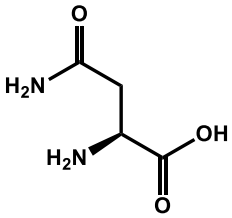
Proteinogenic amino acids



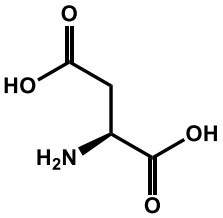
Ala (A): alanine



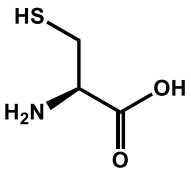
Arg (R): arginine



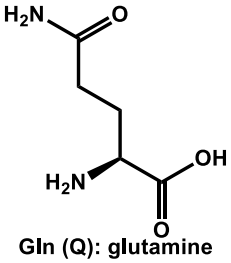
Asn (N): asparagine



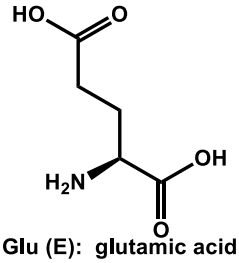
Asp (D): aspartic acid



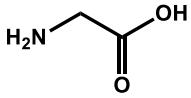
Cys (C): cysteine



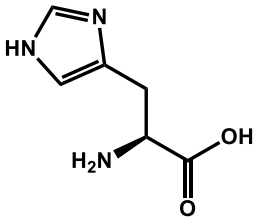
Gln (Q): glutamine



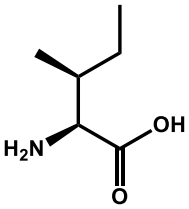
Glu (E): glutamic acid



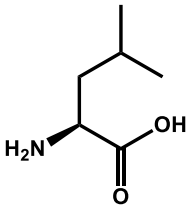
Gly (G): glycine



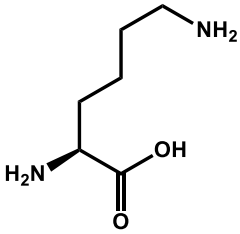
His (H): histidine



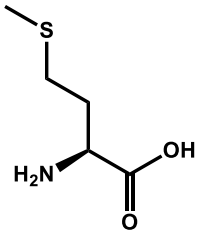
Ile (I): Isoleucine



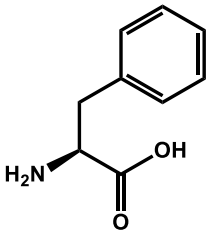
Leu (L): Leucine



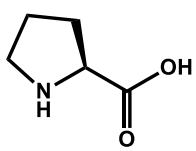
Lys (K) : Lysine



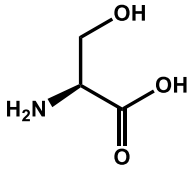
Met (M): methionine



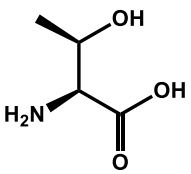
Phe (F): phenylalanine



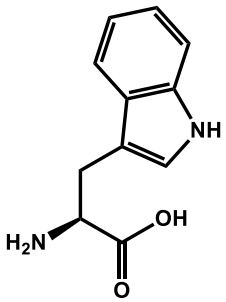
Pro (P): proline



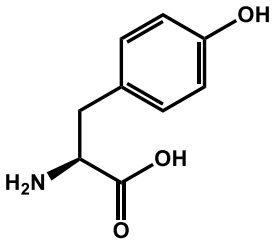
Ser (S): serine



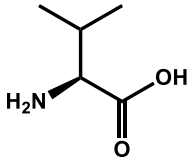
Thr (T): threonine



Trp (W): tryptophan



Tyr (Y): tyrosine



Val (V): valine

Table of contents

Table of contents

Table of contents

Table of contents

Table of contents

Table of contents

General Introduction

General introduction

Peptides and proteins are macromolecules involved in many biological processes that show a big range of different functions and structural roles. For years, the researchers have showed a growing interest in the understanding of the mechanism of action of these compounds. The study of new technologies for their productions thus has been always a priority. Their production from recombinant DNA originating from many organism, has allowed the rapid procurement of these compounds, for their further study and commercialization. This procedure allows to easily producing simplest medium-to-large size peptides and proteins, however it is associated with some limitations such as the use of only proteinogenic amino acids or the difficulties in incorporating post-translational modifications.

As a good alternative, their chemical synthesis offers some advantages and gives the possibility to incorporate any kind of modification in their atomically structure. The most commonly used method for their chemical synthesis is the solid phase peptide synthesis (SPPS). Unfortunately, this method is limited to the synthesis of less than 50 amino acids and required in many cases the realization of a chromatographic purification which not always allows to obtain the pure peptide.

In the last two decades, the industry has showed a growing interest in the production of disulfide-rich peptides (DRPs) bioactive naturally-occurring peptides that due to their high biological and chemical stability, combined with their ability to modulate therapeutically-relevant targets, are considered as promising drug candidates and pharmacological tools.

Similarly as other peptides, the SPPS of DRPs can be difficult and can result in the formation of multiple acetylated truncated peptides (the main co-products of SPPS) which can be complex to separate by standard chromatographic HPLC.

To avoid the realization of these costly and time-consuming HPLC purifications and with the aim to obtain pure peptides, it has been developed an alternative non-chromatographic "*catch*-and-*release*" purification strategy based on the use of N-terminal linkers. Unfortunately, the existing linkers are hardly compatible with the particular reactivity of unprotected cysteines and thus with the purification of these peptides.

It is in this context that takes place the works of my thesis. The main aim has been to adapt the mentioned *catch*-and-*release* strategy to the synthesis and purification of these small-to-medium DRP through the use of the Native chemical ligation (NCL) a chemoselective reaction which occurs between two peptide-fragments in aqueous media at neutral pH.

Because the synthesis of long DRPs is also a real challenge, we have designed our initial strategy for it to be expanded in the future to the synthesis of longer DRPs through the assembly of multiples unprotected fragments via solid phase NCLs which can avoid the costly and time-consuming HPLC purifications.

The first part of chapter 1 consists on a brief description of the main strategies for the chemical synthesis of peptides, either in solution or in solid phase. In the second part of this chapter a detailed revision of the existing N-terminal linkers for the "*catch*-and-*release*" purification of synthetic peptides has been done. The result of this study will be submitted as a review article. Chapter 2 is focused on the design and synthesis of the N-terminal linker: Boc-Cys(Trt)-Dtph-OH for the synthesis and catch-and-release purification of complex DRPs. A detailed study of its stability and cleavage conditions in solution has been also included. In chapter 3 we have proposed a new heterobifunctional acid-thioester linker to functionalize the purification solid support that will serve to anchor the peptides to be purified, decorated with our N-terminal linker. The application of "*catch*-and-*release*" purification of a very difficult DRP has been described in chapter 4. Finally, chapter 5 includes a preliminary application of such methodology to the synthesis and purification of a very long (77 amino acids) biologically-relevant cysteine-rich miniprotein. The experimental data are detailed in chapter 6. The conclusions and perspectives have been included at the end.

Chapter 1

Chapter 1. Literature review

Part 1. Chemical synthesis of peptides

1. Introduction

Over the years, the production of peptides and proteins in bacteria and other microorganisms from recombinant DNA originating from many organism, has allowed the rapid procurement of these compounds, for their further study and commercialization. At least 30% of the commercialized peptides and proteins were produced through the modification of the bacterial genome of *Escherichia coli.*[1] In general, once the procedure is optimized, the production of simplest medium-to-large size peptides and proteins is easy, fast, and does not require complex post-treatments. In contrast, the production of more complex peptides incorporating post-translational modifications, can be difficult and in some cases suffers from proteolytic degradation.[1,2] In this context, the chemical production of these peptides has emerged as a good alternative. The boom of peptide chemistry has been favored by the latest advances in chemical methods as well as the growing interest in the synthesis of more complex and sophisticated peptides, which are not possible to be produced by recombinant approaches. The chemical synthesis of peptides is open to any kind of modification in their atomical structure, and offers some remarkable advantages. For example, the possibility to incorporate non-proteinogenic amino acids, the modification of the backbone structure or the addition of post-translational modifications such as phosphorylation or glycosylation. Since last years, the world of synthetic peptides has entailed a big expansion, and nowadays, numerous research topics are focused on it. However, the breadth of possibilities offered by synthetic peptide chemistry is vast and it is still necessary to continue researching to overcome the limits of the current knowledge.

In the following pages, the most relevant strategies for the chemical synthesis of peptides will be described and how their limitations have led, among other things, to the development of my thesis project.

2. Solid phase peptide synthesis

In 1963, Merrifield developed the stepwise solid phase peptide synthesis (SPPS).[3] This method introduced a breakthrough in the way of synthesizing peptides, facilitating their production and limiting the number of purification steps. In the SPPS strategy, the key step is the immobilization of the C-terminal amino acid: the carboxylic acid moiety of the first amino acid is anchored onto an insoluble solid support or resin through a removable linker and then the peptide is build up. The elongation onto the resin allows to eliminate the excess of reactants and co-products by a simple filtration. Next amino acids are then coupled one by one from their C-terminal to the N-terminal of the previous amino acid. The amino acids have their side chains and N-terminal functions protected with different protecting groups. During the elongation, the N-terminal protecting group is removed before the coupling. The iterative peptide elongation thus consists of the repetition of several cycles of deprotection and coupling. When the elongation is finished, the side chain protecting groups are finally removed and at the same time the peptide is released from the resin (Figure 1).

During the elongation step, some incoming amino acids may be incompletely coupled to the growing chain, resulting in the accumulation of here referred, deleted peptides, which are difficult to separate from the target peptide by usual chromatographic methods (RP-HPLC) A convenient strategy consists in adding after each coupling step an excess of an acetylating reagent (in general acetic anhydride), in order to mask the reactivity of the non-coupled Nα-terminal amino acid. Thus, at the end, the peptide of interest is contaminated with the acetylated truncated peptides, which are in general, easier to separate from the expected peptide than the deleted sequences.

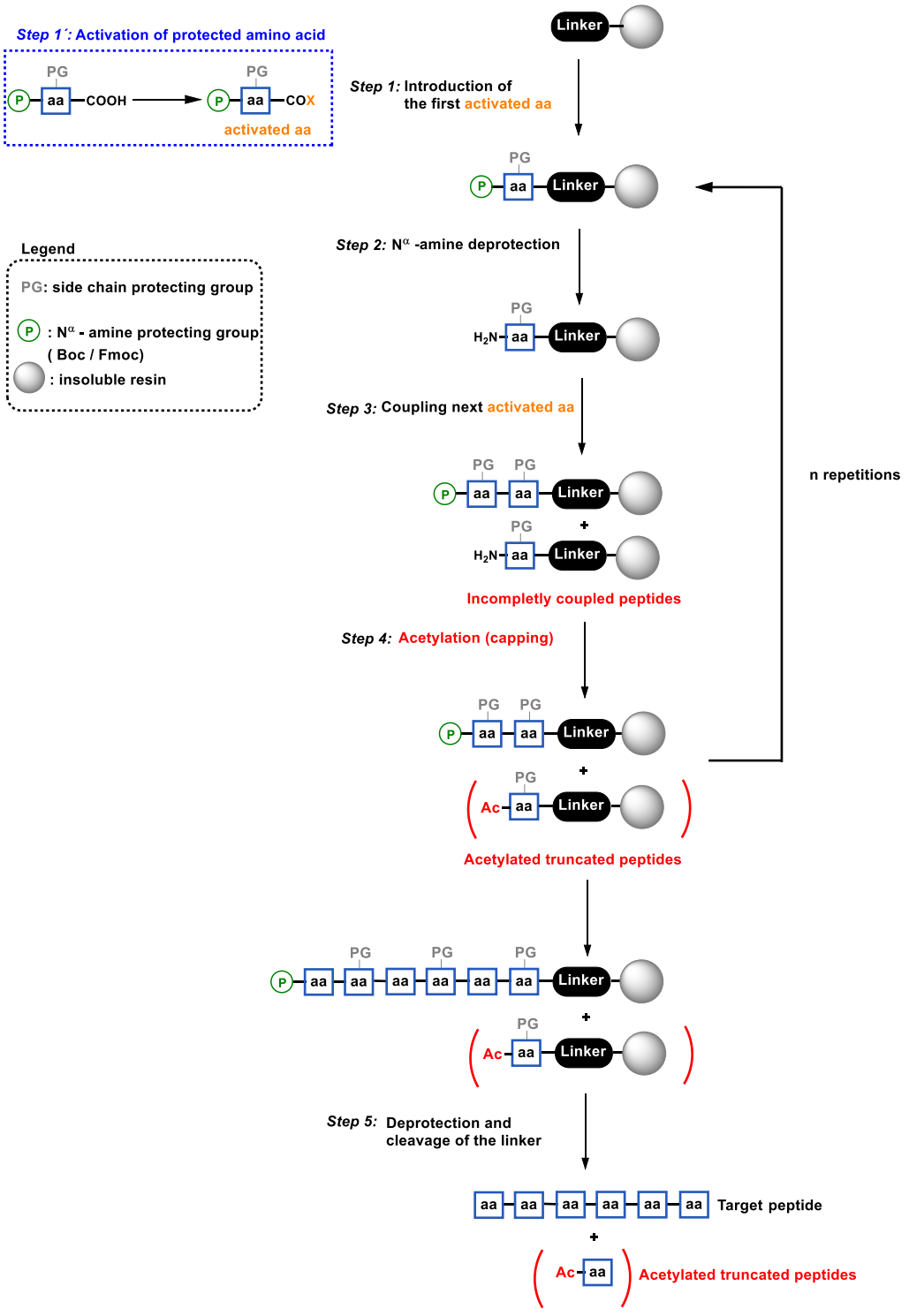


Figure 1. Merrifield's SPPS strategy.

Two different SPPS strategies have been proposed to date depending on the *N*α-protecting group, the Boc/Bzl or the Fmoc/tBu strategy. The pioneering *tert*-butoxycarbonyl (Boc)/benzyl (Bzl) strategy was developed by Merrifield.[3] In this approach, the side chains and N-terminal moieties are protected by Bzl and Boc acid-labile protecting groups. During the elongation, the Boc group is removed from the N-terminal amine upon treatment with

trifluoroacetic acid (TFA). However, the deprotection and cleavage of the peptide require strongly acid conditions with hydrogen fluoride (HF) which is highly corrosive and toxic and limits the application of such a strategy.

Nowadays, the most commonly used is the Fmoc/*t*Bu strategy developed by Carpino and Han.[4] They proposed to use the base-labile 9-fluorenylmethyloxycarbonyl or Fmoc group to protect the α-amines which can be deprotected upon piperidine treatment. On the other hand, the side chain protecting groups are removed with acid-labile TFA. One important advantage is the possibility of quantifying the yield of each coupling step spectrophotometrically, through the measure of the fulvene-piperidine adduct formed during the removal of the protecting group (Figure 2).

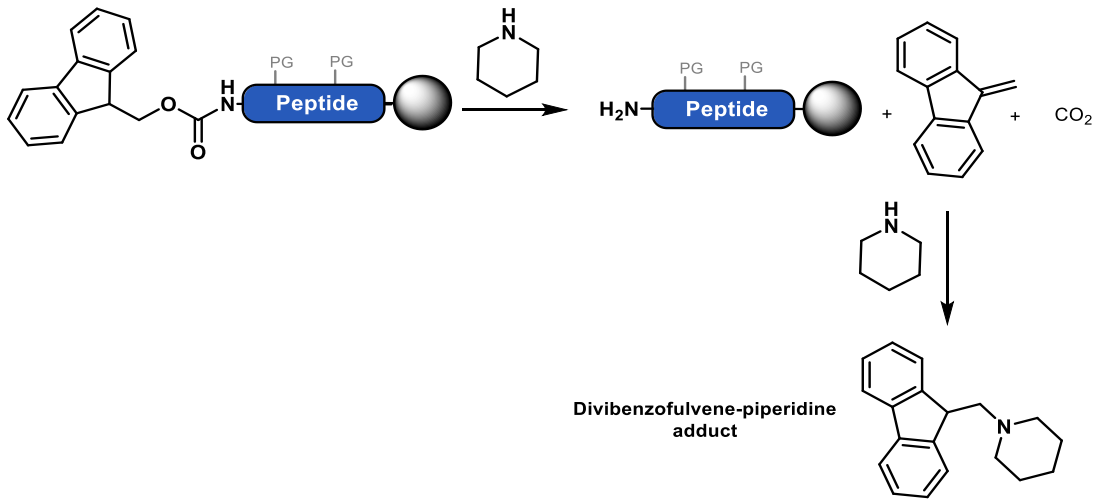


Figure 2. Removal of the Fmoc-protecting group.

Thanks to the latest advances in chemical synthesis of peptides, such as the introduction of new solid supports and linkers,[5] or the development of novel orthogonal protecting groups[6] and coupling agents,[7] automatic SPPS has been implanted nowadays as the method of choice for the synthesis of medium-sized peptides.[8] However, for the synthesis of longer peptides (usually more than 50 amino acids), the application of the Merrifield´s step by step strategy -in a routinely way- usually becomes complicated and results in incomplete couplings and poor yields. In addition, for longer peptides, the accumulation of acetylated peptides is more important and a final reverse-phase HPLC purification is usually necessary to remove them and yield the pure peptide. For some peptides, the similar composition between such impurities and the peptide of interest may result in a co-elution after HPLC thus impairing the purification and homogeneity of the final peptide.[9] To achieve a

successful purification, innovative tools and purification strategies have been developed. One of them is the non-chromatographic "*catch*-and-*release*" purification strategy. This particular topic will be developed in the second part of this chapter.

As the ability of the Merrifield´s SPPS to routinely produce peptides can be compromised when the peptide size exceeds 50 amino acids, novel chemical methodologies have been proposed to achieve long peptides. In the following pages, those methodologies will be briefly described.

3. Chemoselective ligation reactions of peptidyl fragments

3.1. Peptide Chemical ligation

The concept of chemical ligation was introduced by Schnölzer and Kent in the 90s,[10] and makes reference to the chemo- and regioselective coupling between two mutually reactive chemical groups of two peptidyl fragments in solution. The ligation of unprotected fragments, which in general benefits from good solubility and handling, is possible thanks to the chemoselectivity of the reaction.[11] Commonly, the fragments are synthesized by SPPS and then coupled in solution to create a covalent linkage (Figure 3). When the resulting covalent linkage is an amide bond, the chemical ligation is referred to as native ligation.

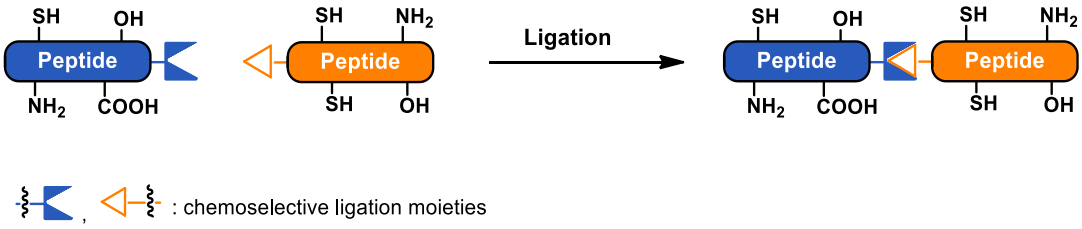


Figure 3. Principle of the chemical ligation between two unprotected-peptide fragments.

This approach has constituted a revolutionary advance in the way of synthesizing peptides, and numerous chemical reactions between chemoselective ligation moieties have been reported to date. The major issue for the researchers has unquestionably been the development of specific chemical groups that could react chemoselectively. In this manuscript, only the reactions resulting in a native peptide bond have been described.

3.2. The Native Chemical Ligation (NCL) reaction

3.2.1. Seminal works

Two works have contributed and inspired the development of the Native Chemical Ligation (NCL) reaction. The first one was the work of Wieland in 1953,[12] who reported for the first time the "thiol/thioester and the *S*-to-*N* acyl shift reaction".[13] The reaction between a cysteine and an α-thioester amine, at neutral pH in water, led to the formation of the native amide bond of the Val-Cys-OH dipeptide (Figure 4).

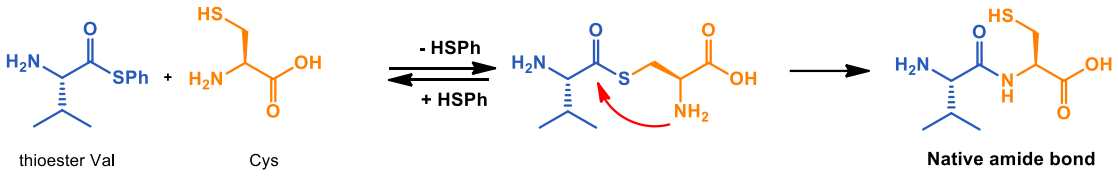


Figure 4. Mechanism of valinyl-cysteine dipeptide formation.

The other work that inspired Kent for the development of the NCL, was the thiol capture strategy described by Kemp and collaborators in the 1980.[14,15] Their strategy is based on the use of an auxiliary template (6-hydroxy-4-mercaptodibenzofuran) which allows the coupling of the two peptidyl fragments by facilitating the thiol-exchange between the thiol moiety of the auxiliary and the disulfide bond of the C-terminal peptide fragment (Figure 5).

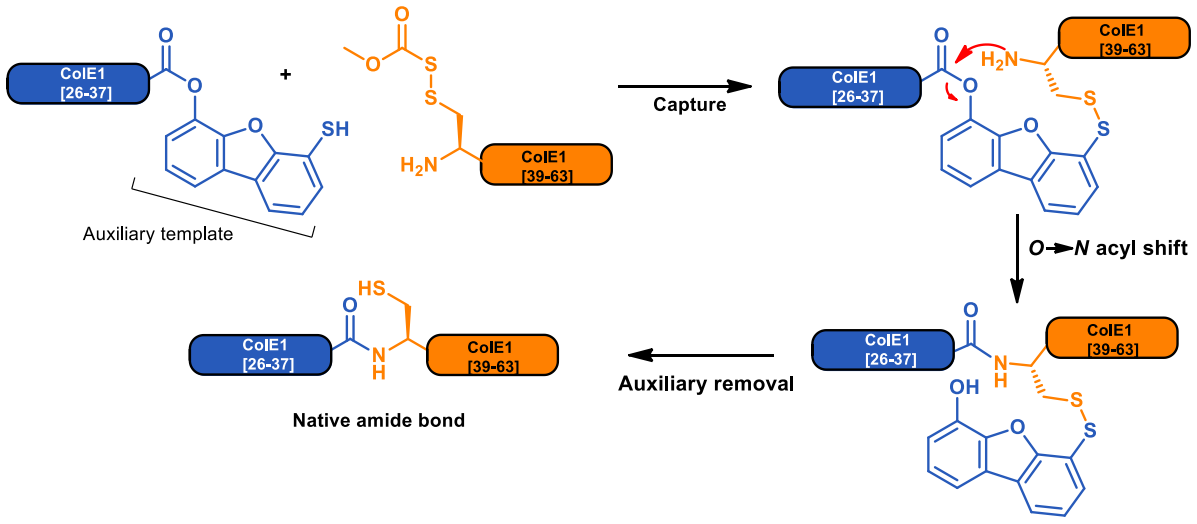


Figure 5. Native ligation promoted by the 6-hydroxy-4-mercaptodibenzofuran auxiliary, for the synthesis of a ColE1 protein.[14,15]

3.2.2. Principle of the NCL

In 1994, Kent[16] proposed the NCL reaction, which was a revolution in the world of peptide chemistry and especially in the synthesis of cysteine-containing peptides. NCL occurs between a C-terminal thioester peptide and a cysteine at the N-terminus of a peptide (cysteinyl peptide) at neutral pH (6 to 8) in water (Figure 6). It is completely adapted and compatible with the synthesis of unprotected cysteine-rich peptides. The mechanism begins with the formation of an intermediate thioester through a reversible trans-thioesterification reaction between both the C-terminal thioester and the N-terminal cysteinyl peptide. This intermediate is rapidly and spontaneously rearranged into a five-membered ring *via* an intramolecular *S*-to-*N* acyl shift to create a stable native amide bond (Figure 6).[16,17] The usefulness of the NCL was demonstrated by Kent with the synthesis of a human interleukin protein of 72 amino acids.[16]

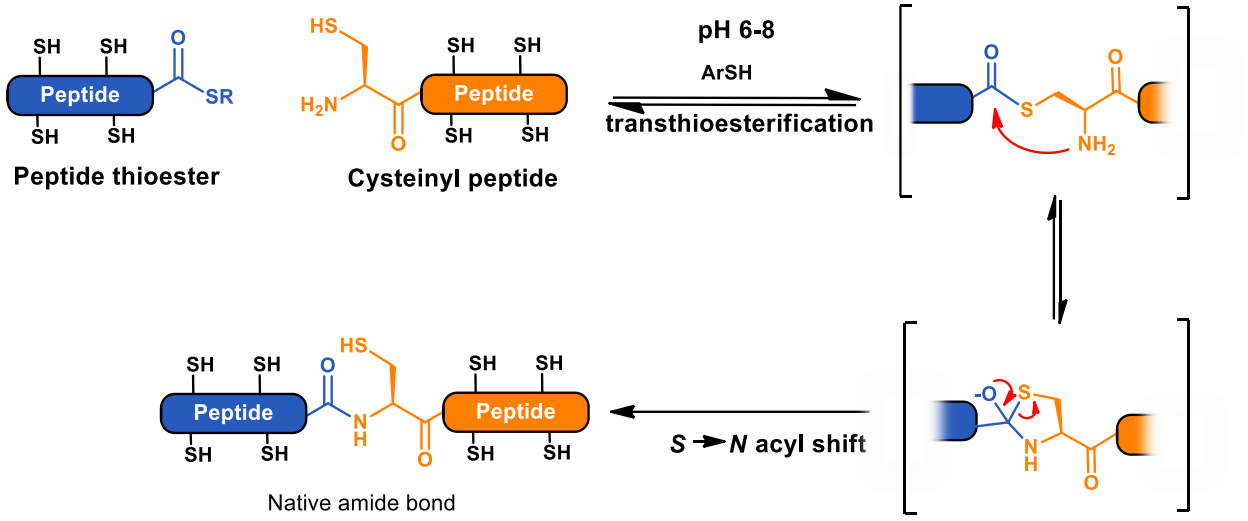


Figure 6. Mechanism of the native chemical ligation reaction.

Nowadays, NCL is established as the method of choice for the total chemical synthesis of many peptides and proteins. For its occurrence, the presence of an N-terminal cysteinyl peptide is needed. However, the abundance of cysteine in proteins is not very high (around 2%). Thus in order to expand the synthesizable targets, other methodologies alternative to the NCL which also result in the formation of a native amide bond have been proposed. Even if they will be not detailed in this thesis, some of them deserve to be mentioned. They are the condensation of α-ketoacids and N-alkylhydroxylamines (or *KAHA reaction*),[18] and the application of the Staudinger reaction.[19]

Extensions of the NCL reaction are for example the realization of a post-ligation desulfurization, proposed by Wan and Danisheksky,[20] and the expressed protein ligation (EPL) reported by Muir,[21] who applied the NCL to the ligation of two peptide fragments which were produced through recombinant strategies.

3.3. Expansion to successive native chemical ligations

As mentioned above, "in routine" SPPS of peptides of more than fifty amino acids is complicated or not possible, and requires specific optimizations. On the other hand, the ligation of two peptidyl fragments can yield peptides until one hundred amino acids. To increase the size of synthesizable peptides, new strategies have been developed based on the successive ligations of multiple fragments. In general, those approaches require multiple purifications and complicated activations and deprotection steps, and they have been the focus of numerous studies.

3.3.1. Multiple segment ligation in the C-to-N direction

The ligation of peptide fragment in the C-to-N direction is an approach easily applicable. The main requirement for successive segment ligation in the C-to-N direction (Figure 7) is the orthogonal protection of the thiol- and/or amine functions of the N-terminal cysteine of the internal segment. A first NCL is carried out between the protected internal fragment (peptide 2, Figure 7) and the C-terminal fragment (peptide 3, Figure 7). Then, the protecting groups are removed and the second NCL continues with the ligation of the N-terminal fragment (peptide 1, Figure 7). Successive ligations and deprotection steps can be repeated to achieve longer peptides.

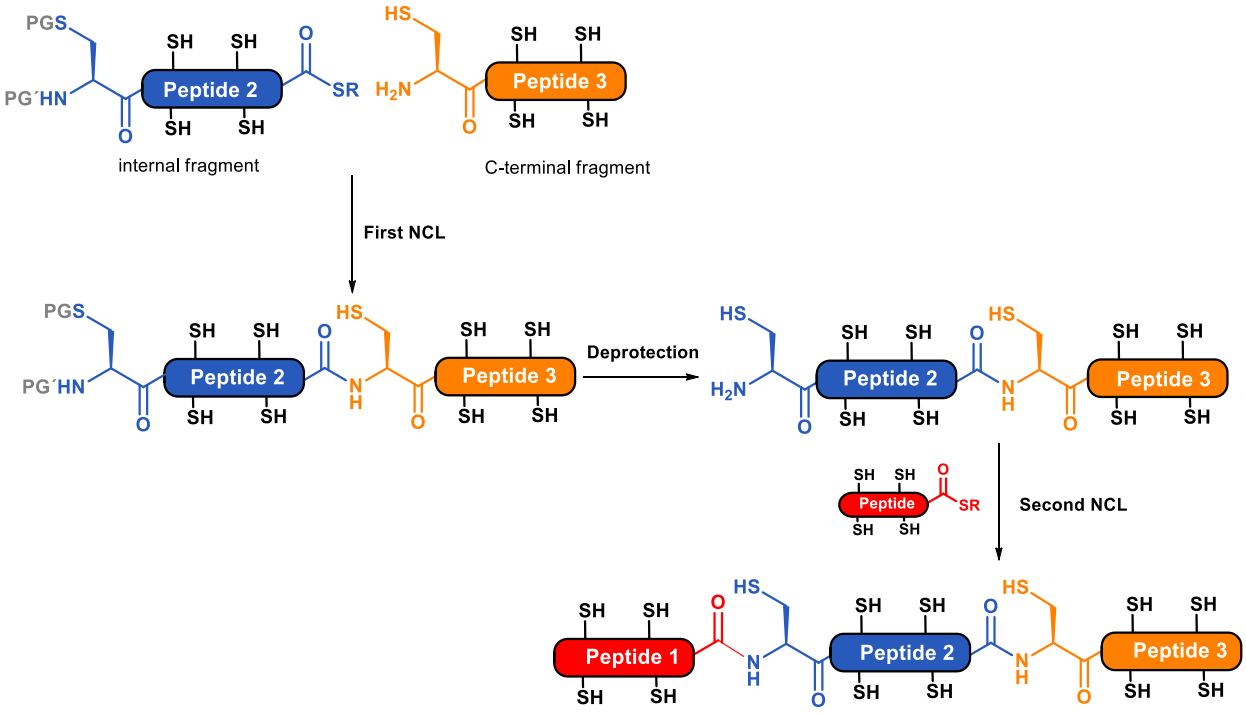


Figure 7. Multiple NCLs in the C-to-N direction.

The most important feature relies on the election of an adequate protecting group of the cysteine internal fragment, which should be not only resistant to the SPPS deprotection and cleavage conditions but also stable during the NCL reaction. In addition, the protecting group must be removed under conditions which do not affect the peptide integrity. In the next figure, the main protecting groups used in the context of the successive NCLs and the removal conditions of the amine and thiol functions are represented (Figure 8).

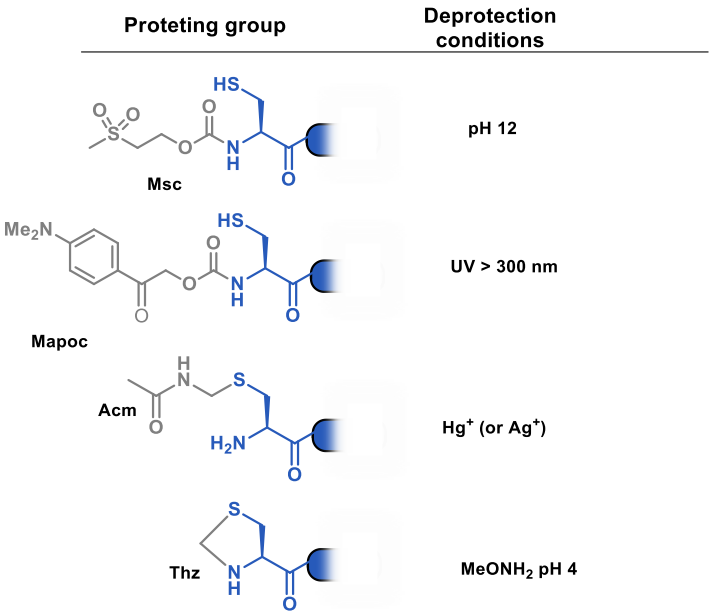


Figure 8. Representation of protecting groups of thiol and amine functions: methylsulfonylethyloxycarbony(Msc),[22,23] 4-(dimethylamino) phenacyloxycarbonyl (Mapoc),[24] acetamidomethyl (Acm),[25] and Thiazolidine (Thz).[26]

3.3.2. Multiple segment ligation in the N-to-C direction

The synthesis of polypeptides *via* successive NCLs of multiple fragments in the N-to-C direction (Figure 9) has been less exploited than in the C-to-N direction. The reaction occurs *via* a first NCL, followed by the generation of the reactive thioester function and ended with a second NCL (Figure 9).

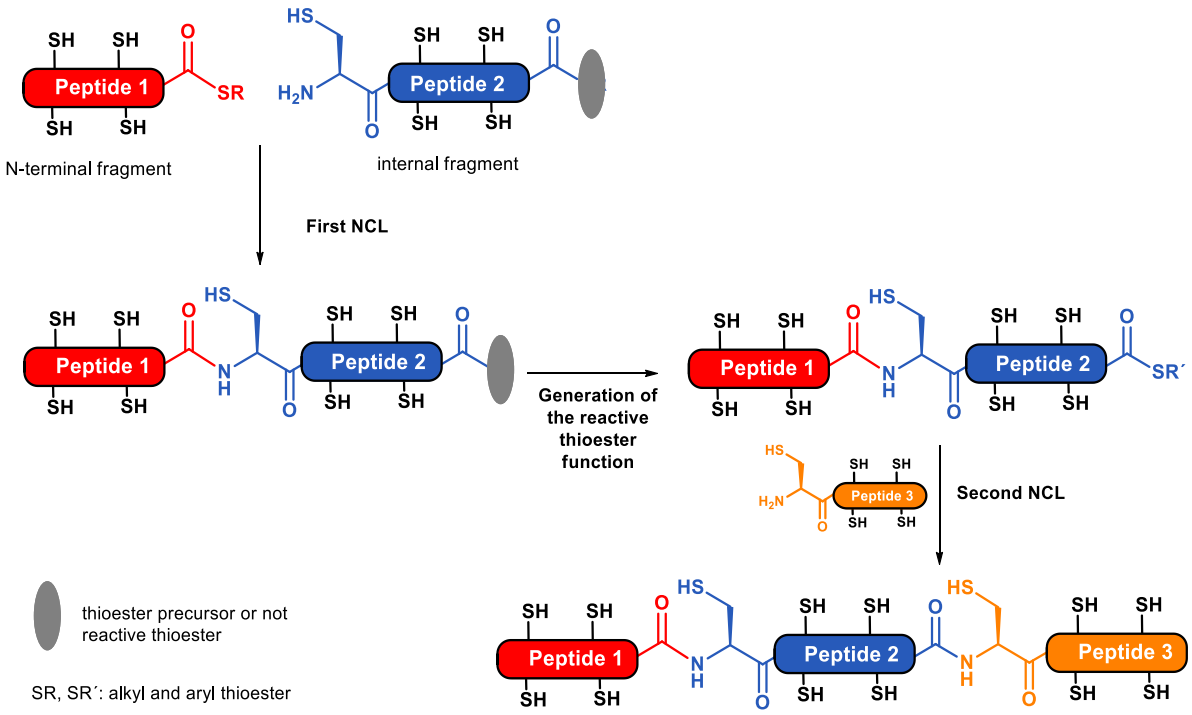


Figure 9. Multiple NCLs in the N-to-C direction.

Such a strategy requires a masked thioester or a precursor of the thioester moiety. Few examples of such strategies have been reported in the literature.

3.3.2.1. Ligation of peptides hydrazides

This method was proposed by Liu,[27] and is based on the utilization of a unprotected peptide containing a C-terminal hydrazide as a precursor of the thioester moiety. Peptide hydrazide can be oxidized under sodium nitrate at acidic pH to afford a peptide azide (Figure 10). A catalyzer thiol is then added under neutral pH for the thioesterification of the peptide,and the subsequent ligation with a cysteinyl peptide.

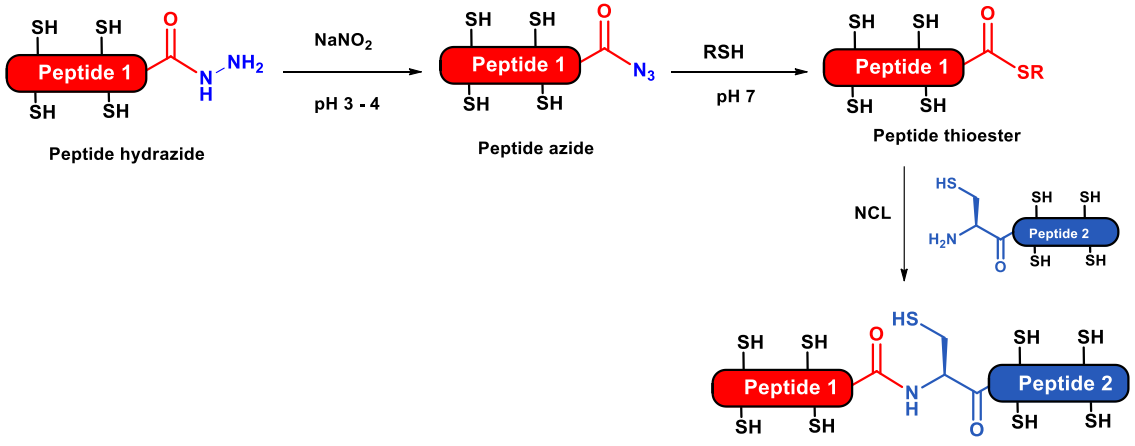


Figure 10. Ligation of peptides hydrazides.

3.3.2.2. The *N*-aminoacyl-*N*-sulfanylethylanilide (SEAlide) ligation

In 2009, Otaka[28] has reported the synthesis of a *N*-aminoacyl-*N*-sulfanylethylanilide (SEAlide) C-terminal peptide thioester which can conduct the acyl *N*-to-*S* thioesterification in acid media in presence of TCEP. The resulted compound can generate the peptide thioester in presence of a thiol (Figure 11).

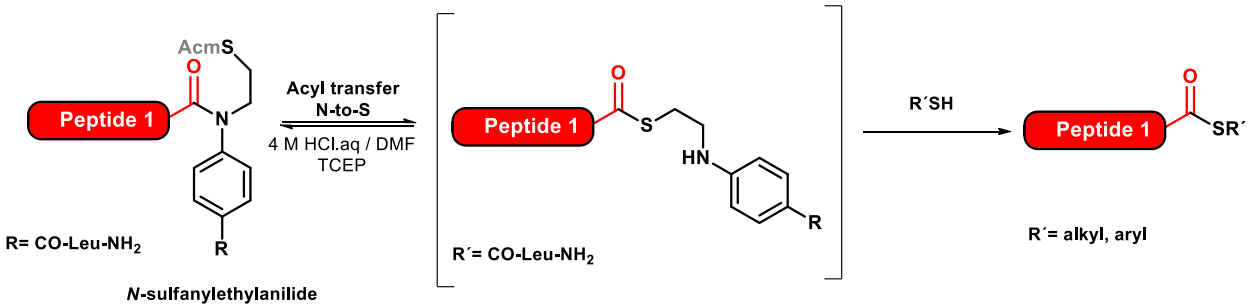


Figure 11. Synthesis of peptides thioester using SEAlide strategy.

Later, Otaka demonstrated that a peptide containing a SEAlide in C-terminal could be directly used in the NCL reaction to afford a native peptide bond.[29] When the thiol of the solid moiety is protected by an acetamidomethyl (Acm) group, the formation of the thioester is inhibited.

The properties of SEAlide peptide have been taking in advantage for the synthesis of a 162 mer peptide (Active monoglycosylated GM2-activator protein analogue) in five fragments via successive NCL in the N-to-C direction.[30]

3.3.2.3. Bis(2-sulfanylethyl)amino SEA Native Peptide ligation

Melnyk[31] and Liu[32] have described independently the synthesis of a *N*,*N*-Bis(2-sulfanylethyl)-amide (SEA) as thioester precursors for its use in the NCL (Figure 12). For the application of multiple NCLs, Melnyk proposed to mask the SEA group through the formation of a disulfide bond to afford a cyclic SEA. The masked SEA group is stable under the NCL conditions and can be also used directly into the next NCL in presence of TCEP. This strategy has been applied to the synthesis *via* NCL in three fragments.[33]

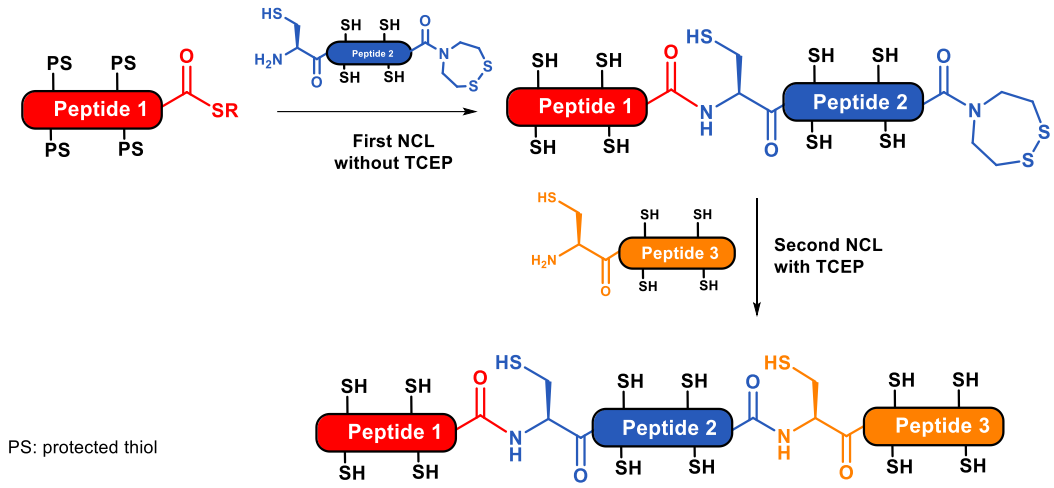


Figure 12. Three fragments NCL through SEA ligations in the N-to-C direction.[33]

Liu have also exploited the SEA ligation to the multiples NCLs.[34]

An alternative methodology is referred to as "*Kinetically Controlled Ligation*" (KCL) [35] and is based on the different reactivities of the aryl thioester ($^{\alpha}$thiophenylester) and the alkyl thioester ($^{\alpha}$thioalkylester), the later playing the role of selective thioester precursor moiety.

In practice, the synthesis of long proteins in solution usually combines the ligation in both directions. These approaches are illustrated by the synthesis of a 312-mer synthase (DapA)[36] or most recently the total synthesis of a DNA polymerase (Dpo49).[37]

4. Solid phase chemical ligation

The synthesis of long peptides *via* successive native chemical ligations is slow and requires multiple intermediate deprotection and activation steps, as well as laborious and time-consuming purifications, which may result at the end in poor yields. In order to overcome these limitations, it is possible to carry on these ligations on a solid phase. Kent and co-workers published for the first time the solid phase chemical ligation (SPCL), for the total synthesis of unprotected peptides.[38] The synthesis onto the solid support allows to decrease the number of intermediate purifications.

4.1. Principle of the solid phase chemical ligation

In the solid supported chemical ligation (SPCL), the first fragment is anchored onto a water-compatible insoluble support and the ligations of the successive peptidyl fragments can occur in solid phase. At the end of the elongation, the peptide is cleaved and recovered in solution. The solid phase approach has some remarkable advantages against the solution approach such as the removal of the excess of reagents by a simple wash step, the minimization of the aggregation and precipitation. As in solution, the elongation of the peptide can be performed in the "N-to-C" or "C-to-N" directions (Figure 13).

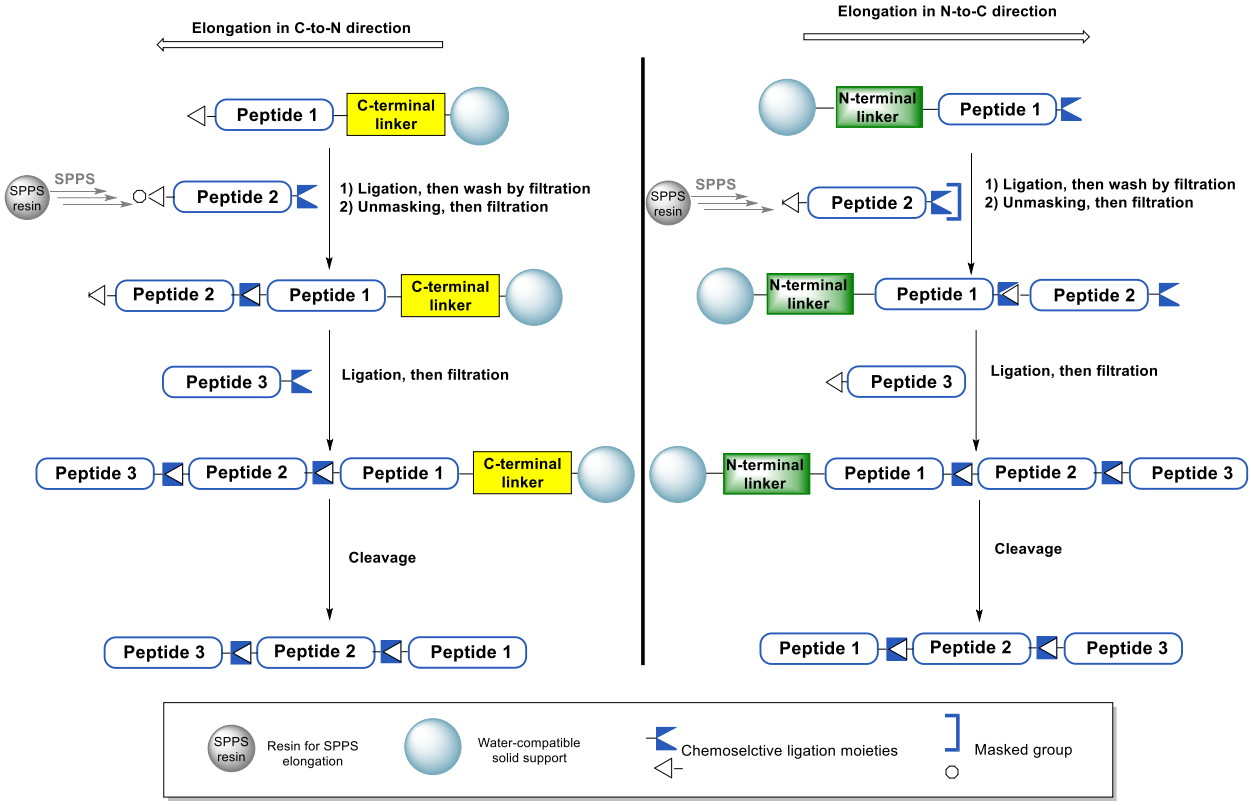


Figure 13. Principle of the solid phase chemical ligation in the C-to-N and N-to-C directions.

The first peptidyl fragment is anchored to the solid support *via* a C-terminal linker or an N-terminal linker, depending on the direction of the elongation, C-to-N or N-to-C direction, respectively. The choice of an adequate linker is required to assure the success of the synthesis.

4.2. C-terminal linkers for SPCL in the C-to-N direction

In general, the first unprotected peptide is introduced onto the solid support through an intermediate C-terminal linker. Then the multiple NCLs are carried out in the C-to-N

direction. In the same way as in solution, the N-terminal cysteine should be protected to avoid secondary reactions. [39] The excess of reactants is easily washed by simple filtration, whereas a chromatographic purification is required after the synthesis of each peptidyl fragments to remove the acetylated truncated peptides generated.

The pioneering work of Kent for the SPCL in the C-to-N direction, proposed the use of a base-labile ester linker compatible with Boc-SPPS chemistry.[38] That linker was synthesized onto the resin prior to the elongation of the first segment, starting from a commercially available Boc-Lys(Fmoc)-OH. That linker contained a C-terminal ketone to immobilize the first segment onto a functionalized aminooxy water-compatible solid support *via* oxime ligation (Figure 14). The Kent´s linker met some important criteria. It was stable under Boc-SPPS conditions, stable to successive NCLs, compatible with the Cys-deprotection steps and it could be removed by aqueous base treatment, to yield an acid C-terminal polypeptide.

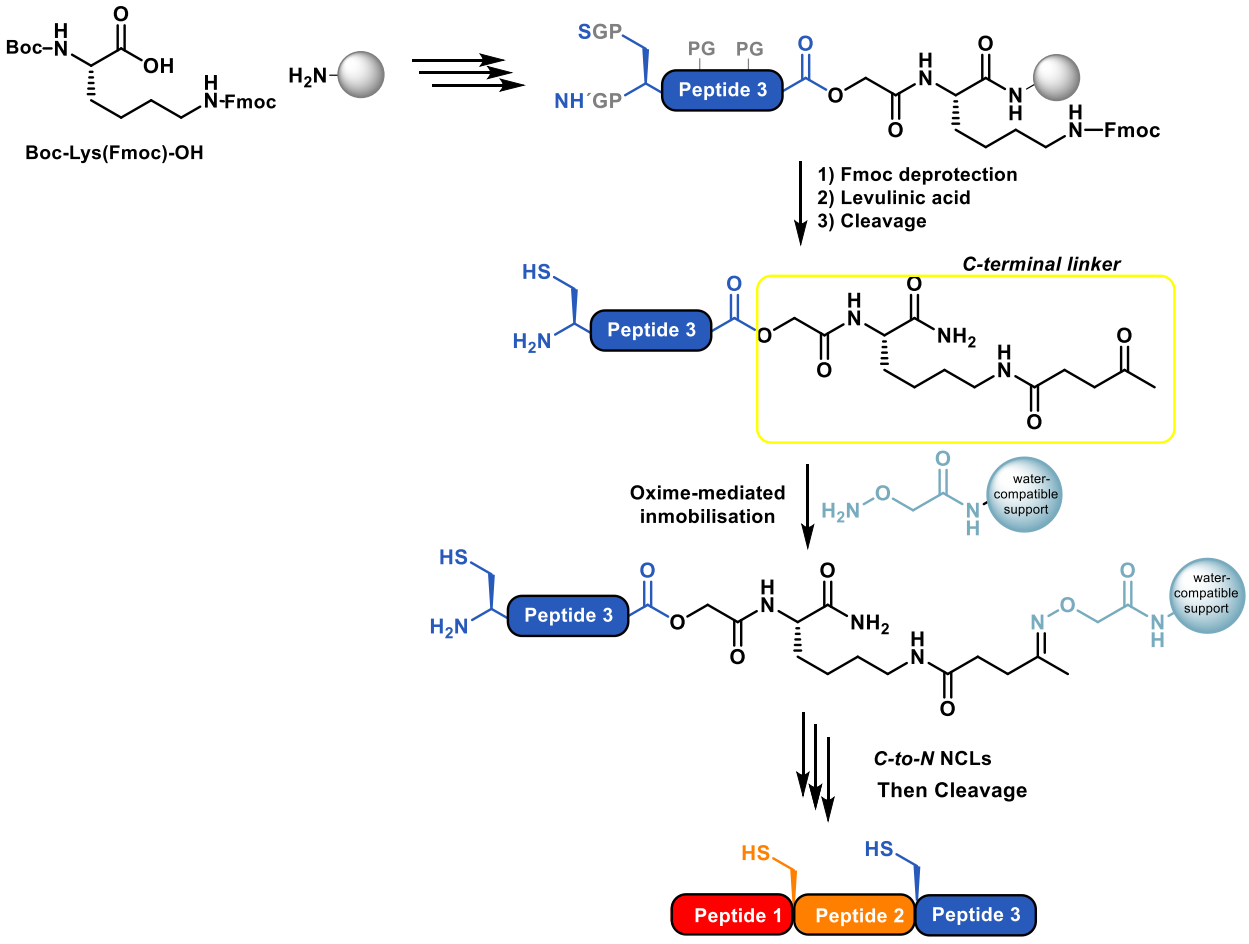


Figure 14. Synthesis of the C-terminal linker from Boc-Lys (Fmoc)-OH. Ketone residue allowed the oxime-mediated immobilization.[38]

Dawson [40] proposed the use of an alternative acid-labile C-terminal linker, the "safety catch acid-labile" (SCAL) linker. That one contained a thioester to allow the immobilization of the first fragment onto a cysteine-functionalized water compatible solid support. The release of the peptide, at the end, occurs under TFA treatment (Figure 15).

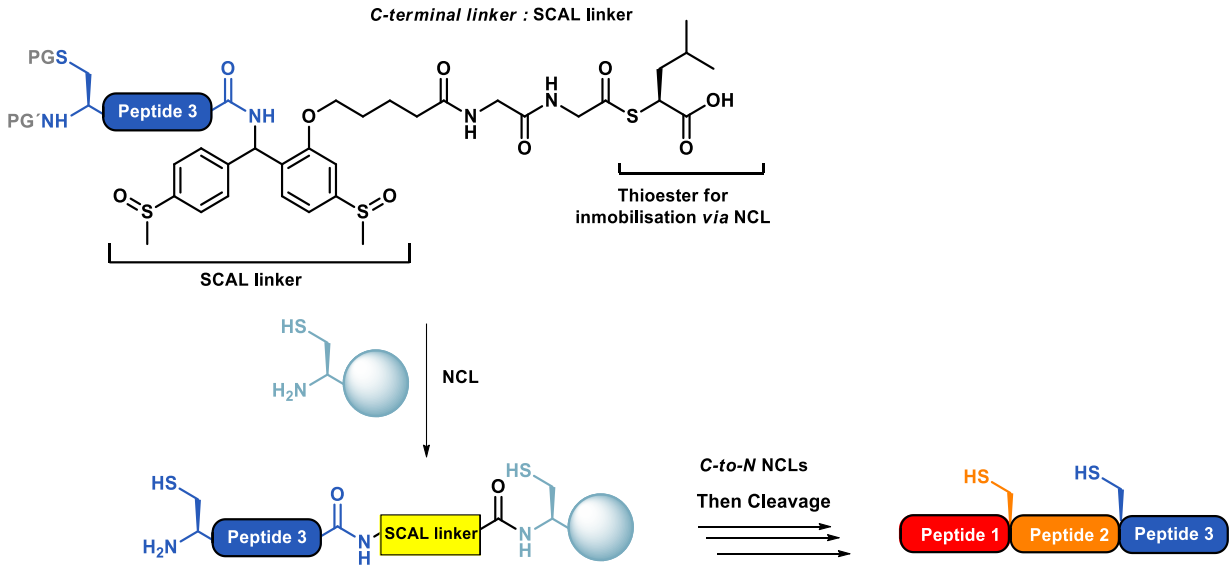


Figure 15. General scheme of Dawson´s strategy based on the use of a C-terminal SCAL linker. [40]

Alternately Cotton and Muir,[41] applied for the first time *the solid phase expressed protein ligation* (SPPL) to the synthesis of long semi-synthetic proteins. That approach offered the advantage of chemically modied recombinant proteins. Their C-terminal linker referred to as "PreSciss linker" was constituted of an nonapeptide (CLEVLFQGP) with an N-terminal Cys for proceeding with successive NCLs. The C-terminal extremity of the linker contained a biotin moiety to allow the binding with an avidin solid support. The liberation of the peptide in solution at the end proceeded through the protease recognition "PreScission protease" of the cleavable site of the linker (Figure 16).

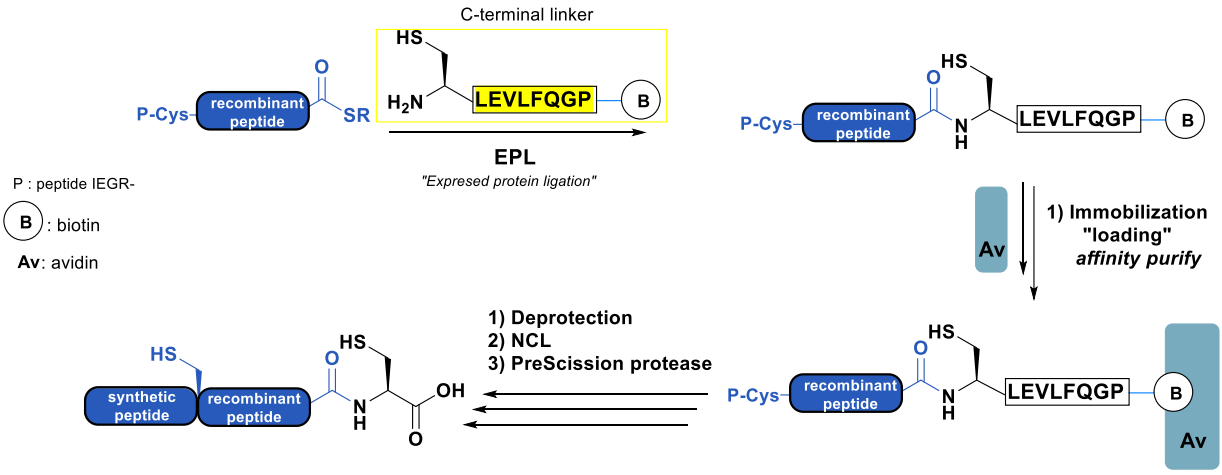


Figure 16. C-terminal linker used by Muir for SPPL.

The application of the SPCL without the first immobilization step has been also reported. For those strategies, the first peptidyl fragment was directly synthesized onto the solid support, which should be compatible at the same time with the SPPS and with the NCL conditions. There are only two examples in the literature. In both cases, the solid support was functionalized with two commercially available C-terminal linkers. In addition, the N-terminal cysteine residue was protected *via* a 1,3-thiazolidine-4-carboxylic acid (Thz) which was converted in the free-cysteine by methoxylamine treatment. (Figure 17).[25] The first report was published by Kent, who made use of a *para*-hydroxybenzyl alcohol linker which was directly anchored onto a SPOCC resin.[42] He reported the SPCL and supported-folding of a 26 amino acids trypsin inhibitor peptide. Later, Brik´s team made use of a acid labile linker, the methoxycarbonylamino]-2,4-dimethoxybenzyl}-phenoxyacetic acid linker (Rink-amide linker) for the total synthesis of a 124 mer protein onto a PEGA resin.[43]

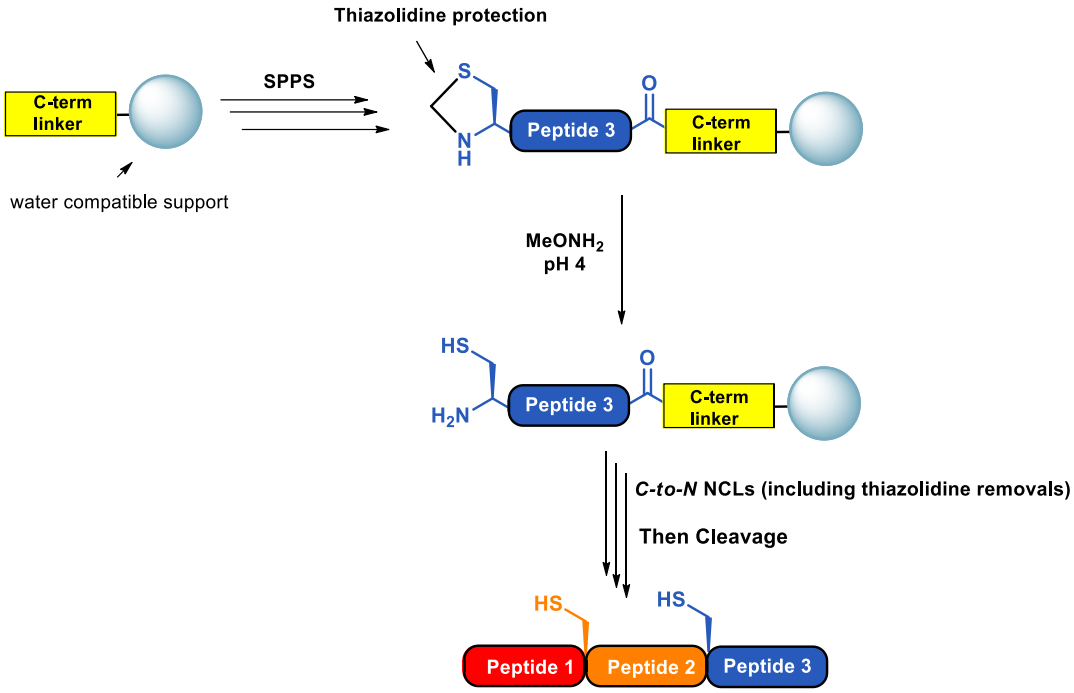


Figure 17. SPCL and thiazoline protection and deprotection strategy used firstly by Kent[34] and then by Brik[36].

More recently, Ottensen´s team[44] published a dual-linker strategy for the synthesis of histone proteins by a hybrid solid-solution phase ligation (Figure 18). Their dual linker consisted in a 4-hydroxymethylbenzoic acid (HMBA) linker attached through a ligation handle to a Rink linker. For one hand, HMBA[45] functioned as a C-terminal cleavable linker, and on the other, Rink-amide served as an analytical tool for quantification of the yield after each ligation cycle.

First, the peptide fragment 4 containing the dual linker is deprotected from the Thz protected Cys under methoxylamine treatment to generate the free Cys which can react by NCL with the next peptide fragments. The peptide is liberated in solution under TFA cleavage of the rink linker. A last NCL is carried out in solution. Base-labile 4-hydroxymethylbenzoic acid (HMBA) which is stable to NCL conditions is then removed in solution and results in a native carboxyl terminus peptide which is needed for the correct folding of the protein in the nucleosome. Finally desulfurization of cysteine in alanine allows yielding the native protein.

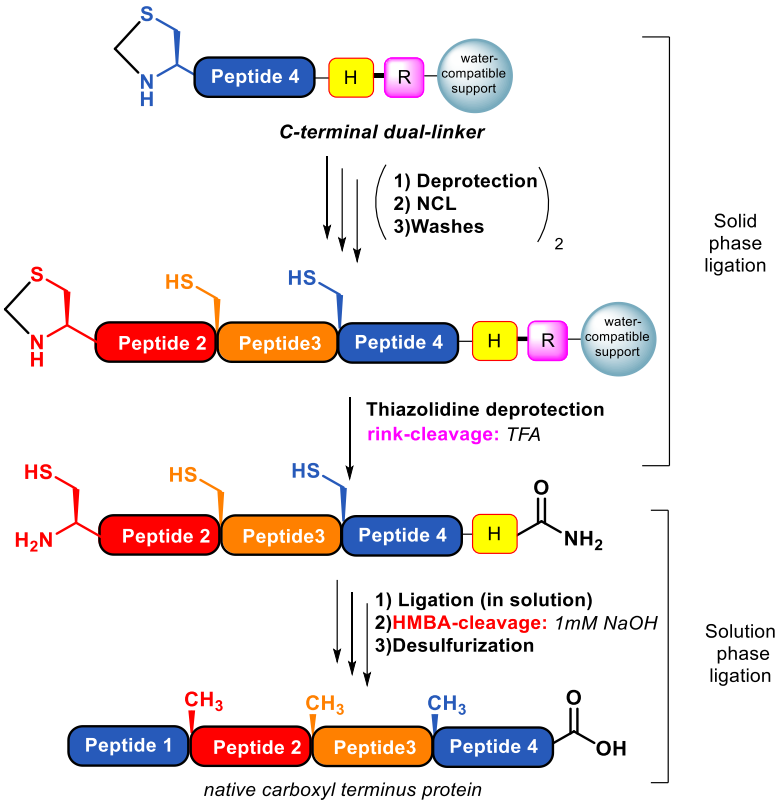


Figure 18. Hybrid solid-solution phase ligation.[44]

The number of publications describing the solid phase chemical ligation into the C-to-N directions *via* the use of C-terminal linkers is low. This is probably due to the difficulty of finding an adequate linker that is at the same time compatible with the SPPS and with the NCL conditions. One of the main advantages of these strategies is the reduction of the intermediate purification steps and the removal of the excess of reagents by simple washings. However, in any case it has been possible to suppress the compulsory purifications after the SPPS of each fragment. For most complex and long peptides, the need of doing multiple SPPS and consequently chromatographic purifications can compromise the final yield.

In order to overcome all the limitations in the synthesis of peptides, a very promising method for the synthesis and purification of peptides was developed. It was referred to as "*catch*-and-*release*" purification. It needs several requirements; the most important is the nature of the N-terminal linker. This method has been expanded to the synthesis of long peptides *via* solid phase chemical ligation in the N-to-C direction. In the next part of this chapter, the existing N-terminal linkers for the *catch*-and-*release* purification of peptides

have been described in detail, as well as applications to further solid phase chemical ligations in the N-to-C-direction.

5. 33References Chapter 1, part 1

(1) Zhao, Q.; Xu, W.; Xing, L.; Lin, Z. Recombinant Production of Medium- to Large-Sized Peptides in Escherichia Coli Using a Cleavable Self-Aggregating Tag. *Microb. Cell Fact.* **2016**, *15* (1), 136.

(2) Hwang, P. M.; Pan, J. S.; Sykes, B. D. Targeted Expression, Purification, and Cleavage of Fusion Proteins from Inclusion Bodies in Escherichia Coli. *FEBS Lett.* **2014**, *588* (2), 247.

(3) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149.

(4) Carpino, L. A.; Han, G. Y. 9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group. *J. Am. Chem. Soc.* **1970**, *92* (19), 5748.

(5) Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F. Handles for Fmoc Solid-Phase Synthesis of Protected Peptides. *ACS Comb. Sci.* **2013**, *15* (5), 217.

(6) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino Acid-Protecting Groups. *Chem. Rev.* **2009**, *109* (6), 2455.

(7) El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem Rev* **2011**, *111* (11), 6557.

(8) Made, V.; Els-Heindl, S.; Beck-Sickinger, A. G. Automated Solid-Phase Peptide Synthesis to Obtain Therapeutic Peptides. *Beilstein J .Org. Chem.* **2014**, *10*, 1197.

(9) Verzele, D.; Madder, A. Patchwork Protein Chemistry: A Practitioner's Treatise on the Advances in Synthetic Peptide Stitchery. *Chem. Bio. Chem.* **2013**, *14* (9), 1032.

(10) Schnölzer, M.; Kent, S. B. Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease. *Science* **1992**, *256* (5054), 221.

(11) Vazquez, O.; Seitz, O. Templated Native Chemical Ligation: Peptide Chemistry beyond Protein Synthesis. *J .Pept. Sci.* **2014**.

(12) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. Über Peptidsynthesen. 8. Mitteilung Bildung von S-Haltigen Peptiden Durch Intramolekulare Wanderung von Aminoacylresten. *Justus Liebigs Ann. Chem.* **1953**, *583* (1), 129.

(13) Liu, F.; Mayer, J. P. Protein Chemical Synthesis in Drug Discovery. *Top. Curr. Chem.* **2015**.

(14) Kemp, D. S.; Galakatos, N. G. Peptide Synthesis by Prior Thiol Capture. 1. A Convenient Synthesis of 4-Hydroxy-6-Mercaptodibenzofuran and Novel Solid-Phase Synthesis of Peptide-Derived 4-(Acyloxy)-6-Mercaptodibenzofurans. *J. Org. Chem.* **1986**, *51* (10), 1821.

(15) Kemp, D. S.; Galakatos, N. G.; Bowen, B.; Tan, K. Peptide Synthesis by Prior Thiol Capture. 2. Design of Templates for Intramolecular O,N-Acyl Transfer. 4,6-Disubstituted Dibenzofurans as Optimal Spacing Elements. *J. Org. Chem.* **1986**, *51* (10), 1829.

(16) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of Proteins by Native Chemical Ligation. *Science.* **1994**, *266*, 776.

(17) Malins, L. R.; Payne, R. J. Recent Extensions to Native Chemical Ligation for the Chemical Synthesis of Peptides and Proteins. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70.

(18) Bode, J. W.; Fox, R. M.; Baucom, K. D. Chemoselective Amide Ligations by Decarboxylative Condensations ofN-Alkylhydroxylamines and α-Ketoacids. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (8), 1248.

(19) Saxon, E.; Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science.* **2000**, *87*, 5460.

(20) Wan, Q.; Danishefsky, S. J. Free-Radical-Based, Specific Desulfurization of Cysteine: A Powerful Advance in the Synthesis of Polypeptides and Glycopolypeptides. *Angew. Chem. Int. Ed. Engl.* **2007**, *46* (48), 9248.

(21) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed Protein Ligation: A General Method for Protein Engineering. *Proc. Natl. Acad. Sci. USA* **1998**, *95* (12), 6705.

(22) Tesser, G. I.; Balvert-Geers, I. C. The Methylsulfonylethyloxycarbonyl Group, a New and Versatile Amino Protective Function. *Int. J. Pept. Protein Res.* **1975**, *7* (4), 295.

(23) Muir, T. W.; Dawson, P. E.; Kent, S. B. Protein Synthesis by Chemical Ligation of Unprotected Peptides in Aqueous Solution. *Methods Enzymol.* **1997**, *289*, 266.

(24) Ueda, S.; Fujita, M.; Tamamura, H.; Fujii, N.; Otaka, A. Photolabile Protection for One-Pot Sequential Native Chemical Ligation. *Chem. Bio. Chem.* **2005**, *6* (11), 1983.

(25) Bang, D.; Kent, S. B. H. A One-Pot Total Synthesis of Crambin. *Angew. Chem. Int. Ed.* **2004**, *43* (19), 2534.

(26) Bang, D.; Chopra, N.; Stephen B. H. Kent. Total Chemical Synthesis of Crambin. *J. Am. Chem. Soc.* **2004**, *126* (5), 1377.

(27) Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L. Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angew. Chem. Int. Ed.* **2011**, *50* (33), 7645.

(28) Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *N* → *S* Acyl-Transfer-Mediated Synthesis of Peptide Thioesters Using Anilide Derivatives. *Org. Lett.* **2009**, *11* (4), 823.

(29) Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. N-Sulfanylethylanilide Peptide as a Crypto-Thioester Peptide. *Chem. Bio. Chem.* **2011**, *12* (12), 1840.

(30) Sato, K.; Shigenaga, A.; Kitakaze, K.; Sakamoto, K.; Tsuji, D.; Itoh, K.; Otaka, A. Chemical Synthesis of Biologically Active Monoglycosylated GM2-Activator Protein Analogue Using N -Sulfanylethylanilide Peptide. *Angew. Chem. Int. Ed.* **2013**, *52* (30), 7855.

(31) Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-Sulfanylethyl)amino

Native Peptide Ligation. *Org. Lett.* **2010**, *12* (22), 5238.

(32) 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.

(33) Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O. A One-Pot Three-Segment Ligation Strategy for Protein Chemical Synthesis. *Angew. Chem. Int. Ed.* **2012**, *51* (1), 209.

(34) Yang, R.; Hou, W.; Zhang, X.; Liu, C.-F. N-to-C Sequential Ligation Using Peptidyl N,N -Bis(2-Mercaptoethyl)amide Building Blocks. *Org. Lett.* **2012**, *14* (1), 374.

(35) Bang, D.; Pentelute, B. L.; Kent, S. B. H. Kinetically Controlled Ligation for the Convergent Chemical Synthesis of Proteins. *Angew. Chem. Int. Ed.* **2006**, *45* (24), 3985.

(36) Weinstock, M. T.; Jacobsen, M. T.; Kay, M. S. Synthesis and Folding of a Mirror-Image Enzyme Reveals Ambidextrous Chaperone Activity. *Proc. Natl. Acad. Sci. USA.* **2014**, *111* (32), 11679.

(37) 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 Discov.* **2017**, *3*, 17008.

(38) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121* (38), 8720.

(39) Raibaut, L.; El Mahdi, O.; Melnyk, O. Solid Phase Protein Chemical Synthesis. *Top. Curr. Chem.* **2015**.

(40) Brik, A.; Keinan, E.; Dawson, P. E. Protein Synthesis by Solid-Phase Chemical Ligation Using a Safety Catch Linker. *J. Org. Chem.* **2000**, *65* (12), 3829.

(41) Cotton, G. J.; Muir, T. W. Generation of a Dual-Labeled Fluorescence Biosensor for Crk-II Phosphorylation Using Solid-Phase Expressed Protein Ligation. *Chem. Biol.* **2000**, *7* (4), 253.

(42) Johnson, E. C. B.; Kent, S. B. H. Towards the Total Chemical Synthesis of Integral Membrane Proteins: A General Method for the Synthesis of Hydrophobic Peptide-Athioester Building Blocks. *Tetrahedron Lett.* **2007**, *48* (10), 1795.

(43) Jbara, M.; Seenaiah, M.; Brik, A. Solid Phase Chemical Ligation Employing a Rink Amide Linker for the Synthesis of Histone H2B Protein. *Chem. Commun.* **2014**, *50* (83), 12534.

(44) Yu, R. R.; Mahto, S. K.; Justus, K.; Alexander, M. M.; Howard, C. J.; Ottesen, J. J.; Hayashi-Takanaka, Y.; Oda, T.; Sato, M.; Park, S.-Y.; et al. Hybrid Phase Ligation for Efficient Synthesis of Histone Proteins. *Org. Biomol. Chem.* **2016**, *14* (9), 2603.

(45) Harris, P. W. R.; Brimble, M. A. Toward the Total Chemical Synthesis of the Cancer Protein NY-ESO-1. *Biopolymers* **2010**, *94* (4), 542.

Part 2. N-terminal linkers as traceless chemical tags for the *catch*-and-*release* purification of synthetic peptides and self-purifying solid-phase chemical ligations.

Note that this part of the manuscript is planned to be submitted as a review article, and has consequently been written independently from the previous and following sections of the manuscript.

1. Introduction: the *catch*-and-*release* peptide purification concept

From Fisher's synthesis of glycylglycine[1] to the few recent achievements concerning the syntheses of proteins containing more than 300 amino acid (aa) residues,[2–7] the biology-driven constant need in peptides and proteins have inspired many conceptual breakthroughs in organic chemistry, such as the development of coupling reagents, orthogonal protective groups, solid phase synthesis or chemoselective ligations reactions. Strikingly, these synthetic methodologies could not have emerged without the parallel continuous improvement of adapted analysis and purification technologies, many of which having been initially developed for application to bio-sourced peptides and proteins.

In this context, Cuatrecasas *et al.* first exploited in 1968 non-covalent enzyme-substrate interactions for the purification of natural proteins through selective immobilization of an enzyme onto a solid support grafted with a substrate-derived competitive inhibitor, followed by elution using either an excess of a soluble ligand, or salt or pH conditions disfavoring protein-ligand interactions.[8] This so-called affinity chromatography method represented a major advance in comparison with the traditional precipitation techniques used for the purification of proteins and greatly accelerated the identification of new proteins and the understanding of important biological processes.[9] In a related approach, yet not relying on interaction with a specific substrate, Jellum reported in 1963 the purification of cysteine-containing proteins through specific thiol-metal interactions with an organomercurial-functionalized solid support.[10]

Inspired by these two techniques, Merrifield introduced in 1976[11] the *catch*-and-*release* concept for the purification of synthetic peptides obtained through his newly-developed stepwise solid phase peptide synthesis (SPPS) technology.[11,12] The nowadays widely used SPPS is based on the stepwise elongation of the peptide chain through cycles consisting on the coupling of *N*α-carbamate-protected amino acids, followed by carbamate deprotection. Incomplete couplings are, by far, the most common source of by-products, leading to the formation of "deleted peptides" (e.g. sequences lacking one or several amino acids) which are difficult to separate from the target peptide due to similar physicochemical properties. For this reason, a "capping" step can be added after each coupling, in order to stop the elongation of uncoupled peptides. This is generally achieved by acetylation using a large excess of acetic anhydride, but other capping reagents have also been used.[13] Capping results in the formation of "truncated peptides" which are easier to separate from the target peptide, the most common purification method being reverse phase HPLC. However, the purification can sometimes be very complicated, and co-elution of truncated peptides can impair the homogeneity of the purified target peptide, this problem being prominent for medium-sized and long peptides.

The *catch*-and-*release* purification strategy described by Merrifield[11] is based on the introduction of two additional N-terminal amino acids after the SPPS elongation of the target sequence: methionine then cysteine. After cleavage from the SPPS resin, the crude peptide mixture (truncated acetylated byproducts and Cys-Met-tagged target peptide) is incubated with an organomercurial-agarose support, resulting in the selective immobilization of the tagged peptides through thiol-metal interactions. The solid support is then washed to remove the non-tagged impurities. Finally, the tagged peptide can be eluted from the support using an excess of cysteine, and the Met-peptide bond is selectively cleaved through treatment with cyanogen bromide, to give the purified untagged target peptide (Figure 1).[11]

Note that this tag-mediated purification concept has later been extended to the purification of recombinant proteins. A large collection of fusion tags is now available,[14–16] and the tag-based strategy has become the most widely used method for protein isolation.

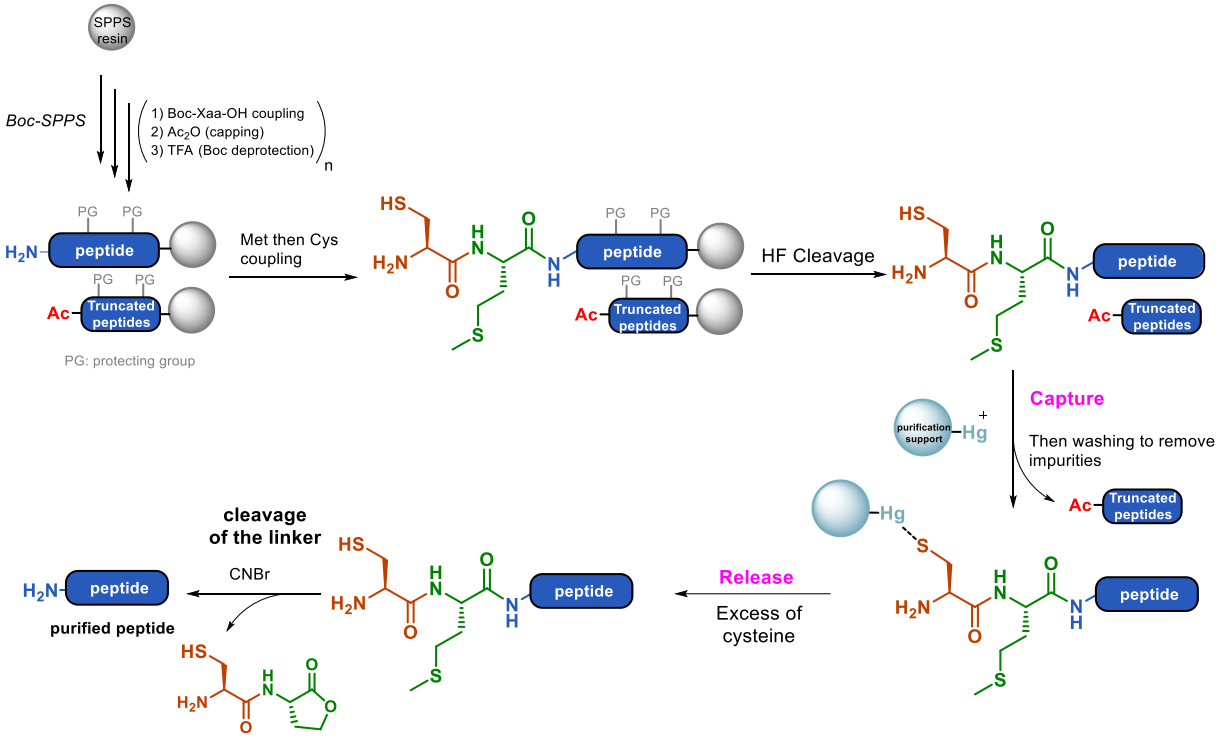


Figure 1. Merrifield´s proof of concept of *catch*-and-*release* purification through the use of an N-terminal cysteinyl-methionine linker.[11]

Merrifield's approach suffers from obvious limitations, as peptides containing internal cysteines or methionines cannot be purified using this strategy. During the last 40 years, many efforts have been devoted to expand the methodology to the purification of more complex peptides.

The key aspect for these strategies is the design of N-terminal cleavable heterobifunctional linkers, containing both a "binding moiety" aimed at immobilizing the target tagged peptide on a solid support, and a "cleavable moiety" for the subsequent traceless liberation of the purified peptide into solution (Figure 2).

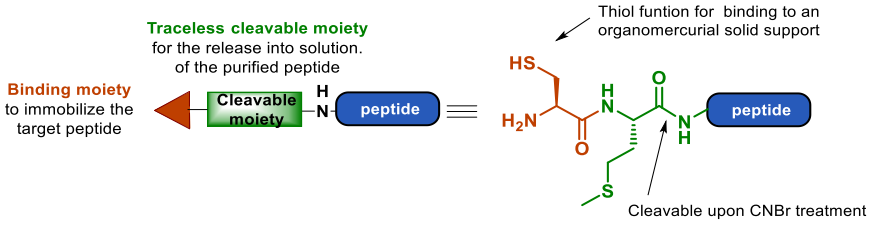


Figure 2. General structure of N-terminal linkers developed for *catch*-and-*release* purification.

A large collection of N-terminal linkers is now available, and the *catch*-and-*release* purification strategy has been used to purify a wide range of synthetic peptides (Figure 3).

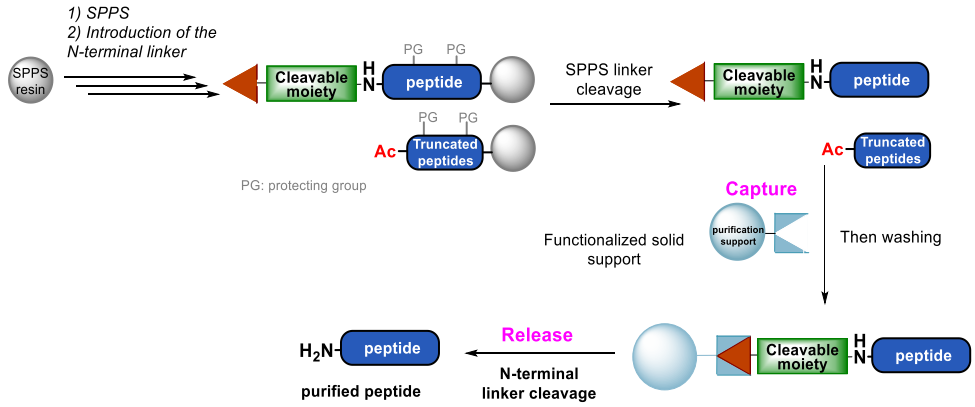


Figure 3. General principle of the *catch*-and-*release* purification approach.

The aim of this review is to give an exhaustive overview of all the linkers having been developed to date for this purpose. Beyond a simple linker catalog, we chose to present separately the different "catch" and "release" strategies, given that most linkers were designed in a modular fashion, thus incorporating two independent units, excluding only a few exceptions.

The vast majority of both immobilization and cleavage strategies have in common the highly chemoselective nature of both bond formation (including selective non-covalent interactions) and cleavage, which is required for the handling of unprotected peptides considering the potential cross-reactivity of many amino acid side chains.

Two different classes of methods used for the chemoselective binding to the purification solid-supports can be differentiated, depending on the use of either non-covalent or covalent captures. This boundary is however somewhat artificial, given the strength of some of the non-covalent interactions and the intrinsic reversible nature of some covalent bonds used for capture step. We chose to classify the metal-ligand interactions in the "non-covalent" interactions subchapter. Note that we also included a few closely related linkers aimed at facilitating chromatographic purification by either reverse-phase, fluorous phase or ion-exchange chromatography, by increasing the retention of the tagged peptides rather

than a true catch/release mechanism.

Apart from a very few exceptions, the chemoselective bond cleavage methods used for the release of the immobilized target can be classified in a more clear-cut manner, depending on the chemical nature of the reagents used: mainly electrophiles, oxidants, acids, bases, nucleophiles, or photons.

Beyond *catch*-and-*release* purification, conceptually-related methodologies are further described, such as the internal resin capture and the "cap-tag" concepts, as well as the promising extension of the *catch*-and-*release* concept to the assembly of very long peptides *via* successive "self-purifying" solid phase chemical ligations of multiple peptide segments.

2. Capture strategies

2.1. Metal-ligand and noncovalent interactions

2.1.1. Metal-ligand

In addition to the seminal proof-of-concept of Merrifield (Figure 1), thiol-organomercurial interactions have been explored for the immobilization of protected peptides by Lansbury.[17] However, the acidic conditions used for the cleavage of the linker causes a protodemercuration of polystyrene-bound mercury, resulting in the leaching of mercury salts which could not be separated from the released peptides.

N-terminal linkers containing the so-called "His_6-tag" were also developed (Figure 4).[18–20] This tag consisting in six consecutive histidine residues is one the most extensively used for the purification of recombinant proteins: this peptide sequence is able to strongly bind solid support functionalized with a nitrilotriacetic acid moiety that chelates nickel (II) cations (Ni-NTA), through cooperative multiple ligand-metal interactions.

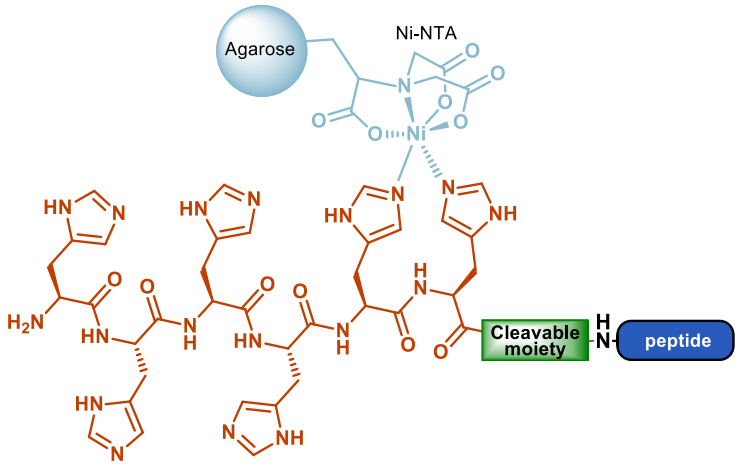


Figure 4. His_6-funtionalized linker binding to Ni-NTA agarose.

2.1.2. Ionic interactions

Merrifield developed in 1978 a derivative of the base-labile 9-fluorenylmethoxycarbonyl[21] (Fmoc) protecting group functionalized with a sulfonate moiety (Sulfmoc),[22] for application to the *catch*-and-*release* purification of peptides synthesized through Boc-SPPS.[22] At low pH, sulfonate-containing peptides can be selectively captured onto a cationic polymer: given the strong acidity of the sulfonic acid group, it remains in the sulfonate form at pH < 2 while the side chains of aspartates and glutamates and the C-terminal carboxylic acids are protonated. To our knowledge, this promising method was not further exploited.

In a related example, Mascagni reported two Fmoc group analogues derived with either lysines or glutamates, whose side chains were protected with HF-labile 2-chlorobenzyloxy and benzyl protecting groups, respectively.[23] The introduction of these linkers on target peptides synthesized by Boc-SPPS resulted in the modification of their net charge and facilitated their purification *via* ion-exchange chromatography (Figure 5). Note that in this case, the linker is cleaved in solution after elution.

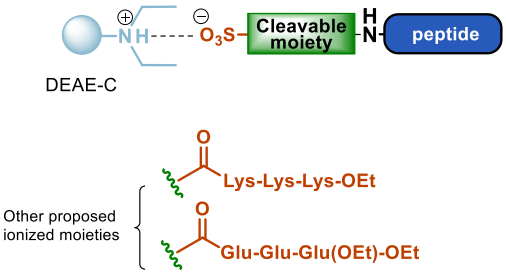


Figure 5. Ionic interactions-based linkers. DEAE-C: Diethylaminoethyl cellulose.

2.1.3. Avidin-biotin interaction

The interaction between the egg-white glycoprotein avidin (and its bacterial analog streptavidin) with biotin (vitamin H) is one of the strongest known non-covalent bonds (dissociation constant ~ 10^{-15} M). It has been extensively exploited by biochemists and molecular biologists as a Swiss army knife-like tool for labeling, imaging or immobilization purposes. In 1987, Lobl and co-workers[24] reported the impressive purification of a synthetic 153 residues N-terminally biotinylated interleukin, but without subsequent cleavage of the biotin from the target peptide. A large range of biotinylated N-terminal linkers cleavable under basic[25–29] or nucleophilic[30] conditions, or through autoproteolysis[31] or photocleavage[32] have been later proposed for the *catch*-and-*release* purification of synthetic peptides (Figure 6).

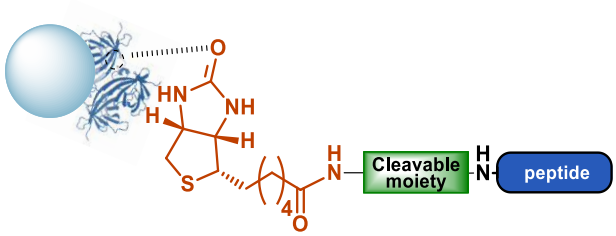


Figure 6. Biotinylated linkers bound to immobilized streptavidin.

2.1.4. Π-stacking and hydrophobic interactions

Ramage introduced in 1992 a derivative of the Fmoc group incorporating four extra benzene rings fused to the fluorene core: tetrabenzo [*a*,*c*,*g*,*i*]fluorenyl-17-methoxycarbonyl (Tbfmoc).[33–36] Given its large, flat and π-electron-rich nature, the Tbfmoc group strongly binds to graphite surfaces through non-covalent π-π interactions. Tagged peptides can thus be selectively adsorbed onto a porous graphitized carbon (PGC) matrix (Figure 7), while truncated peptides remain in solution.

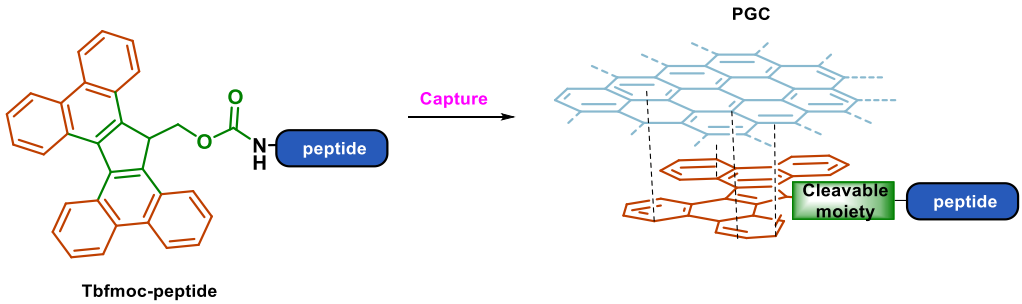


Figure 7. Tbfmoc interaction with a porous graphitized carbon.

The highly lipophilic nature of Tbfmoc has also been exploited to facilitate the purification of tagged peptides through standard reverse phase HPLC, the linker is cleaved in solution after elution. A number of related base-labile lipophilic linkers have been proposed for the same purpose (Figure 8).[13,23,25,37]

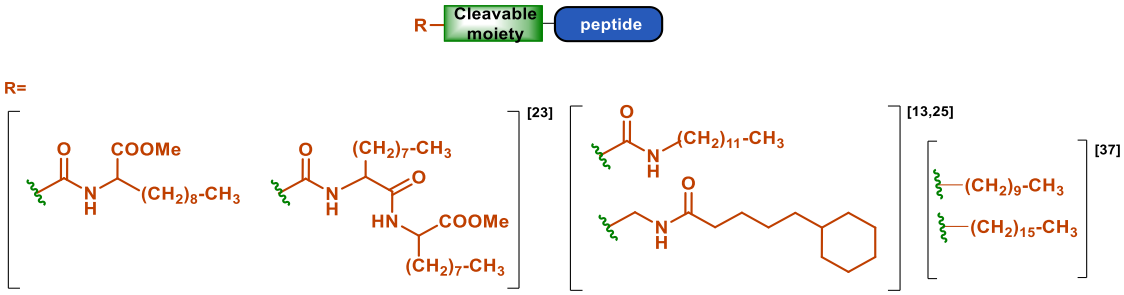


Figure 8. Lipophilic linkers to facilitate RP-HPLC purification.

2.1.5. Fluorous interactions

The unique property of perfluoroalkyl groups to bind to each-others through so-called fluorous interactions has been largely exploited by synthetic organic chemists to facilitate the separation of complex reaction mixtures.[38,39] Van Boom introduced in 2002 a perfluorooctylated acid-labile N-terminal linker for the *catch*-and-*release* purification of synthetic peptides, based on its selective adsorption onto fluorous silica gel (Figure 9).[40] However, this first generation linker required the use of protected peptides for the immobilization step and a more generally usable base-labile version has been further developed.[41] Note that fluorous-HPLC (FHPLC) can also be used to discriminate the tagged peptides from truncated impurities. In the latter case, the linker is cleaved in solution after elution.

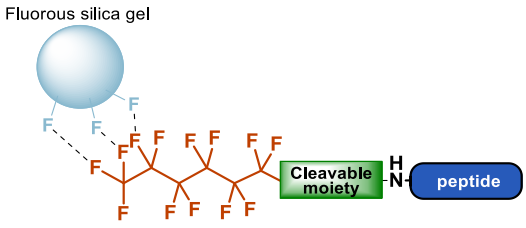


Figure 9. Perfluoroalkyl-functionalized linkers.

2.2. Covalent bond formation

2.2.1. Thiol-disulfide interchange and thiol alkylation

In 1978 Ragnarsson described an alternative to the seminal organomercurial/cysteine approach of Merrifield, based on disulfide interchange between a Cys-tagged peptide and a disulfide-functionalized solid support (Figure 10, strategy 1).[42] A similar strategy was later applied by Lansbury to the immobilization of protected peptides. The later also explored a reverse version making use of a thiolated-solid support and a disulfide-functionalized N-terminal linker (Figure 10, strategy 2).[17]

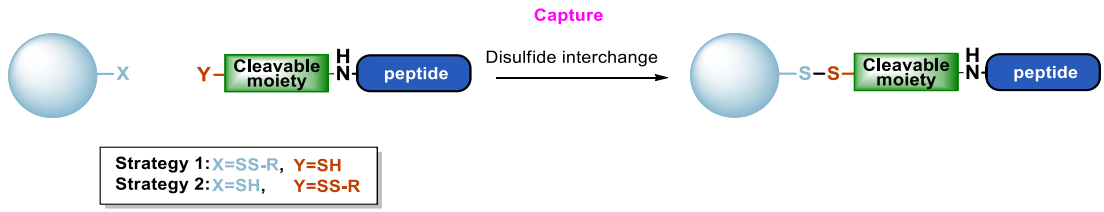


Figure 10. *Catch*-and-*release* purification *via* disulfide bond formation.

Alternatively, thiol-tagged peptides can be immobilized onto an iodoacetamide-functionalized solid support through the formation of an irreversible thioether bond, as pioneered in 1991 by Fujii employing a base-cleavable linker[43] (Figure 11) and later exploited by Lansbury with his before mentioned acid-labile linker.[17]

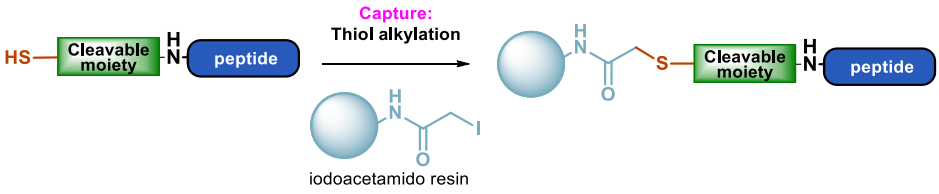


Figure 11. Thiol alkylation mediated capture.

2.2.2. Carbonyl-based ligations: oxime or thiazolidine formation

Besides thiol-based chemistry which is obviously limited to peptides devoid of any cysteine residue, ketone/aldehyde-based chemoselective ligations have largely been employed for the capture step.

In his 1997 pioneering work, Kent introduced a ketone-functionalized base-labile linker used for the immobilization of the tagged peptides onto an aminooxy-functionalized solid support.[44] Aimoto[45] then Melnyk[46] later developed alternative linkers cleavable under nucleophilic conditions (Figure 12). The reverse strategy can also be used: two different aminooxy-containing N-terminal linkers cleavable under either basic or reductive conditions, respectively, has been very recently described.[47]

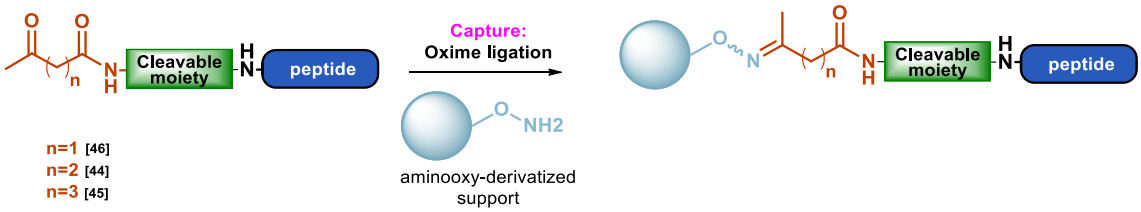


Figure 12. Oxime bond formation-mediated capture.

In 2001, Rose described an elegant linker-free approach for the *catch*-and-*release* purification of N-terminal cysteine-containing peptides from synthetic or recombinant origin, through selective reaction with an aldehyde-functionalized solid support to generate a thiazolidine.[48] The peptide can be released by treatment with excess methoxyamine to reverse the equilibrium (Figure 13).

Note that the same strategy has been used for the selective enrichment of tryptic digestion mixtures for proteomics analysis purposes.[49] In addition to cysteines, peptides containing an N-terminal serine or threonine residue are also purified, through the reversible formation of an oxazolidine.

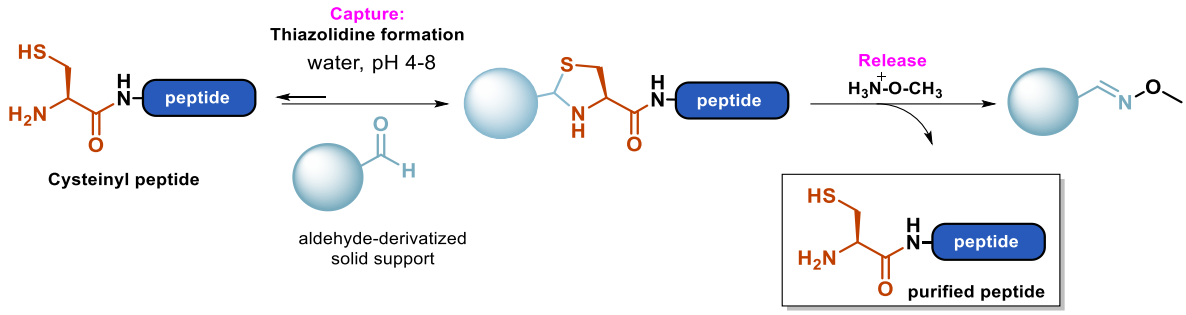


Figure 13. *Catch*-and-*release* purification of N-terminal Cys-containing peptides through thiazolidine formation.

To extend the methodology to peptides containing any N-terminal amino acid, a cleavable moiety was later incorporated between the cysteine residue and the target peptide (Figure 14).[50]

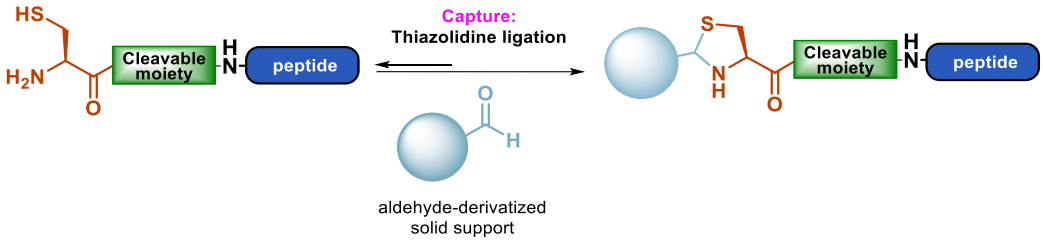


Figure 14. Capture of cysteine-functionalized liner *via* thiazolidine formation.

2.2.3. Cycloadditions

2.2.3.1. Cu^{I}-catalyzed azide/alkyne cycloaddition (CuAAC)

The highly chemoselective copper(I)-catalyzed azide/alkyne cycloaddition (CuAAC) variant of the Huisgen cycloaddition independently discovered by Meldal[51] and Sharpess[52] rapidly emerged as an efficient method for conjugation of biomolecules including peptides and proteins.[53] CuAAC ligation has been applied to the *catch*-and-*release* purification of synthetic peptides by Aucagne and Delmas who synthetized an azido-functionalized base-cleavable N-terminal linker (Figure 15).[54]

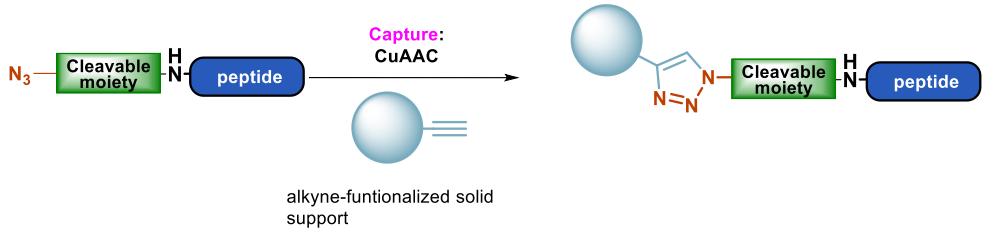


Figure 15. Azido-functionalized linker for CuAAC-mediated capture.

2.2.3.2. Inverse electron demand Diels-Alder (iEDDA)

Jacob *et al.*[55] reported in a patent an N-terminal cyclobutene-functionalized linker for the *catch*-and-*release* purification of peptide nucleic acids (PNA). A tetrazine-derivatized solid support allowed for the covalent capture of the tagged PNA through an inverse demand Diels-Alder cycloaddition (iEDDA) (Figure 16).

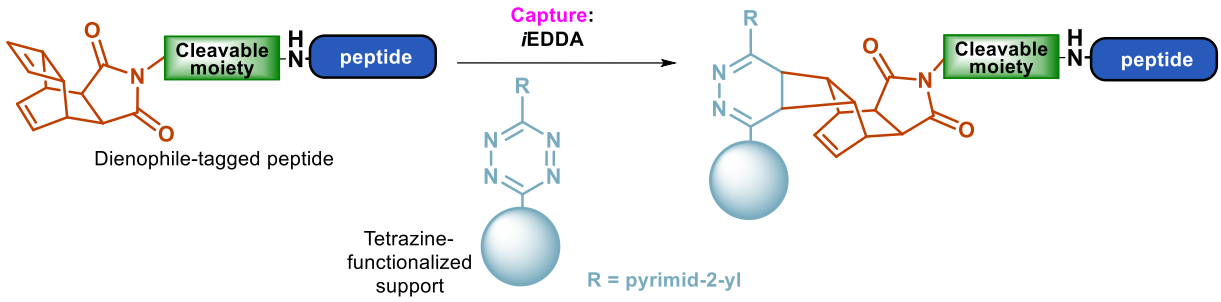


Figure 16. Covalent capture *via* inverse electron demand Diels-Alder cycloaddition.

2.2.4. Methoxyamine-isothiocyanate reaction

Aimoto proposed a *catch*-and-*release* purification method relying on the selective immobilization of N-terminal *N*-methoxyglycine-containing peptides into an isothiocyanate-functionalized resin (Figure 17).[56] The low pKa of the oxyamine (pKa 4 to 5) makes it nucleophilic at low pH where amines (Lys) or imidazoles (His) are protonated, thus allowing the chemoselective capture of the methoxyamino-peptide.[57,58] However, cysteines thiols must be protected. Aimoto[56] used oxidation into disulfides, the immobilized peptide being then reduced through treatment with TCEP to regenerate free cysteines.

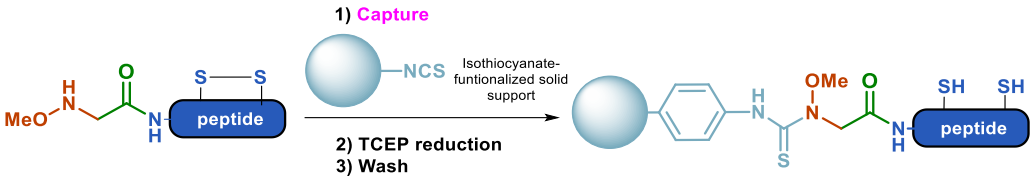


Figure 17. Methoxyamine-isothiocyanates mediated capture

2.2.5. Polymerization

Fang has exploited polyacrylamide gel formation to develop an alternative *catch*-and-*release* purification method.[59] He developed an acid-labile N-terminal linker functionalized with an acrylamide moiety which can polymerize in the presence of *bis*-acrylamide to form a polyacrylamide solid support *in situ* (Figure 18). Protection of the peptide side chains is required to avoid interference with the polymerization process.

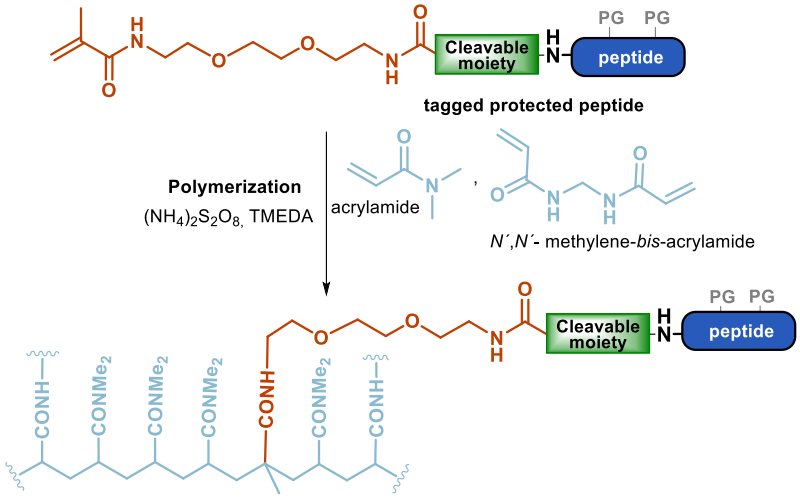


Figure 18. Capture by polymerization. TMEDA: tetramethylethylenediamine.

3. Release strategies

3.1. Electrophilic cleavage

In Merrifield's seminal example of *catch*-and-*release* peptide purification (N-terminal Cys-Met sequence)[11] the liberation of the purified peptide occurred through selective cleavage of the Met-Xaa peptide bond (Xaa = any amino acid) by reaction with cyanogen bromide (Figure 19),[60] a classical protein biochemistry reaction. Later, Corradin[18] used the same cleavage strategy for the purification of a 69 aa peptide corresponding to the C-terminal

region of the circumsporozoite protein of *Plasmodium berghei* (PbCS 242-310) [61] using a N-terminal His_6-Met sequence. These cleavage conditions are obviously limited to peptides which do not contain any internal Met in their sequences.

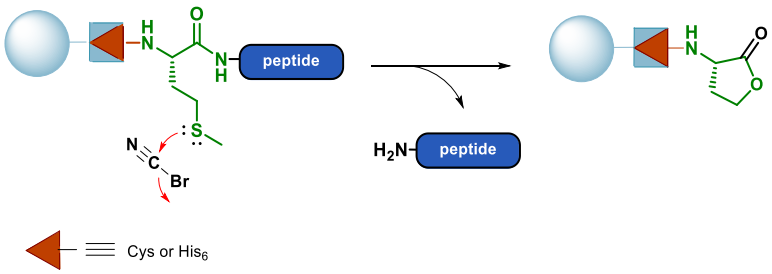


Figure 19. Cyanogen bromide-mediated electrophilic cleavable.

3.2. Oxidative cleavage.

β-amino-alcohols can be cleaved upon treatment with sodium periodate,[62] in a similar way as the Malaprade reaction of 1,2-diols. Exploiting this reactivity, Rose developed the 4-amino-2-hydroxy-butyryl (Ahb) linker, whose oxidation leads to an imine that spontaneously hydrolyzed into the amine (Figure 20).[50] The procedure is compatible with oxidation-sensitive tryptophanes and can be adapted to Met residues, but internal cysteines must be protected to avoid oxidation into a sulfonate derivative.

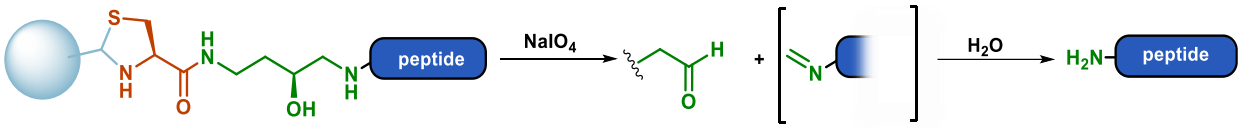


Figure 20. Release *via* periodate oxidation.

Building blocks were synthesized, incorporating the N-terminal amino acid and the Ahb moiety suitably protected for Boc-SPPS (Figure 21).Conveniently, any binding moiety could be thus potentially introduced through solid phase coupling. Rose coupled Boc-Cys(*p*MeBzl)-OH to generate a Cys-Ahb linker for the thiazolidine *catch*-and-*release* purification of a set of model polypeptides.

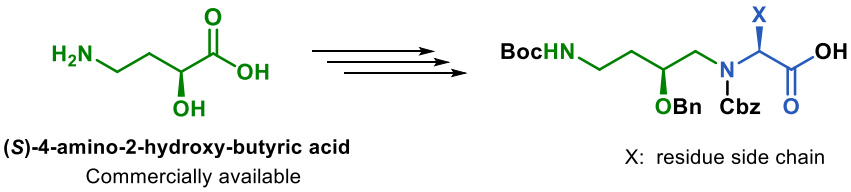


Figure 21. Ahb-derived amino acid building block for Boc-SPPS.

3.3. Reductive cleavage

A N-terminal linker cleavable under reductive conditions has been recently developed.[47] The TCEP-mediated reduction of the aryl azide into an amine is followed by a spontaneous 1,6 elimination leading to the liberation of the peptide (Figure 22).

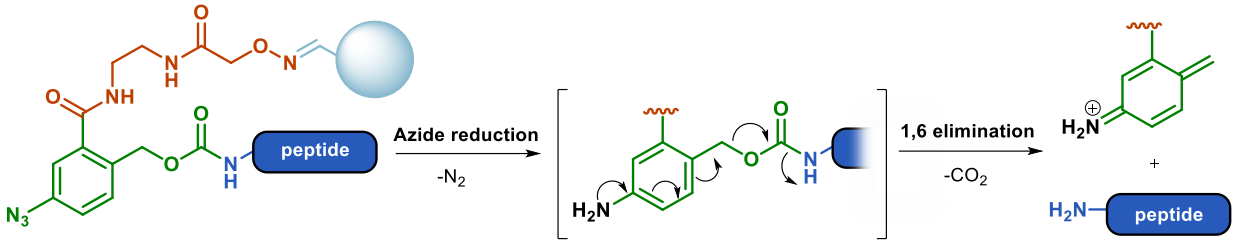


Figure 22. N-terminal linker cleavable under reductive conditions.

3.4. Acidolytic cleavage

To date, all the strategies based on an acidolytic cleavage of the linker can only be applied to the immobilization of protected peptides. Such peptides can be synthesized *via* Fmoc chemistry using a highly acid-labile 2-Cl-trityl or 4-hydroxymethyl-3-methoxyphenoybutyric acid (HMPB) SPPS linker, cleavable under conditions tolerated by the side chain protections and the N-terminal linker. Deprotection occurs concomitantly with the release of the peptide, upon TFA-mediated linker cleavage (Figure 23).

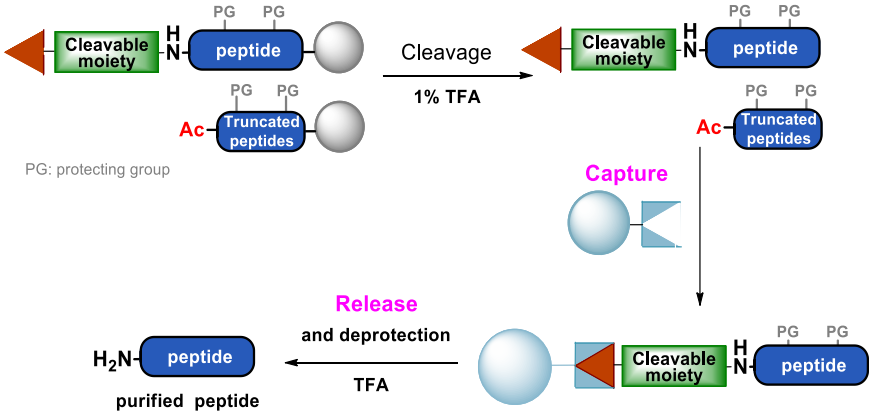


Figure 23. *Catch*-and-*release* purification using an acid-labile linker.

Five different benzyl carbamate structures have been proposed for that purpose (Figure 24), exploiting either thiol/disulfide exchange, thiol alkylation (linkers A-B),[17,59] or fluorous interactions (linkers D-F)[40] for the capture step.

Lansbury applied linkers A and B to a small fragment (9 amino acids) of the β amyloid protein which proved to be extremely difficult to purify in its unprotected form because of its β-sheet forming properties.[10]

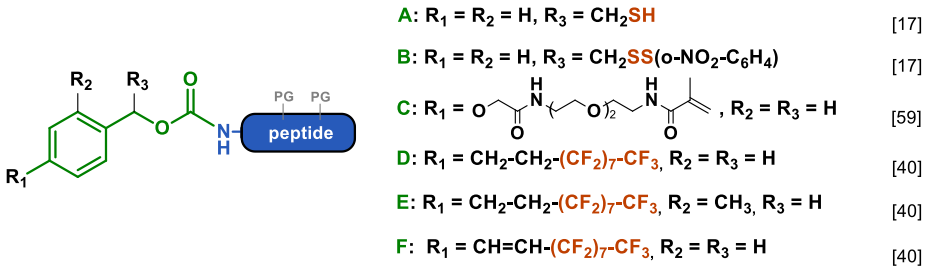


Figure 24. Acid-labile linkers.

3.5. Base-mediated cleavage

Two types of base-labile carbamate structures have been used, both releasing the peptide from the solid support through a β-elimination mechanism: fluorenylmethyl (Fmoc) and alkylsulfonylethyl carbamates. Due to the Michael acceptor properties of the resulting dibenzofulvenes and vinyl sulfones, respectively, internal cysteines need to be protected to avoid their nucleophilic attack.

3.5.1 Fluorenylmethyl carbamates

A wide range of Fmoc-derived N-terminal linkers has been described and applied to many challenging targets. In addition to the aforementioned Sulfmoc[22] and Tbfmoc,[34,35] respectively functionalized with a sulfonate group and fused benzene rings, biotinylated,[25,26,29] lipophilic,[23,25] and multicharged[23] versions have been developed by various groups. The liberation of the purified peptides from the solid support can be triggered by various aqueous conditions such as 1% Et_3N, 1% Na_2CO_3, 0.1 M NH_4OH or 0.1 M NaOH, and cleavage is also effective under organic conditions, typically morpholine or piperidine in dichloromethane or DMF. A compilation of the capture strategies, cleavage conditions, as well as some representative examples of purified peptides is presented in Table 1.

Table 1. Fmoc-derived N-terminal linkers.

Entry [Ref]	Immobilization strategy	Typical-cleavage conditions	Example of purified peptides	Number of AA
1[22]	Ionic interactions	Morpholine/CH_2Cl_2 Piperidine/CH_2Cl_2 Et_3N/H_2O	Model peptides	4-7
2[23]	Ionic interactions*	Et_3N/H_2O	cyclic analogue of the FMDV Vp1 capsid protein[63] (142-157) HIV-1 gp120 peptide[64]	15 25
3[34,35]	Π-stacking	Piperidine/CH_3CN Piperidine/6 M Gu.HCl/*i*PrOH	Hepatitis B surface antigen (1-23) ayw α-CGRP (8-37) [65] Bacteriophage λ-Ral[66] $(Acm)_4$ Gastrin releasing peptide[67] MBD from MeCP2 (78-162)[68]	23 30 66 77 85
4[29]	Avidin-biotin	Et_3N/H_2O	Rat cpn10[69]	101
5[30]	Hydrophobic interactions*	Piperidine/CH_3CN	Bacteriophage λ-Ral[66] $(Trt)_4$ MBD from MeCP2 (78-162)[68]	66 85
6[25]	Hydrophobic interactions*	Et_3N/H_2O	β2 subunit of CD18 [70] Rat cpn10[69]	46 101
7[23]	Hydrophobic interactions*	Et_3N/H_2O	FMDV Vp1 capsid protein (142-157) [63] HIV-1 gp120 peptide [64] HIV-1 gag p24[71]	15 25 104

Abbreviations: FMDV: food-and-mouse disease virus. HIV-1: human immunodeficiency virus type 1. CGRP: calcitonin gene-related peptide. MBD: methylated DNA binding domain. MeCP2: human methyl-CpG binding protein 2. Rat cpn 10: chaperonin 10 from *Rattus norvegicus.* Ral protein $(Trt)_4$ or $(Acm)_4$: restriction alleviation protein with trityl-protected Cys or acetamidomethyl.

***Strategies applied for the facilitation of the chromatographic purification (not *catch*-and-*release* purification)**

3.5.2. Sulfonylethyl carbamates

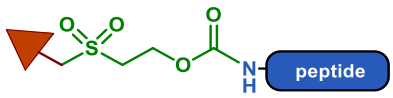


Figure 25. Alkylsulfonylethyl carbamate N-terminal linkers.

The methylsulfonylethyloxycarbonyl (Msc)[72] protective group can be easily derivatized and has consequently been used as a scaffold for the development of a large range of N-terminal linkers (Figure 25). Functionalization with a thiol,[43] a biotin,[28] a ketone,[44,73] a perfluoroalkyl group,[41] lipophilic groups[37] or an azide moiety[54] has been described. A compilation of the above-cited N-terminal linkers, the strategies used for the capture step, the cleavage conditions as well as some representative examples of purified peptides is presented in Table 2.

Table 2. Alkylsulfonylethyl carbamates N-terminal linkers

Entry [Ref]	Immobilization strategy	Typical-cleavage conditions	Example of purified peptide	Number of AA
1[43]	Thiol alkylation	NH_4-OH/ MeOH NaOH/ MeOH	Poly-phemusin II[74] $(Acm)_4$ hCCK[75] hGRF[76]	18 33 44
2[28]	Avidin-biotin	NaOH/ MeOH	Magainin 2[77] hGRF[76]	23 44
3[37]	Hydrophobic	NH_4-OH/TFE	Magainin 1[77]	23
4[44]	Oxime ligation	NaOH pH 12 hydrazine pH 14	model peptide	12
5[41]	Fluorous interactions	NH_4-OH	model peptides	6-35
6[54]	CuAAC	CAPS pH 11.7	AvBD2[78] $(Acm)_6$	36

Abbreviations: hCCKF: brain/gut peptide cholecystokinin. hGRF: hypothalamic growth hormone releasing factor. CAPS: N-Cyclohexyl-3-aminopropanesulfonic acid. AvBD2: chicken avian beta defensin 2, Cys protected by Acm groups.

3.6. Nucleophilic cleavage

Three types of N-terminal linkers cleavable upon treatments with nucleophiles have been reported in the literature, deriving from three amine protecting groups: the 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl group (Dde),[79] the 1,3-dimethyl-2,4,6-trioxopyrimidine-5-ylidene)methyl group (DTPM),[80,81] and the acetoacetyl group (AcA) (Figure 26).

Dde has been functionalized with four different binding moieties (biotin, lipophilic chains, cyclobutene and azide).[30,55,82] The azide-functionalized Dde-derived linker was found to be unstable under neutral and mildly acidic aqueous conditions,[82] thus, Aucagne and Delmas described another linker based on the closely related DTPM, the N_3-Dtpp which proved to be completely stable at pH 2-8.[82]

Both Dde and DTPM are stabilized by an intramolecular hydrogen bond.[83,84] The enamine moiety is stable towards acids and bases, thus completely compatible with Fmoc-or Boc-SPPS conditions. Both scaffolds can be readily cleaved through an aqueous hydrazine treatment. Alternatively, N_3-Dtpp can also be cleaved using aqueous hydroxylamine.[82]

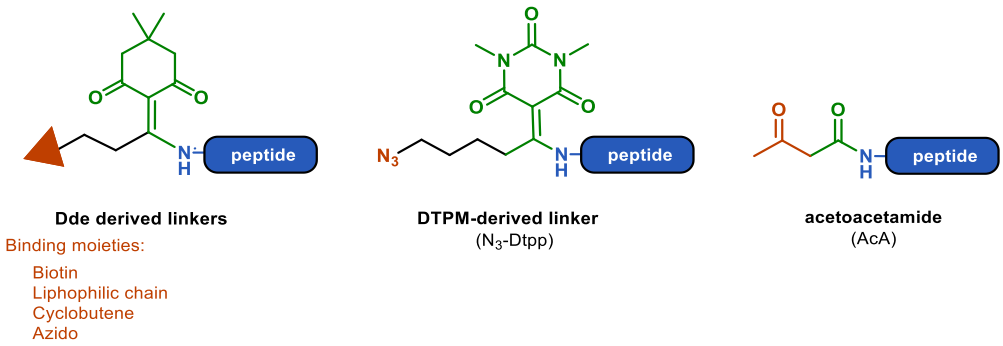


Figure 26. Linkers cleavable upon treatment with nucleophiles.

The *N*-acetoacyl (AcA) protecting group developed by Scoffone[85] has been recently proposed for the *catch*-and-*release* purification of peptides by Melnyk.[46] Advantageously, the ketone group of AcA can be used as the binding moiety, for immobilization through oxime ligation. Specific conditions are needed for the TFA-mediated cleavage from the SPPS resin to preserve the intact linker moiety from degradation.[86]

This N-terminal linker can be removed upon cleavage with hydroxylamine, through a transoximation followed by spontaneous intramolecular nucleophilic amide bond cleavage (Figure 27).

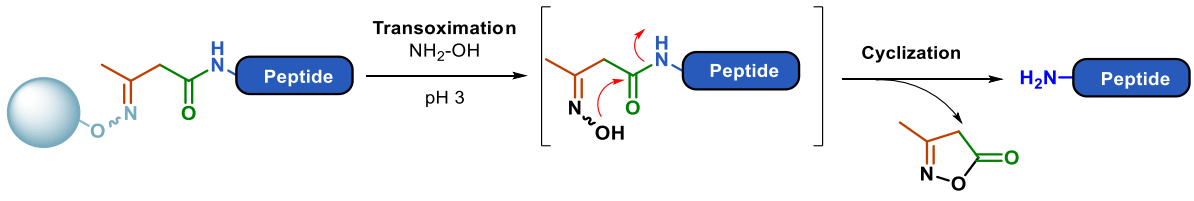


Figure 27. Transoximation-mediated AcA linker cleavage.

3.7. Transamination followed by nucleophilic cleavage

An alternative method has been proposed by Aimoto and co-workers [45] based on the use of a "safety-catch" nucleophile-cleavable moiety, inspired by the methodology developed by Dixon to remove the N-terminal residue of proteins.[87–89] In Aimoto´s strategy, an α-amino-amide precursor (*N*$^{\varepsilon}$-acyl-lysyl) linker is converted into an α-ketoamide linker *via* solid-phase transamination through a Nickel-catalyzed reaction with glyoxylic acid.[45,87] Then, cleavage of the linker is triggered by nucleophilic attack of the α-ketoamide with phenylenediamine under acidic aqueous conditions to liberate the peptide into solution (Figure 28).[45,88]

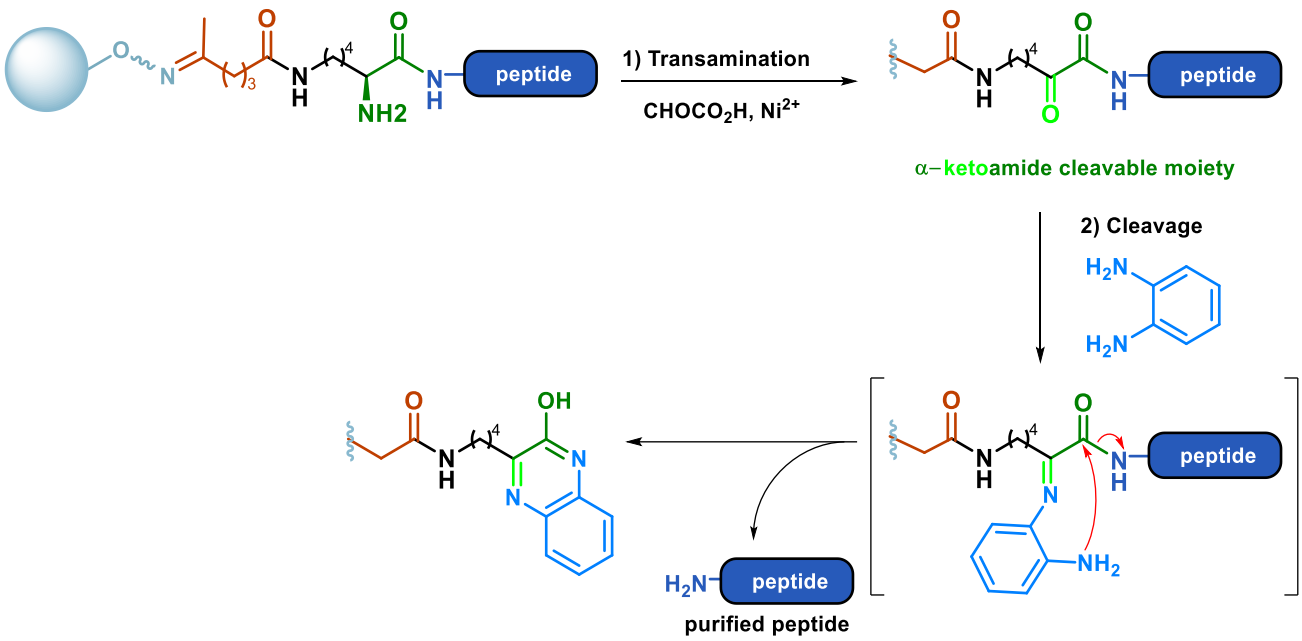


Figure 28. Two step-cleavage through transamination then nucleophilic cleavage.[45]

That strategy was illustrated with the purification of a highly hydrophobic transmembrane domain of opioid receptor-like 1 (41 AA) that could not be effectively purified by conventional chromatography methods due to its hydrophobicity and high tendency to aggregate.

3.8. Edman degradation

N-methoxythioureas can cleave the neighboring amide bond under highly acidic conditions through Edman degradation-type mechanism (Figure 29).[56] This method was applied to the

catch-and-*release* purification of the (240-272) loop region of the human dopamine D2 receptor long isoform (34 mer D2LR peptide).[90]

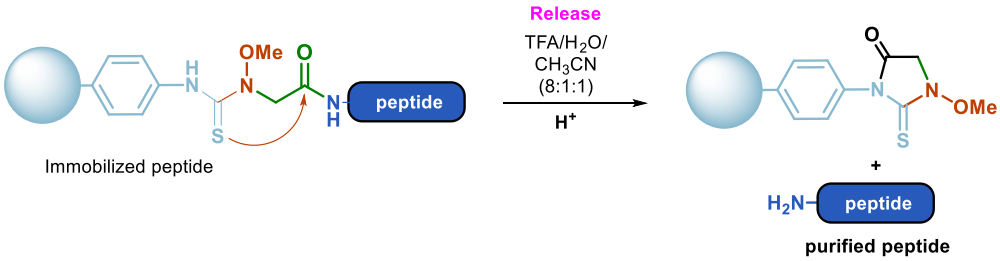


Figure 29. Edman degradation-mediated release.

Unlike the other described methods, there is not a well-differentiated binding and cleavable moiety in the N-terminal linker since the formed thiourea is responsible for the intramolecular nucleophilic attack.

3.9. Photocleavage

The photolabile 1-(2-nitrophenyl)-ethyl group (Figure 30), has been derivatized with a biotin moiety, and the linker applied to reversible immobilization of the pentapeptide Leu-enkephalin used as a model.[32] Very recently, a similar linker incorporating an azide binding group for a capture *via* strain-promoted alkyne-azide cycloaddition (SPAAC) has been developed application to the purification of oligonucleotides.[91]

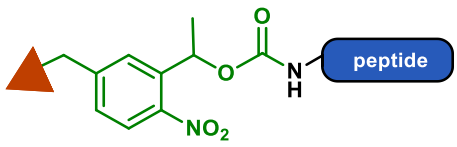


Figure 30. Photocleavable linkers.[32]

3.10. Autocatalytic cleavage

Heinrickson synthesized a 99 AA protease from simian immunodeficiency virus (SIV)[92,93] using a linker consisting in a N-terminally biotinylated sequence of eight amino acids (GGDRGFAA) which correspond to a propeptide sequence cleaved during the viral maturation. After the immobilization of the peptide *via* biotin-avidin interactions, the folding of the protease conducts to its activation and its own liberation from the solid support.[31]

4. Other related strategies

4.1. Internal resin capture

In 1997, Bradley introduced the concept of internal resin capture for the non-chromatographic purification of C-terminally modified peptides.[94] The technique was initially applied to the synthesis of C-terminal amides then extended to the synthesis of C-terminal peptide alcohols and nitriles.[95]

In the internal resin capture method, the *catch*-step occurs through a cyclization between the N-terminal linker and a reactive moiety included in the same SPPS resin. The SPPS linker is then cleaved under conditions leaving intact the N-terminal linker and the side-chain protecting groups. This results in the selective release of the truncated acetylated peptides from the resin. Bradley used the HMPB SPPS linker removable using 1% TFA, leading to a C-terminal carboxylic acid which can be further modified through coupling with different amines. After that, the purified C-terminally modified peptide can be deprotected and released into solution through TFA-mediated cleavage of the N-terminal linker (Figure 31).

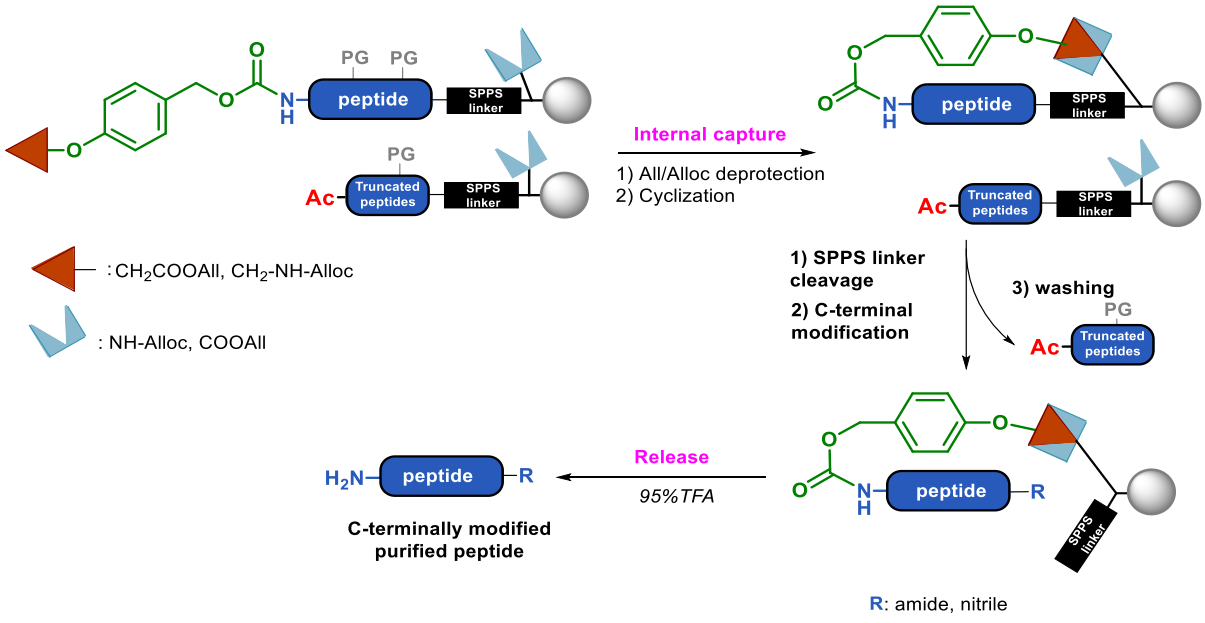


Figure 31. General principle of internal resin capture.

More recently, this strategy has been adapted by Seitz to the synthesis of C-terminal thioesters,[96–98] making use of the Kenner's sulfonamide "safety-catch" linker.[99,100] Cyclization

is also achieved through amide coupling, and SPPS linker cleavage proceeds concomitantly with thioester formation, through thiolysis of the alkylated Kenner's linker.[101]

4.2. Cap-tag strategy

The reverse "Cap-tag" strategy consists in tagging truncated sequences by using a capping reagent incorporating a binding moiety.[102] This agent has to rapidly react with the incomplete sequences and has to remain stable during the subsequent cycles of the SPPS. After cleavage from the resin, the tagged truncated peptides are separated from the non-tagged target peptide through their immobilization onto a solid support. In contrast with the standard *catch*-and-*release* purification methods, the target peptide remains in solution (Figure 32). This concept was originally introduced by Trudelle for the purification of peptide libraries, using biotin as the capping agent.[103] Kumar later proposed a fluorous capping agent compatible consisting in fluorinated trivalent iodonium salts.[104,105]

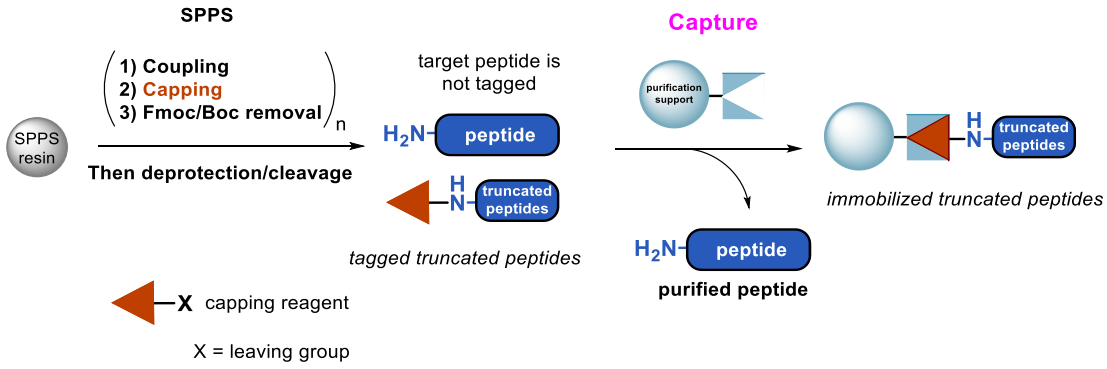


Figure 32. General principle of the "cap-tag" strategy.

5. Extension of the *catch*-and-*release* purification concept to solid supported chemical ligations with concomitant purifications

In the late 90´s, Kent introduced the solid supported chemical ligation (SPCL) concept for the total synthesis of very long peptides *via* successive chemical ligations of unprotected segment on a solid support.[106] The production of peptides through SPCL benefits for some important advantages compared to solution phase approaches, such as the avoidance of intermediate purifications, the facility to remove the excess of reagents, or the diminution in some cases of the peptide aggregation and precipitation. Several strategies have been

proposed for the total synthesis of peptide *via* SPCL in the "C-to-N" direction[73,107–111] and in "N-to-C" direction.[46,112]

Recently, Aucagne and Delmas proposed a "self-purification" method based on the *catch*- and *release* principle.[54] The first fragment is derivatized with an N-terminal linker and selectively captured onto a solid support and consequently purified. Then the successive N-to-C chemical ligations benefit from the same "self-purification" effect and can be carried out with unpurified peptide segments (Figure 33).

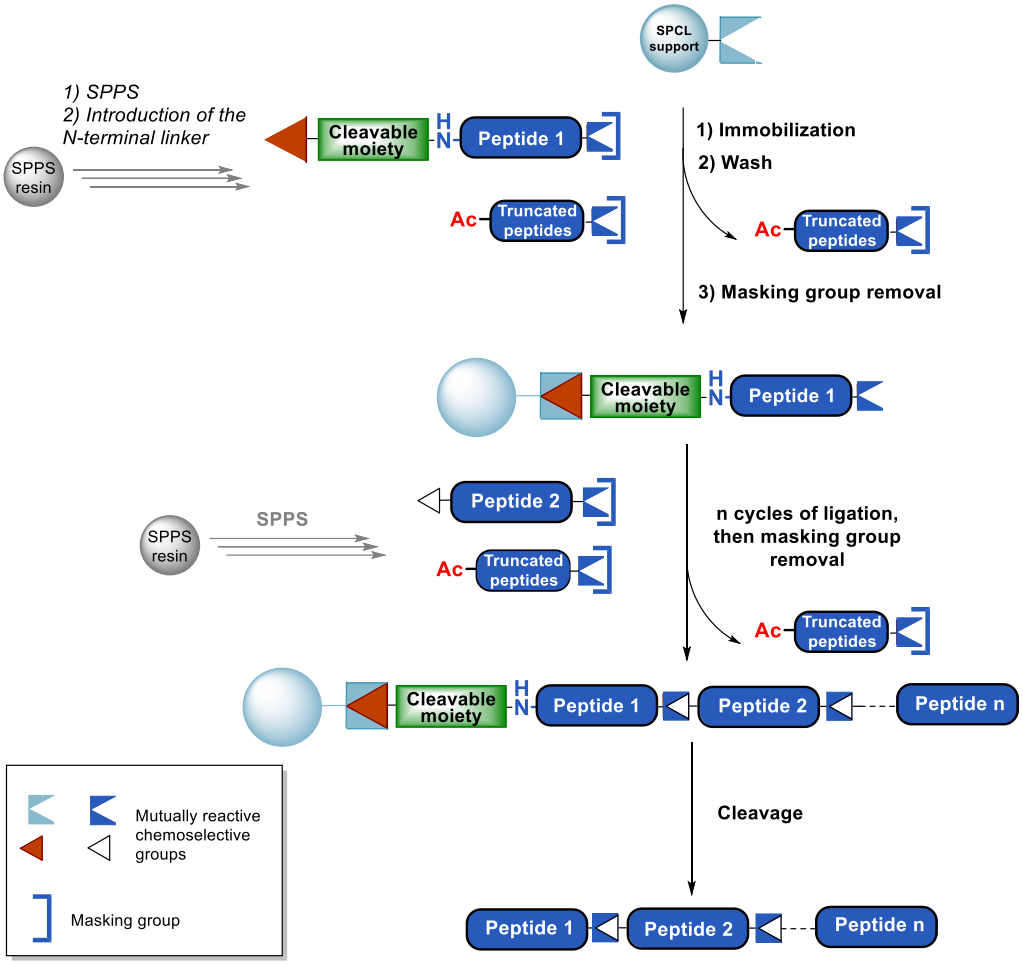


Figure 33. N-to-C solid supported chemical with concomitant purifications.

Similarly to successive ligations in solution, iterative ligations in the N-to-C direction require the temporary masking of the C-terminal chemoselectively reacting function of the peptide segments to avoid their cyclization or oligomerization.

The "self-purification" effect was first demonstrated with the synthesis of an analogue of the extracellular domain of the human mucin 1 (MUC1) peptide (160 amino acids, 8 tandem

repeat sequences.) MUC1 peptide was synthesized from four segments *via* N-to-C peptidomimetic triazole ligations.[113] For that purpose, they made use of the (2-[2-(2-azide-ethoxy)-ethyl-sulfonyl]-2-ethoxycarbonyl (N_3-Esoc) base-labile N-terminal linker (Figure 34). The methodology was later extended to the synthesis of *O*-glycosylated MUC1 peptides through solid supported enzymatic glycosylation, using the N_3-Dtpp linker which is compatible with the carbohydrates.[82]

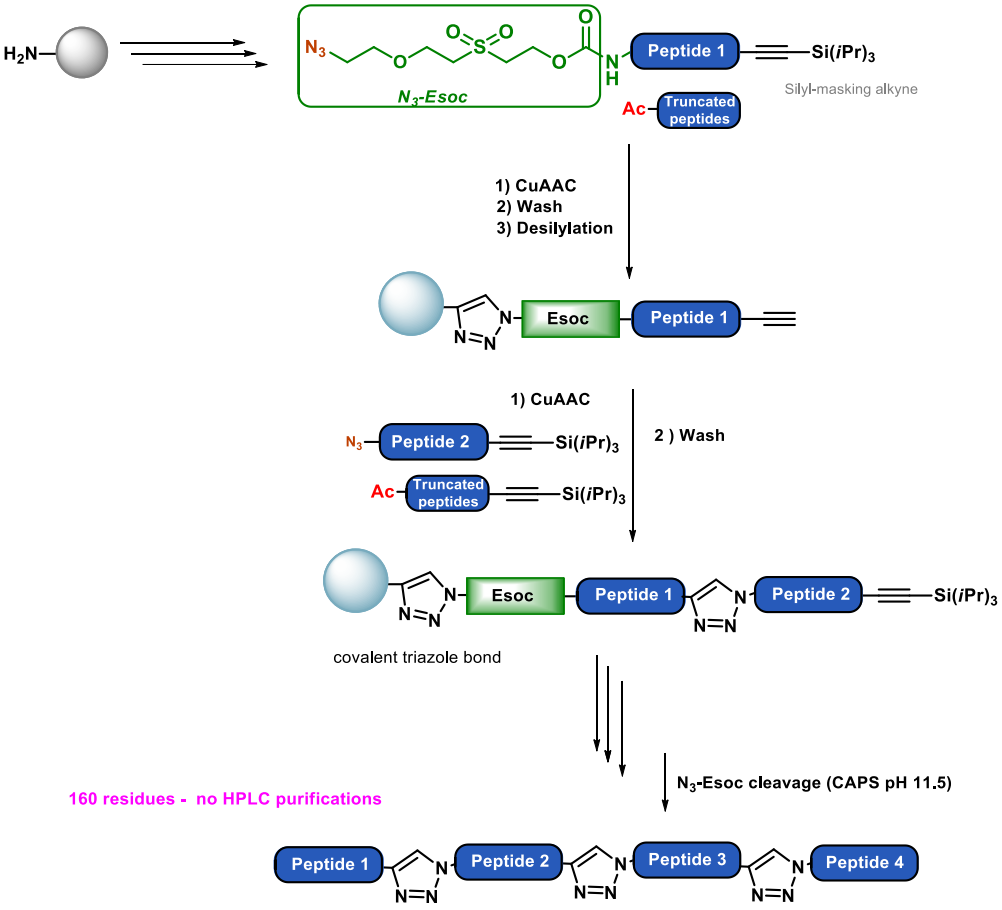


Figure 34. N-to-C solid supported peptidomimetic CuAAC ligations with concomitant purifications.

An alternative method that combines metal-ligand (Ni-NTA/His_6) and hydrazone capture with solid supported native chemical ligation (NCL) of two fragments, has been recently proposed by Seitz for the synthesis and HPLC-free purification of a 126 residues MUC1 peptide (Figure 35). [20] For that purpose, the C^{α}-peptide thioester and the cysteinyl peptide were respectively tagged with an N-terminal His_6-linker and a C-terminal hydrazide. Firstly, His_6-tagged peptide thioester is captured on a NTA-Ni agarose and isolated from the acetylated peptides originated during the SPPS by simple washing. Before cleavage, an auxiliary mediated NCL with the hydrazide-tagged cysteinyl peptide allowed to remove the

acetylated peptides from the SPPS of this C-terminal fragment. Non-ligated byproducts originating from the hydrolysis of the thioester during the NCL can be removed through a second capture *via* the hydrazide tag, onto an aldehyde-functionalized solid support. The purified peptide which contains the N-terminal His_6 and C-terminal hydrazide is finally liberated into solution by trans-hydrazonation. This method constitutes the first example of a double *catch*-and-*release* purification. However, contrarily to the previously proposed methods, the tags remain in the purified peptide.

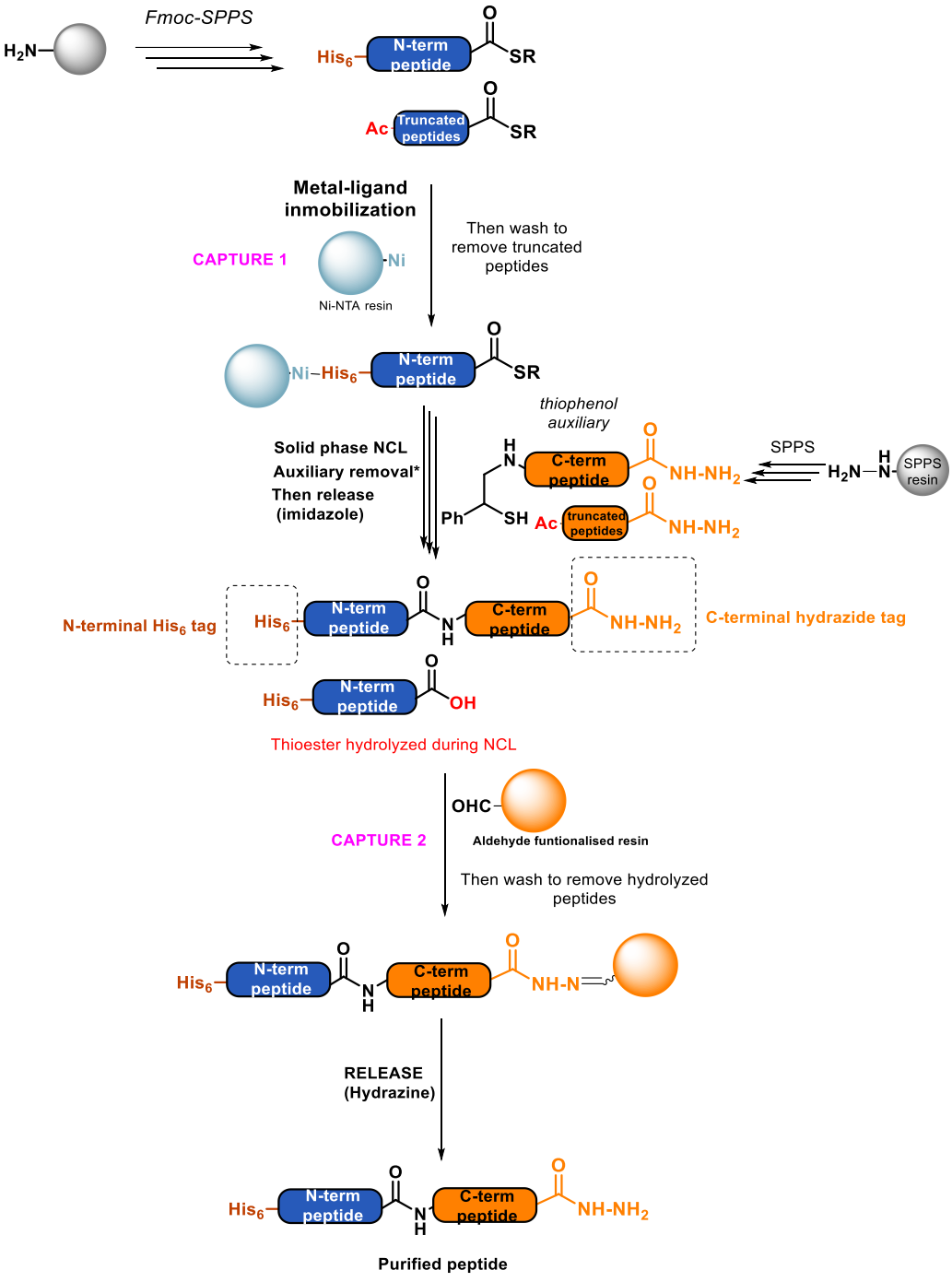


Figure 35. General principle of the Seitz´s double *catch*-and-*release* solid phase NCL strategy. *Auxiliary removal conditions: aqueous solution of TCEP (200 mM), and morpholine (800 M) pH = 8.5, 40°C.

6. Conclusion

The automated solid phase synthesis of small to medium-sized peptides, up to 50 amino acids is a well-established method. The SPPS byproducts, mostly truncated acetylated peptides, formed during the elongation are removed by HPLC to yield the purified target peptide. However, in some cases, the similar composition between such impurities and the peptide of interest may result in a co-elution after HPLC thus impairing the purification.

To overcome that limitation, *catch*-and-*release* purification methods, based on N-terminal linkers, have been developed as an alternative. In this review, we have studied in detail the different chemical functions found in these linkers which are responsible for the capture of the peptide onto a purification solid support, and then of the release of the purified peptide in solution.

These linkers have been also used for the extension of the concept to successive chemical ligations with self-purifications on a solid support which has allowed to afford long peptides and to limit or completely eliminate the HPLC-purifications. We have also described the existing linkers for these purposes. Although the implementation of the solid phase chemical ligation approach offers great advantages against the solution methods, there are only a few numbers of publications found in the literature. This fact can be due to an ensemble of reasons. The first one is the challenge of finding an adequate linker which is stable under the SPPS and the iterative chemical ligations conditions. Another issue is that the linker should allow the efficient immobilization of the peptide onto the solid support and the final cleavage under mild conditions. Other difficulties are related to the design of chemoselectively reacting functions that can be activated in a determinate step to allow the chemical ligation between the peptide segments.

The synthesis of long peptides *via* iterative NCLs on solid support offers the possibility to produce long natural peptides with self-purification. It is necessary to continue developing innovative tools which are completely adapted to these methodologies.

7. References chapter 1, part 2

(1) Fischer, E.; Fourneau, E. Ueber Einige Derivate Des Glykocolls. *Ber. Dtsch. Chem. Ges.* **1901**, *34* (2), 2868.

(2) Bavikar, S. N.; Spasser, L.; Haj-Yahya, M.; Karthikeyan, S. V.; Moyal, T.; Ajish Kumar, K. S.; Brik, A. Chemical Synthesis of Ubiquitinated Peptides with Varying Lengths and Types of Ubiquitin Chains to Explore the Activity of Deubiquitinases. *Angew. Chem. Int. Ed.* **2012**, *124* (3), 782.

(3) Wintermann, F.; Engelbrecht, S. Reconstitution of the Catalytic Core of F-ATPase (Aβ) 3 γ from Escherichia Coli Using Chemically Synthesized Subunit γ. *Angew. Chem. Int. Ed.* **2013**, *52* (4), 1309.

(4) Hemantha, H. P.; Bavikar, S. N.; Herman-Bachinsky, Y.; Haj-Yahya, N.; Bondalapati, S.; Ciechanover, A.; Brik, A. Nonenzymatic Polyubiquitination of Expressed Proteins. *J. Am. Chem. Soc.* **2014**, *136* (6), 2665.

(5) Weinstock, M. T.; Jacobsen, M. T.; Kay, M. S. Synthesis and Folding of a Mirror-Image Enzyme Reveals Ambidextrous Chaperone Activity. *Proc. Natl. Acad. Sci. USA.* **2014**, *111* (32), 11679.

(6) 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 Discov.* **2017**, *3*, 17008.

(7) Tang, S.; Liang, L.-J.; Si, Y.-Y.; Gao, S.; Wang, J.-X.; Liang, J.; Mei, Z.; Zheng, J.-S.; Liu, L. Practical Chemical Synthesis of Atypical Ubiquitin Chains by Using an Isopeptide-Linked Ub Isomer. *Angew. Chem. Int. Ed.* **2017**, *56* (43), 13333.

(8) Cuatrecasas, P.; Wilchek, M.; Anfinsen Christian B. Selective Enzyme Purification by Affinity Chromatography. *Proc. Natl. Acad. Sci. USA* **1968**, *61*, 636.

(9) *Affinity Chromatography Methods and Protocols. Second Edition*, Second.; Zachariou, M., Ed.; Humana Press: California, **2008**.

(10) Eldjarn, L.; Jellum, E. Organomercurial-Polysaccharide, a Chromatographic Material for the Separation and Isolation of SH-Proteins. *Acta Chem. Scand.* **1963**, *17*, 2610.

(11) Kriege, D. E.; Erickson, W.; Merrifield, R. B. Affinity Purification of Synthetic Peptides. *Proc. Nati. Acad. Sci. USA.* **1976**, *73* (9), 3160.

(12) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149.

(13) Ball, H. L.; Mascagni, P. N-(2-Chlorobenzyloxycarbonyloxy)-Succinimide as a Terminating Agent for Solid-Phase Peptide Synthesis: Application to a One-Step Purification Procedure. *Lett. Pept. Sci.* **1995**, *2* (1), 49.

(14) Nordlund, P.; Weigelt, J.; Bray, J.; Gileadi, O.; Knapp, S.; Gräslund, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J. Protein Production and Purification. *Nat. Methods*

2008, *5* (2), 135.

(15) Terpe, K. Overview of Tag Protein Fusions: From Molecular and Biochemical Fundamentals to Commercial Systems. *Appl. Microbiol. Biotechnol.* **2003**, *60* (5), 523.

(16) Kosobokova, E. N.; Skrypnik, K. A.; Kosorukov, V. S. Overview of Fusion Tags for Recombinant Proteins. *Biochem.* **2016**, *81* (3), 187.

(17) Sucholeiki, I.; Lansbury, P. T. An Affinity Chromatographic Method for the Purification of Water-Insoluble Peptides. *J. Org. Chem.* **1993**, *58* (6), 1318.

(18) Roggero, M. A.; Servis, C.; Corradin, G. A Simple and Rapid Procedure for the Purification of Synthetic Polypeptides by a Combination of Affinity Chromatography and Methionine Chemistry. *FEBS Lett.* **1997**, *408* (3), 285.

(19) Bang, D.; Kent, S. B. His6 Tag-Assisted Chemical Protein Synthesis. *Proc. Nati. Acad. Sci. USA.* **2005**, *102* (14), 5014.

(20) Loibl, S. F.; Harpaz, Z.; Zitterbart, R.; Seitz, O. Total Chemical Synthesis of Proteins without HPLC Purification. *Chem. Sci.* **2016**, *7* (11), 6753.

(21) Carpino, L. A.; Han, G. Y. 9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group. *J. Am. Chem. Soc.* **1970**, *92* (19), 5748.

(22) Merrifield, R. B.; Bach, A. E. 9-(2-Sulfo)fluorenylmethyloxycarbonyl Chloride, a New Reagent for the Purification of Synthetic Peptides. *J. Org. Chem.* **1978**, *43* (25), 4808.

(23) Ball, H. L.; Mascagni, P. Purification of Synthetic Peptides Using Reversible Chromatographic Probes Based on the Fmoc Molecule. *Int. J. Pept. Protein Res.* **1992**, *40* (5), 370.

(24) Lobl, T. J.; Deibel, M. R.; Yem, A. W. On-Resin Biotinylation of Chemically Synthesized Proteins for One-Step Purification. *Anal. Biochem.* **1988**, *170* (2), 502.

(25) Ball, H. L.; Bertolini, G.; Levi, S.; Mascagni, P. Purification of Synthetic Peptides with the Aid of Reversible Chromatographic Probes. *J. Chromatogr. A* **1994**, *686* (1), 73.

(26) Mascagni, P.; Ball, H. L.; Bertolini, G. Selective Purification of Synthetic Proteins by the Use of FMOC- and Biotin-Based Reversible Chromatographic Probes. *Anal. Chim. Acta* **1997**, *352* (1–3), 375.

(27) Bianchi, E.; Sollazzo, M.; Tramontano, A.; Pessi, A. Affinity Purification of a Difficult-Sequence Protein. *Int. J. Pept. Protein Res.* **1993**, *42* (1), 93.

(28) Funakoshi, S.; Fukuda, H.; Fujii, N. Affinity Purification Method Using a Reversible Biotinylating Reagent for Peptides Synthesized by the Solid-Phase Technique. *J. Chromatogr. A* **1993**, *638* (1), 21.

(29) Ball, H. L.; Bertolini, G.; Mascagni, P. Affinity Purification of 101 Residue Rat cpn10 Using a Reversible Biotinylated Probe. *J. Pept. Sci.* **1995**, *1* (5), 288.

(30) Kellam, B.; Chan, W. C.; Chhabra, S. R.; Bycroft Barrie W. Transient Affinity Tags Based

on the Dde Protection/Deprotection Strategy: Synthesis and Application of 2-Biotinyl- and 2-Hexanoyidimedone. *Tetrahedron Lett.* **1997**, *38* (30), 5391.

(31) Tomasselli, A. G.; Bannow, C. A.; Deibel, M. R.; Hui, J. O.; Zurcher-Neely, H. A.; Reardon, I. M.; Smith, C. W.; Heinrikson, R. L. Chemical Synthesis of a Biotinylated Derivative of the Simian Immunodeficiency Virus Protease. Purification by Avidin Affinity Chromatography and Autocatalytic Activation. *J. Biol. Chem.* **1992**, *267* (15), 10232.

(32) Olejnik, J.; Sonar, S.; Krzymañska-Olejnik, E.; Rothschild, K. J. Photocleavable Biotin Derivatives: A Versatile Approach for the Isolation of Biomolecules. *Proc. Natl. Acad. Sci. USA.* **1995**, *92* (16), 7590.

(33) Ramage, R.; Raphy, G. Design on an Affinity-Based Nα-Amino Protecting Group for Peptide Synthesis: Tetrabenzo[a,c,g,i]fluorenyl-17-Methyl Urethanes (Tbfmoc). *Tetrahedron Lett.* **1992**, *33* (3), 385.

(34) Brown, A. R.; Irving, S. L.; Ramage, R. Affinity Purification of Synthetic Peptides and Proteins on Porous Graphitised Carbon. *Tetrahedron Lett.* **1993**, *34* (44), 7129.

(35) Brown, A. R.; Irving, S. L.; Ramage, R.; Raphy, G. (17-Tetrabenzo[a,c,g,i]fluorenyl)methylchloroformate (TbfmocCl) a Reagent for the Rapid and Efficient Purification of Synthetic Peptides and Proteins. *Tetrahedron* **1995**, *51* (43), 11815.

(36) Liu, Q.; Jin, Z.; Hu, N. F.; Wang, P. Tbfmoc Methioine: Its Synthesis and Application for the Solid Phase Synthesis and Purification of Ubiquitin C-Hydrazide. *Asian J. Chem.* **2013**, *25* (2).

(37) García-Echeverría, C. On the Use of Hydrophobic Probes in the Chromatographic Purification of Solid-Phase-Synthesized Peptides. *J. Chem. Soc., Chem. Commun.* **1995**, *34* (7), 779.

(38) Zhang, W. Fluorous Tagging Strategy for Solution-Phase Synthesis of Small Molecules, Peptides and Oligosaccharides. *Curr. Opin. Drug Discov. Devel.* **2004**, *7* (6), 784.

(39) Gladysza, J. .; Curran, D. P. Fluorous Chemistry: From Biphasic Catalysis to a Parallel Chemical Universe and beyond. *Tetrahedron* **2002**, *58* (20), 3823.

(40) Filippov, D. V.; van Zoelen, D. J.; Oldfield, S. P.; van der Marel, G. A.; Overkleeft, H. S.; Drijfhout, J. W.; van Boom, J. H. Use of Benzyloxycarbonyl (Z)-Based Fluorophilic Tagging Reagents in the Purification of Synthetic Peptides. *Tetrahedron Lett.* **2002**, *43* (43), 7809.

(41) de Visser, P. C.; van Helden, M.; Filippov, D. V; van der Marel, G. A.; Drijfhout, J. W.; van Boom, J. H.; Noort, D.; Overkleeft, H. S. A Novel, Base-Labile Fluorous Amine Protecting Group: Synthesis and Use as a Tag in the Purification of Synthetic Peptides. *Tetrahedron Lett.* **2003**, *44* (50), 9013.

(42) Lindeberg, G.; Tengborn, J.; Bennich, H.; Ragnarsson, U. Purification of a Synthetic Peptide with the Aid of Covalent Chromatography. *J. Chromatogr. A* **1978**, *156* (2),

366.

(43) Funakoshi, S.; Fukuda, H.; Fujii, N. Chemoselective One-Step Purification Method for Peptides Synthesized by the Solid-Phase Technique. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 6981.

(44) Canne, L. E.; Winston, R. L.; Ken, S. B. H. Synthesis of a Versatile Purification Handle for Use with Boc Chemistry Solid Phase Peptide Synthesis. *Tetrahedron Lett.* **1997**, *38* (19), 3361.

(45) Hara, T.; Tainosho, A.; Nakamura, K.; Sato, T.; Kawakami, T.; Aimoto, S. Peptide Purification by Affinity Chromatography Based on Alpha-Ketoacyl Group Chemistry. *J. Pept. Sci.* **2009**, *15* (5), 369.

(46) Ollivier, N.; Desmet, R.; Drobecq, H.; Blanpain, A.; Boll, E.; Leclercq, B.; Mougel, A.; Vicogne, J.; Melnyk, O. A Simple and Traceless Solid Phase Method Simplifies the Assembly of Large Peptides and the Access to Challenging Proteins. *Chem. Sci.* **2017**, *11*, 135.

(47) a) Reimann, O.; Regnery, A.; Sarma, D.; Zitterbart, R. HPLC-Free Peptide Purification.A General Method for Potential Large-Scale and High-Throughput Purification of Chemically Synthesized Peptides by Catch & Release Filtration. **2017**, 6th Chemical Protein Synthesis meeting (poster presentation). http://www.belyntic.com/wp-content/uploads/2017/09/Poster_Haifa.pdf b) Zitterbart, R; Seitz, O. **2017**, *EPA patent,* WO 2017/129818A1

(48) Villain, M.; Vizzavona, J.; Rose, K. Covalent Capture: A New Tool for the Purification of Synthetic and Recombinant Polypeptides. *Chem. Biol.* **2001**, *8* (7), 673.

(49) Giron, P.; Dayon, L.; David, F.; Sanchez, J. C.; Rose, K. Enrichment of N-Terminal Cysteinyl-Peptides by Covalent Capture. *J Proteomics* **2009**, *71* (6), 647.

(50) Vizzavona, J.; Villain, M.; Rose, K. Covalent Capture Purification of Polypeptides after SPPS via a Linker Removable under Very Mild Conditions. *Tetrahedron Lett.* **2002**, *43*, 8693.

(51) 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.

(52) 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.

(53) Tang, W.; Becker, M. L. 'Click' reactions: A Versatile Toolbox for the Synthesis of Peptide-Conjugates. *Chem. Soc. Rev.* **2014**, *43* (20), 7013.

(54) 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.

(55) Jacob, A.; Hoheisel, J.; Dauber, M.; Wiessler, M.; Lorenz, P.; Fleischhacker, H.; Kliem, H.-C. Dimedon Derivative and a Method for the Purification of PNA and Peptide Oligomers. U. S. Patent 0,274,440, Oct 17, 2013.

(56) Hara, T.; Tainosyo, A.; Kawakami, T.; Aimoto, S.; Murata, M. Peptide Purification Using the Chemoselective Reaction between N-(Methoxy)glycine and Isothiocyanato-Functionalized Resin. *J. Pept. Sci.* **2016**, *22* (6), 379.

(57) Toshiyuki, M.; Toshifumi, T. Selective Isolation of N-Blocked Peptides by Isocyanate-Coupled Resin. *Anal. Chem.* **2007**, *79*, 7910.

(58) Hara, T.; Purwati, E. M.; Tainosyo, A.; Kawakami, T.; Hojo, H.; Aimoto, S. Site-Specific Labeling of Synthetic Peptide Using the Chemoselective Reaction between *N* -Methoxyamino Acid and Isothiocyanate. *J. Pept. Sci.* **2015**, *21* (10), 765.

(59) Zhang, M.; Pokharel, D.; Fang, S. Purification of Synthetic Peptides Using a Catching Full-Length Sequence by Polymerization Approach. *Org. Lett.* **2014**, *16* (5), 1290.

(60) Schroeder, W. A.; Shelton, J. B.; Shelton, J. R. An Examination of Conditions for the Cleavage of Polypeptide Chains with Cyanogen Bromide: Application to Catalase. *Arch. Biochem. Biophys.* **1969**, *130*, 551.

(61) Lanar, D. E. Sequence of the Circumsporozoite Gene of Plasmodium Berghei ANKA Clone and NK65 Strain. *Mol. Biochem. Parasitol.* **1990**, *39* (1), 151.

(62) Nicolet, B. H.; Shinn, L. A. The Action of Periodic Acid on α-Amino Alcohols. *J. Am. Chem. Soc.* **1939**, *61* (6), 1615.

(63) Bittle, J. L.; Houghten, R. A.; Alexander, H.; Shinnick, T. M.; Sutcliffe, J. G.; Lerner, R. A.; Rowlands, D. J.; Brown, F. Protection against Foot-and-Mouth Disease by Immunization with a Chemically Synthesized Peptide Predicted from the Viral Nucleotide Sequence. *Nature.* **1982**, *298* (5869), 30.

(64) Robson, B.; Fishleigh, R.; Morrison, C. A. Prediction of HIV Vaccine. *Nature* **1987**, *325* (6103), 395.

(65) Wang, H.; Xing, L.; Li, W.; Hou, L.; Guo, J.; Wang, X. Production and Secretion of Calcitonin Gene-Related Peptide from Human Lymphocytes. *J. Neuroimmunol.* **2002**, *130* (1), 155.

(66) Wilhelmine, A. M.; Murray, N. E. Modification Enhancement by the Restriction Alleviation Protein (Ral) of Bacteriophage λ. *J. Mol. Biol.* **1986**, *190* (1), 11.

(67) Spindel, E. R.; Chin, W. W.; Price, J.; Rees, L. H.; Besser, G. M.; Habener, J. F. Cloning and Characterization of cDNAs Encoding Human Gastrin-Releasing Peptide. *Proc. Natl. Acad. Sci. USA* **1984**, *81* (18), 5699.

(68) Lewis, J. D.; Mechan, R. R.; Henzel, W. J.; Maurer-Fogy, Ingrid; Jeppesen, P.; Klein, F.; Bird, A. Purification, Sequence, and Cellular Localization of a Novel Chromosomal Protein That Binds to Methylated DNA. *Cell* **1992**, *69* (6), 905.

(69) Hartman, D. J.; Hoogenraad, N. J.; Condron, R.; Høj, P. B. The Complete Primary

Structure of Rat Chaperonin 10 Reveals a Putative Beta Alpha Beta Nucleotide-Binding Domain with Homology to p21ras. *Biochim. Biophys. Acta* **1993**, *1164* (2), 219.

(70) Weitzman, J. B.; Wells, C. E.; Wright, A. H.; Clark A., P.; Law Alex, S. K. The Gene Organisation of the Human β2 Integrin Subunit (CD18). *FEBS Lett.* **1991**, *294* (1-2), 97.

(71) Mascagni, P.; Sia, D. Y.; Coates, A. R.; Gibbons, W. A. Protein Engineering of Hiv Viral Proteins by Total Chemical Synthesis: The C-Terminal 104 Residue Peptide from Gag p24. *Tetrahedron Lett.* **1990**, *31* (32), 4637.

(72) Tesser, G. I.; Balvert-Geers, I. C. The Methylsulfonylethyloxycarbonyl Group, a New and Versatile Amino Protective Function. *Int. J. Pept. Protein Res.* **1975**, *7* (4), 295.

(73) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121* (38), 8720.

(74) Miyata, T.; Tokunaga, F.; Yoneya, T.; Yoshikawa, K.; Iwanaga, S.; Niwa, M.; Takao, T.; Shimonishi, Y. Antimicrobial Peptides, Isolated from Horseshoe Crab Hemocytes, Tachyplesin II, and Polyphemusins I and II: Chemical Structures and Biological Activity. *J. Biochem.* **1989**, *106* (4), 663.

(75) Takahashi, Y.; Kato, K.; Hayashizaki, Y.; Wakabayashi, T.; Ohtsuka, E.; Matsuki, S.; Ikehara, M.; Matsubara, K. Molecular Cloning of the Human Cholecystokinin Gene by Use of a Synthetic Probe Containing Deoxyinosine. *Proc. Natl. Acad. Sci. USA* **1985**, *82* (7), 1931.

(76) Brazeau, P.; Ling, N.; Böhlen, P.; Esch, F.; Ying, S. Y.; Guillemin, R. Growth Hormone Releasing Factor, Somatocrinin, Releases Pituitary Growth Hormone in Vitro. *Proc. Natl. Acad. Sci. USA* **1982**, *79* (24), 7909.

(77) Zasloff, M. Magainins, a Class of Antimicrobial Peptides from Xenopus Skin: Isolation, Characterization of Two Active Forms, and Partial cDNA Sequence of a Precursor. *Proc. Natl. Acad. Sci. USA* **1987**, *84* (15), 5449.

(78) Derache, C.; Meudal, H.; Aucagne, V.; Mark, K. J.; Cadène, M.; Delmas, A. F.; Lalmanach, A.-C.; Landon, C. Initial Insights into Structure-Activity Relationships of Avian Defensins. *J. Biol. Chem.* **2012**, *287* (10), 7746.

(79) Bycroft, B. W.; Chan, W. C.; Chhabra, S. R.; Teesdale-Spittle, P. H.; Hardy, P. M. A Novel Amino Protection-Deprotection Procedure and Its Application in Solid Phase Peptide Synthesis. *J. Chem. Soc., Chem. Commun.* **1993**, *81* (9), 776.

(80) Dekany, G.; Bornaghi, L.; Papageorgiou, J.; Taylor, S. A Novel Amino Protecting Group: DTPM. *Tetrahedron Lett.* **2001**, *42* (17), 3129.

(81) da Silva, E. T.; Lima, E. L. S. Reaction of 1,3-Dimethyl-5-Acetyl-Barbituric Acid (DAB) with Primary Amines. Access to Intermediates for Selectively Protected Spermidines. *Tetrahedron Lett.* **2003**, *44* (18), 3621.

(82) 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*, 3617.

(83) Nash, I. A.; Bycroft, B. W.; Chan, W. C. Dde - A Selective Primary Amine Protecting Group: A Facile Solid Phase Synthetic Approach to Polyamine Conjugates. *Tetrahedron Lett.* **1996**, *37* (15), 2625.

(84) Sharma, A.; Jad, Y.; Ghabbour, H.; de la Torre, B.; Kruger, H.; Albericio, F.; El-Faham, A. Synthesis, Crystal Structure and DFT Studies of 1,3-Dimethyl-5-Propionylpyrimidine-2,4,6(1H,3H,5H)-Trione. *Crystals* **2017**, *7* (1), 31.

(85) Filira, F.; D'Angeli, F.; Ernesto, S. The Acetoacetyl Group, an Amino Protective Group of Potential Use in Peptide Synthesis. *Tetrahedron Lett.* **1965**, *6* (10), 605.

(86) Boll, E.; Ebran, J.-P.; Drobecq, H.; El-Mahdi, O.; Raibaut, L.; Ollivier, N.; Melnyk, O. Access to Large Cyclic Peptides by a One-Pot Two-Peptide Segment Ligation/Cyclization Process. *Org. Lett.* **2015**, *17* (1), 130.

(87) Dixon, H. Transamination of Peptides. *Biochem. J.* **1964**, *92* (3), 661.

(88) Dixon, H. B. F. Removal of the N-Terminal Residue of a Peptide after Transamination. *Biochem. J.* **1964**, *90* (2).

(89) Dixon, H. B. F. N-Terminal Modification of Proteins?a Review. *J. Protein Chem.* **1984**, *3* (1), 99.

(90) Shioda, N.; Takeuchi, Y.; Fukunaga, K. Advanced Research on Dopamine Signaling to Develop Drugs for the Treatment of Mental Disorders: Proteins Interacting with the Third Cytoplasmic Loop of Dopamine D2 and D3 Receptors. *J. Pharmacol. Sci.* **2010**, *114* (1), 25.

(91) Igata, Y.; Saito-Tarashima, N.; Matsumoto, D.; Sagara, K.; Minakawa, N. A 'catch and Release' Strategy towards HPLC-Free Purification of Synthetic Oligonucleotides by a Combination of the Strain-Promoted Alkyne-Azide Cycloaddition and the Photocleavage. *Bioorg. Med. Chem.* **2017**, *In press*.

(92) Kestler, H.; Kodama, T.; Ringler, D.; Marthas, M.; Pedersen, N.; Lackner, A.; Regier, D.; Sehgal, P.; Daniel, M.; King, N. Induction of AIDS in Rhesus Monkeys by Molecularly Cloned Simian Immunodeficiency Virus. *Science.* **1990**, *248* (4959), 1109.

(93) Lackner, A. A.; Smith, M. O.; Munn, R. J.; Martfeld, D. J.; Gardner, M. B.; Marx, P. A.; Dandekar, S. Localization of Simian Immunodeficiency Virus in the Central Nervous System of Rhesus Monkeys. *Am. J. Pathol.* **1991**, *139* (3), 609.

(94) Davies, M.; Bradley, M. C-Terminally Modified Peptides and Peptide Libraries-Another End to Peptide Synthesis. *Angew. Chem. Int. Ed.* **1997**, *36* (10), 1097.

(95) Davies, M.; Bradley, M. Internal Resin Capture - A Self Purification Method for the Synthesis of C-Terminally Modified Peptides. *Tetrahedron* **1999**, *55* (15), 4733.

(96) Mende, F.; Seitz, O. Solid-Phase Synthesis of Peptide Thioesters with Self-Purification. *Angew. Chem. Int. Ed.* **2007**, *46* (24), 4577.

(97) Mende, F.; Beisswenger, M.; Seitz, O. Automated Fmoc-Based Solid-Phase Synthesis of Peptide Thioesters with Self-Purification Effect and Application in the Construction of Immobilized SH3 Domains. *J. Am. Chem. Soc.* **2010**, *132* (32), 11110.

(98) Wade, W. S.; Yang, F.; Sowin, T. J. Application of Base Cleavable Safety Catch Linkers to Solid Phase Library Production. *J. Comb. Chem. 2* (3), 266.

(99) Kenner, G. W.; McDermott, J. R.; Sheppard, R. C. The Safety Catch Principle in Solid Phase Peptide Synthesis. *J. Chem. Soc. D Chem. Commun.* **1971**, *12*, 636.

(100) Bradley J. Backes; Ellman, J. A. An Alkanesulfonamide 'Safety-Catch' Linker for Solid-Phase Synthesis. *J. Org. Chem.* **1999**, *64* (7), 2322.

(101) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc/t-Bu Chemistry. *J. Am. Chem. Soc.* **1999**, *121* (49), 11369.

(102) Quesnel, A.; Delmas, A.; Trudelle, Y. Purification of Synthetic Peptide Libraries by Affinity Chromatography Using the Avidin-Biotin System. *Anal. Biochem.* **1995**, *231* (1), 182.

(103) Palmacci, E. R.; Hewitt, M. C.; Seeberger, P. H. 'Cap-Tag'-Novel Methods for the Rapid Purification of Oligosaccharides Prepared by Automated Solid-Phase Synthesis. *Angew. Chem. Int. Ed.* **2001**, *40* (23), 4433.

(104) Montanari, V.; Kumar, K. Just Add Water: A New Fluorous Capping Reagent for Facile Purification of Peptides Synthesized on the Solid Phase. *J. Am. Chem. Soc.* **2004**, *126* (31), 9528.

(105) Montanari, V.; Kumar, K. Enabling Routine Fluorous Capping in Solid Phase Peptide Synthesis. *J. Fluor. Chem.* **2006**, *127* (4-5), 565.

(106) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121* (38), 8720.

(107) Brik, A.; Keinan, E.; Dawson, P. E. Protein Synthesis by Solid-Phase Chemical Ligation Using a Safety Catch Linker. *J. Org. Chem.* **2000**, *65* (12), 3829.

(108) Cotton, G. J.; Muir, T. W. Generation of a Dual-Labeled Fluorescence Biosensor for Crk-II Phosphorylation Using Solid-Phase Expressed Protein Ligation. *Chem. Biol.* **2000**, *7* (4), 253.

(109) Jbara, M.; Seenaiah, M.; Brik, A. Solid Phase Chemical Ligation Employing a Rink Amide Linker for the Synthesis of Histone H2B Protein. *Chem. Commun.* **2014**, *50* (83), 12534.

(110) Yu, R. R.; Mahto, S. K.; Justus, K.; Alexander, M. M.; Howard, C. J.; Ottesen, J. J. Hybrid Phase Ligation for Efficient Synthesis of Histone Proteins. *Org. Biomol. Chem.* **2016**, *14* (9), 2603.

(111) Decostaire, I. E.; Lelièvre, D.; Aucagne, V.; Delmas, A. F. Solid Phase Oxime Ligations for the Iterative Synthesis of Polypeptide Conjugates. *Org. Biomol. Chem.* **2014**, *12* (29), 5536.

(112) Raibaut, L.; Adihou, H.; Desmet, R.; Delmas, A. F.; Aucagne, V.; Melnyk, O. Highly Efficient Solid Phase Synthesis of Large Polypeptides by Iterative Ligations of bis(2-Sulfanylethyl)amido (SEA) Peptide Segments. *Chem. Sci.* **2013**, *4* (10), 4061.

(113) Valverde, I. E.; Lecaille, 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.

Chapter 2

Chapter 2. Development of an N-terminal linker for the synthesis of disulfide-rich peptides through capture-and-release purification and solid supported NCLs

1. Biologically-relevant disulfide-rich peptides

1.1 Disulfide-rich peptides in nature

Disulfide-rich peptides (DRPs) are naturally occurring constrained peptides of around 15 to 100 amino acids, typically containing more than 10% of cysteine residues. They form multiple disulfide bonds which result in highly constrained and well-defined three-dimensional structure.[1–3]

They are found in animal and plant kingdoms showing a very large range of different biological activities and properties (most of them are related to predation and defense roles). On a lesser extent, they are also found in bacteria and fungi.

Due to their high biological and chemical stability, combined with their ability to modulate therapeutically-relevant targets, they are considered as promising drug candidates and pharmacological tools.[2]

Hereafter, a few representative examples of naturally occurring DRPs are briefly described.

a) Defensins

Defensins are disulfide-rich cationic miniproteins found in vertebrates, invertebrates, and plants.[4,5] These host defense antimicrobial peptides are part of the innate immune system, playing a central role in the defense against pathogens such as bacteria, fungi and viruses. They are considered as potential alternative to current antibiotic treatments.[1,5–8]

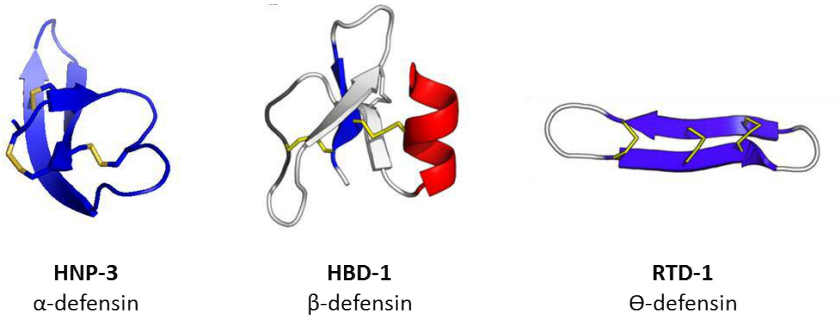
They typically consist of 18-50 amino acids including six to eight cysteine residues forming three or four disulfide bonds. α, β and Θ defensins constitute the three subfamilies of defensins in vertebrates. The latter, Θ defensins, are specific to non-human primate and

possess a cyclic backbone. All three classes contain six cysteines and differ from their disulfide connectivity (Table 1).[4]

Table 1. Similarities and differences between the three families of peptide defensins.[4]

Defensin	Aas	Disulfide connections
Θ-defensin	18	I-IV, II-V, II-VI
β-defensin	32-42	I-VI, II-IV, III-V
α-defensin	29-45	I-V, II-IV, III-VI

Three representative examples of each family are shown in figure 1.



HNP-3
α-defensin

HBD-1
β-defensin

RTD-1
Θ-defensin

Figure 1. Three-dimensional structures of the dimer human neutrophil peptide 3 (HNP-3)[4], human beta-defensin (HBD-1)[9] and rhesus theta defensin 1.

b) DRP in animal venoms

Venoms from animals like snakes,[10] scorpions,[11] cone snails[12] or spiders[13]are rich reservoirs of bioactive molecules including many DRPs with highly diverse structures and functions. They are usually classified into different families according to their three-dimensional structure (fold type), length, cysteine distribution along the sequence, and disulfide connectivity.[14] A few representative examples are shown in figure 2. Many of these DRPs have proved to be potent and selective ligands targeting pharmacologically important receptors, including several from the central nervous system.[12,14]

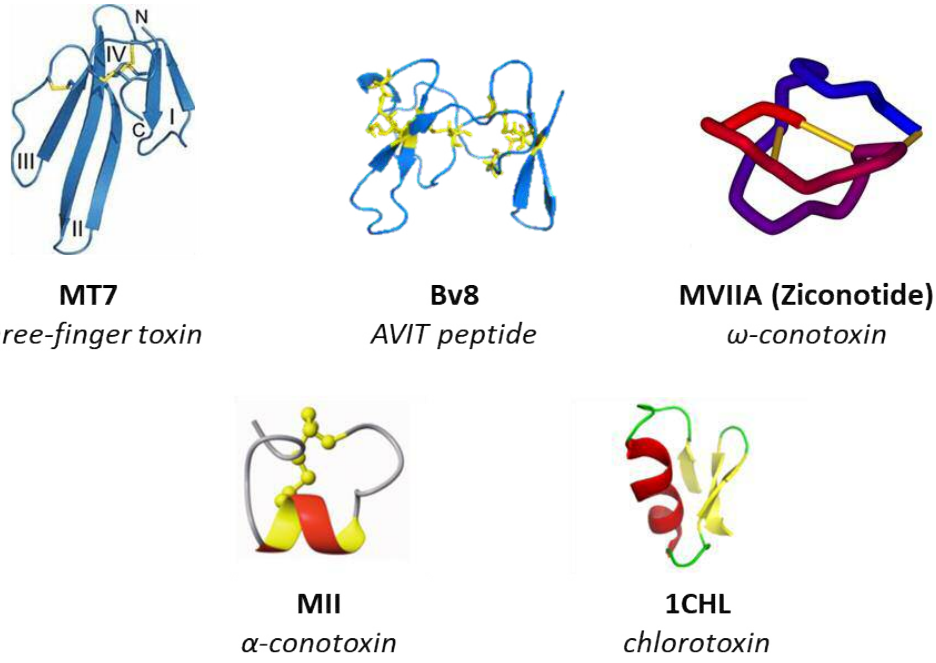


Figure 2. Three-dimensional structures of Muscarinic toxin 7 (MT7),[10] *Bombina variegata* protein 8 KDa (Bv8),[15] ω-Conotoxin MVIIA,[16] *α-conotoxin* MII and *chlorotoxin-1* (1CHL).[14,17] (Images extracted from ref 5,15,16,17 and Wikipedia)

c) Cyclotides

Cyclotides are the most abundant DRPs found in plants (Rubiaceae, Cucurbitaceae, Fabaceae or Violaceae).[18] They have the particularity to present a cyclic backbone with a highly conserved and well-defined topological motif: the cyclic cystine knot (CCK) which are responsible for the exceptional stability of these peptides (Figure 3).[19–21] They are involved in host defense and they show a wide range of interesting pharmaceutical activities such as antimicrobial, cytotoxic or uterotonic activities.[21,22] They are also interest for agronomical applications such as biopepticides.[21]

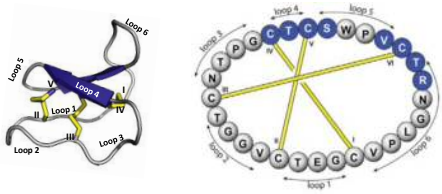


Figure 3. Structure of Kalata B1, the first isolated cyclotide. (Image extracted from ref 3)[3].

1.2. DRP as promising drugs candidates

In terms of size, DRPs are positioned in a chemical space between classical small molecules and large biologics such as monoclonal antibodies or other therapeutics proteins (Figure 4).[23,24] Consequently, DRPs exhibit the main advantages of both families but not their major

limitations. Hence they share the bioavailability, physicochemical and biological stability, and low immunogenicity of small molecules and the high selectivity and low toxicity of biologics. In addition, the high conformational stability of DRPs confer to them important properties such high selective, potency, thermal stability or better permeability.[25] Consequently they are considered as excellent drug candidates.

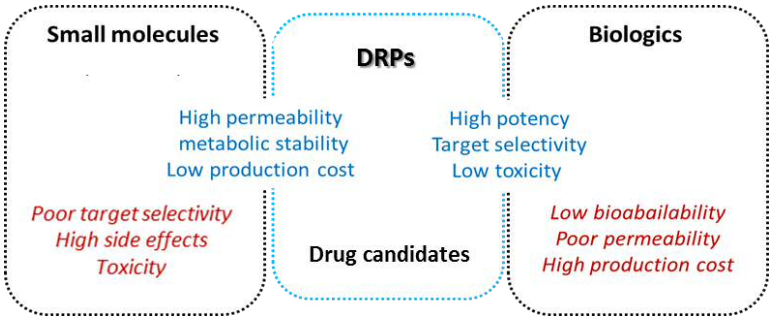


Figure 4. Druggable chemical space of constrained peptides.

One of the main limitations of peptide-based drugs is their vulnerability to proteolysis that can dramatically decrease their bioavailability.[26] In the case of the highly-constrained DRPs, the peptide bonds susceptible to proteolytic attack are not accessible to proteases. Hence, they are highly resistant compounds with long half-lives in serum.[3,24] In some cases, they can even resist to digestive enzymes and be orally administrated, a route which still represents a persistent challenge for peptide drugs.

Their important pharmacological properties have favored their inclusion in the drug market. Three synthetic DRPs have reached the market since 2000: Ziconotide, Linaclotide and Plecanatide (Table 2).[2]

The first one, marketed under the name of Prialt, is the synthetic form of the ω-*conotoxi*ns MVIIA which is a 25 amino acids peptide present in the venoms of the marine cone (*Conus magus*).[16,27] It shows a very potent analgesic activity and it is used in the treatment of the severe and chronic pain.

Linaclotide (commercialized as Linzess or Constella) and Plecanatide (commercialized as Trulance) are synthetic tetradecapeptide agonists of the guanylate cyclase 2C with a potent laxative activity.[28]

Table 2. Approval DRPs in 2014.[2]

Drug Name	Company	aa	Disulfide bonds	Indications	Approval year
Ziconotide	Elan	25	3	Severe and chronic pain	2004
Linaclotide	Ironwood	14	3	Chronic constipatiion	2012
Plecanatide	Synergy Pharmaceuticals	16	2	Chronic constipatiion	2017

Currently, they are many DRPs which are in clinical and preclinical trials for new pharmacological applications: treatment of inflammatory diseases (multiple or amyotrophic lateral sclerosis), cardiovascular diseases (heart failure or coagulation disorders), antifungal or antibacterial activities (Table 3).[2] These last ones are longer peptides and their production have been only possible through recombinant approaches.

Table 3. Example of DRPs in clinical and preclinical trials in 2014.[2]

Drug Name	Company	aa	Disulfide bonds	Indications	Phase
Chlorotoxin	TransMolecular	36	4	Cancer	I
Serelaxin	Novartis	50	3	Pre-eclampsia Heart failure	I III
Elafin	Proteo Biotech	57	4	Inflammation in cancer	II
α-Cobrotoxin	ReceptoPharm	62	4	HIV Herpes virus	I II
α-Cobratoxin	ReceptoPharm	71	5	Multiple sclerosis Herpes virus Demyelinating diseases	II II III

Naturally occurring DRPs can be used as scaffolds for the grafting of active peptide sequences, resulting in new chimeric structures with present all the advantages of DRPs such as their thermal stability or resistance to proteolysis while maintaining or enhancing the binding properties of the original peptide.[25,29] For example, Tam used the cyclotide kalata B1 as a scaffold to increase the stability and half-life in circulation of a bradykinin B_1 receptor antagonist, which is useful for the treatment of chronic pain and inflammatory pain (Figure 5).[30]

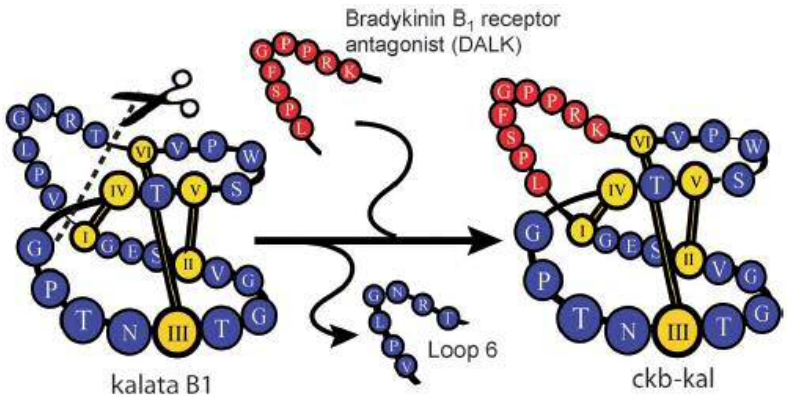


Figure 5. Cyclotide scaffold for the treatment of inflammatory pain. (Image extracted from ref 30).[30]

1.3. Chemical synthesis of DRPs

Another point that unquestionably has called the interest of the industry is the capacity of DRPs to be chemically prepared. As mentioned in chapter 1, the chemical synthesis of peptides offers in general some advantages in comparison with their recombinant production: in particular introduction of non-natural amino acids can greatly improve their pharmacological profiles and in particular their selectivity towards a receptor subtype.[2,3,29]

In the last two decades, the research of new methods for the chemical synthesis of these peptides has greatly increased, in order to facilitate their production and later manufacturing.

The specific point in the synthesis of DRPs is the regioselective formation of the disulfide bridges which are crucial for their biological activities.[2] The number of possible isomers exponentially increases with the number of disulfide bonds (figure 6). In some cases the isolation of the correct folded structure is complicated.

Another limitation of the oxidative folding is associated with the highly diluted conditions (10 µM) that are required to avoid the formation of intermolecular disulfide bonds. These conditions can result in low recoveries and can limit the expansion of the process to their large-scale manufacturing.

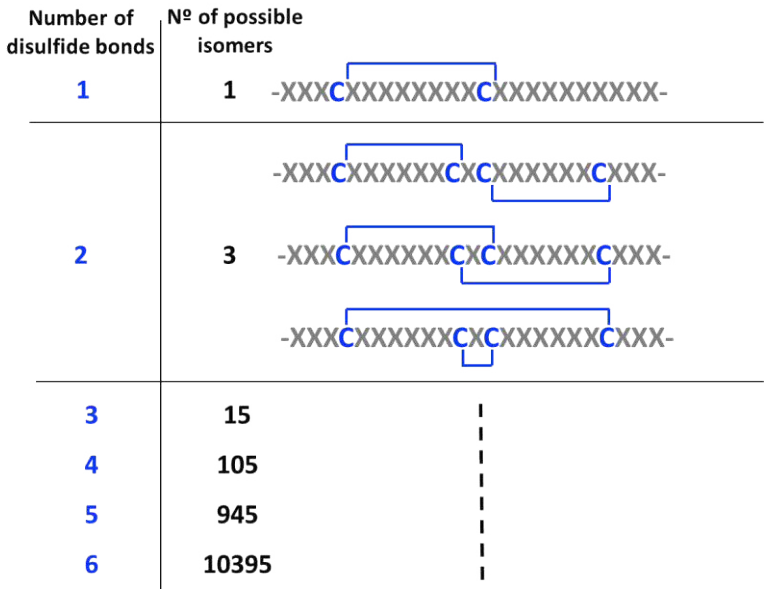


Figure 6. Number of possible isomers as a function of their number of cysteines.

2. Development of an N-terminal linker for the solid phase native chemical ligation (SP-NCL) of DRPs

Small DRPs up to 50 amino acids are, in general, accessible by standard SPPS. However, like other peptides, their HPLC purification can sometimes be highly challenging due to the presence of high amounts of acetylated truncated peptides. This problem is even exacerbated with DRPs, due to the large number of cysteine residues that are classically protected with a very bulky trityl group that slows down kinetics of coupling, leading to an increase in acetylated truncated peptides

To facilitate their purification, one alternative could be the application of the non-chromatographic "*catch*-and-*release*" purification strategy, which is based in the use of N-terminal linkers (described in chapter 1). Unfortunately, existing linkers are hardly compatible with the particular reactivity of unprotected cysteines. The main aim of this work has been to adapt the mentioned *catch*-and-*release* strategy to the synthesis and purification of these small-to-medium DRPs (Figure 7).

In addition, the synthesis of long DRPs peptides is complicated and requires the assembly of multiples unprotected fragments *via* successive ligations (usually NCLs) and intermediate purifications which are known to be costly, time-consuming and associated with low yields.

The capture of the peptide on the solid support could be exploited for the future expansion of the methodology to the synthesis of longer DRPs *via* successive solid phase NCLs in the N-to-C direction, which benefits from a "self-purification effect" (chapter 1) and avoids intermediate purifications.

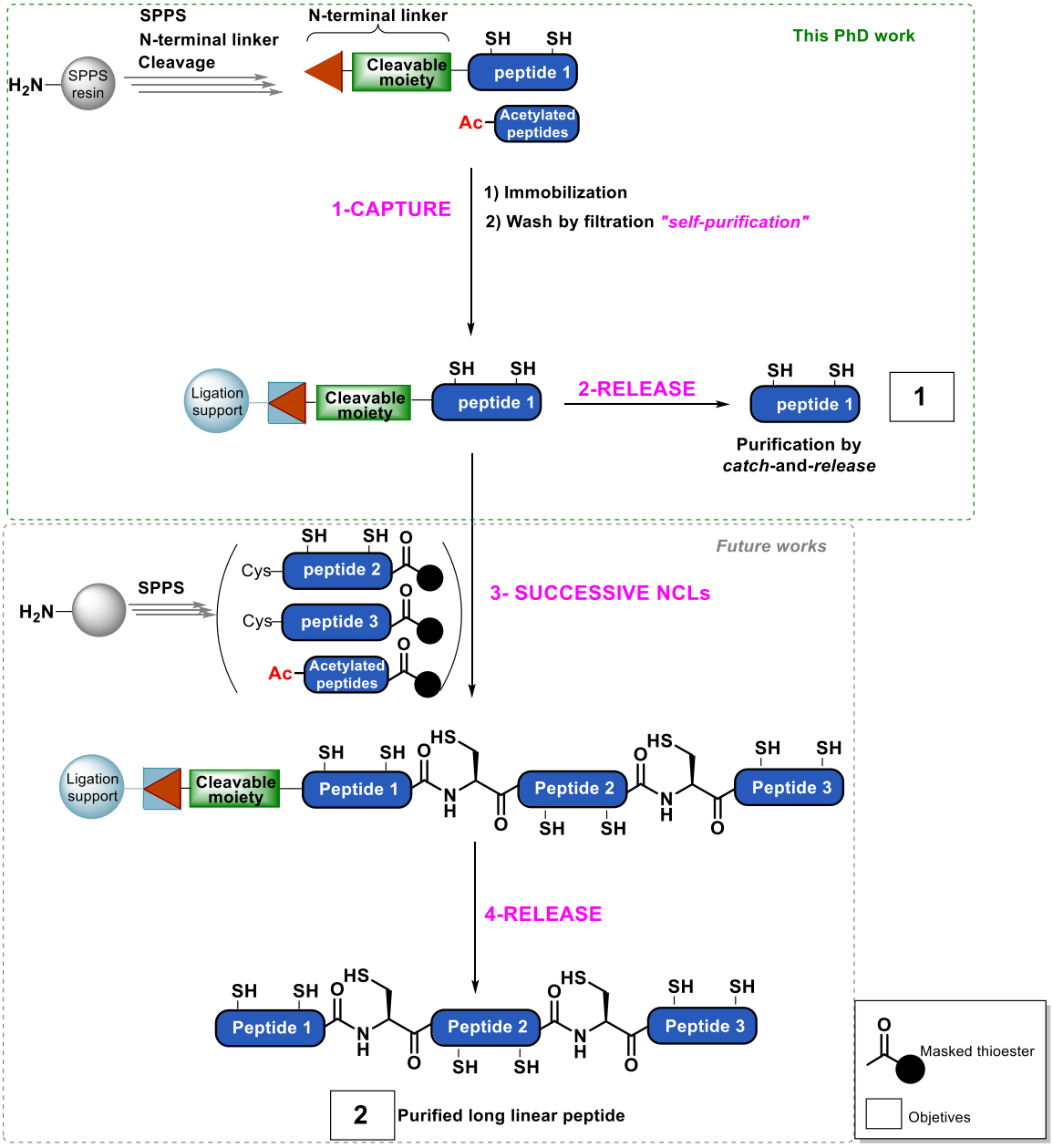


Figure 7 . *Catch*-and-*release* purification of cysteine-rich peptides (this thesis) and expansion to the synthesis of longer peptides though successive solid phase NCLs.[a]: in the case of successive NCLs, peptide 1 must be functionalized with a (masked) thioester.

2.1. Study of the existing capture and cleavable strategies compatibles with cysteines and with the successive NCLs

2.1.1. Capture strategies

Immobilization strategies which involve non-covalent reactions (chapter 1, part 2, page 35), could limit the range of treatments on the solid support (strong acids or bases, denaturing agents, etc.) and consequently the extension of the methodology to multiples NCLs.

For the same reason the capture through reversible covalent reactions, *via* oxime, oxazolidine, thiazolidine or thiol-disulfide interaction (chapter 1, part 2, page 39) would also limit the expansion of the methodology. Accordingly, we have excluded them.

Obviously reactions involving thiol functions, like cysteine-organomercurial interaction or thiol alkylation (chapter 1, part 2, page 39) cannot be applied to unprotected Cys-rich peptides and thus, they have been also excluded.

The polymerization was neither taken into account it has been only applied to protected peptides and consequently is hardly compatible with the reaction *via* NCL.

Finally, methods based on cycloaddition reactions such SPAAC, *i*EDDA, and CuAAC (chapter 1, part 2, page 41) are difficulty compatibles with these peptides. SPAAC and *i*EDDA involve strained alkynes or alkenes that can be attacked by the nucleophilic cysteines. CuAAC requires catalysis with Cu(I) salts which can be chelated with the unprotected cysteines.

In conclusion, none of the methods that have been proposed to date for the capture of the tagged peptides into the solid support are compatible with our goals, and we thus need to develop a specific linker.

2.1.2. Release strategies

a) Electrophilic cleavage: (chapter 1, part 2, page 43) with cyanogen bromide cannot be generalized to Met-containing peptides and consequently we decided not to take it into consideration.
b) Oxidative cleavage: (chapter 1, part 2, page 44) using sodium periodate is not compatible with free cysteines that get oxidized into sulfonic acids and therefore, it was also excluded.

c) Acidic cleavage: (chapter 1, part 2, page 45). All the methodology reported to date involving this kind of cleavage are only applicable to protected peptides and are consequently difficultly compatible with the NCL reaction.

d) Basic cleavage: (chapter 1, part 2, page 46). The β-elimination mechanism for the cleavage of the reported linkers, leads to the formation of electrophilic Michael acceptors (vinyl sulfone or dibenzofulvene) which can be attacked by the thiols.[31] In addition, strong basic conditions can cause a base-assisted desulfurization of cysteines.[32]

e) Nucleophilic cleavage: (chapter 1, part 2, page 48). There are three N-terminal linkers described in the literature which contains cleavable moieties susceptible to react under nucleophilic conditions (Figure 8).[33–35]

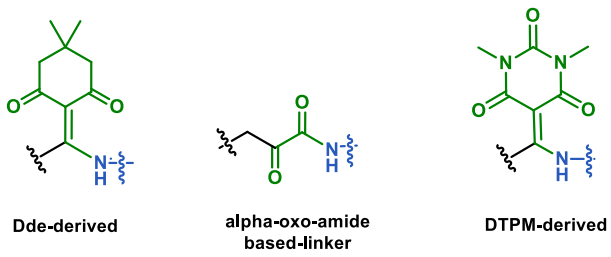


Figure 8. Existing linkers containing a cleavable moiety susceptible to a nucleophilic attack.[33–35]

As mentioned in chapter 1, Dde-derived linkers are unstable under aqueous conditions, and are therefore not applicable to solid supported NCLs.

Both α-oxo-amide and DTPM-based linkers, could in principle be used for our goals. However, the strategy for the cleavage of the first linker implies two step-reactions (transamination then nucleophilic cleavage). The DTPM-derived structures, reported by the team, seem to be the most adapted for our purpose.

2.2. Design of the Boc-Cys(Trt)-1-{6-[1,3-dimethyl-2,4,6 (1*H*,3*H*,5*H*) trioxo-pyrimidine -5-ylidene]hexyl}-OH linker

We have chosen the NCL reaction for the immobilization of peptides which is a chemoselective reaction completely compatible with the cysteine-rich peptides. The NCL occurs between a thioester moiety and a 2 amino thiol-containing moiety (such as a N-terminal cysteine) in water at neutral pH and results in the formation of a stable native amide bond.[36]

For our strategy we decided to incorporate the cysteine moiety into the DTPM linker and not the thioester which would be more difficult since a synthetically point of view.

The new N-terminal linker: H-Cys-1-{6-[1,3-dimethyl-2,4,6 (1*H*,3*H*,5*H*) trioxopyrimidine-5-ylidene]hexyl}-OH linker, hereafter referred as **Cys-Dtph-linker** and capture strategy are schematize in Figure 9. This new linker should be compatible with the synthesis of cysteine-rich peptides and should also be useful for further successive chemical ligations in solid phase.

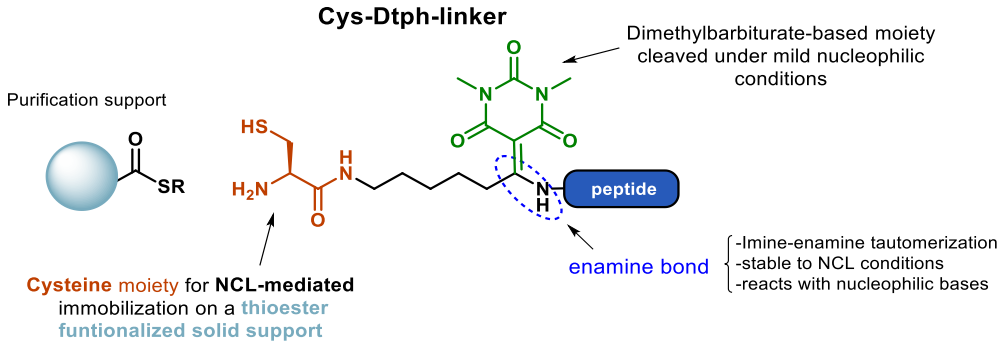


Figure 9. Structure of Cys-Dtph-OH linker and thioester functionalised solid support (the synthesis of the thioester solid support is explained in detail in chapter 3).

2.3. Synthesis of Boc-Cys(Trt)-Dtph-OH heterobifunctional linker

Enol precursor **3** was obtained through a three-step synthesis starting from commercially available methyl 6-aminohexanoate hydrochloride in a 64% overall yield (Figure 10). The first step consisted in a simple coupling between the amine group of the amino ester and the acid group of Boc-Cys(Trt)-OH in presence of DCC and HOBt, to lead to amide **1**. Saponification of the ester then formed the carboxylic acid **2**. The last step consisted in a *C*-acylation reaction between **2** and dimethyl barbituric acid mediated by EDCI and DMAP,[35] to successfully yield the enol **3.**

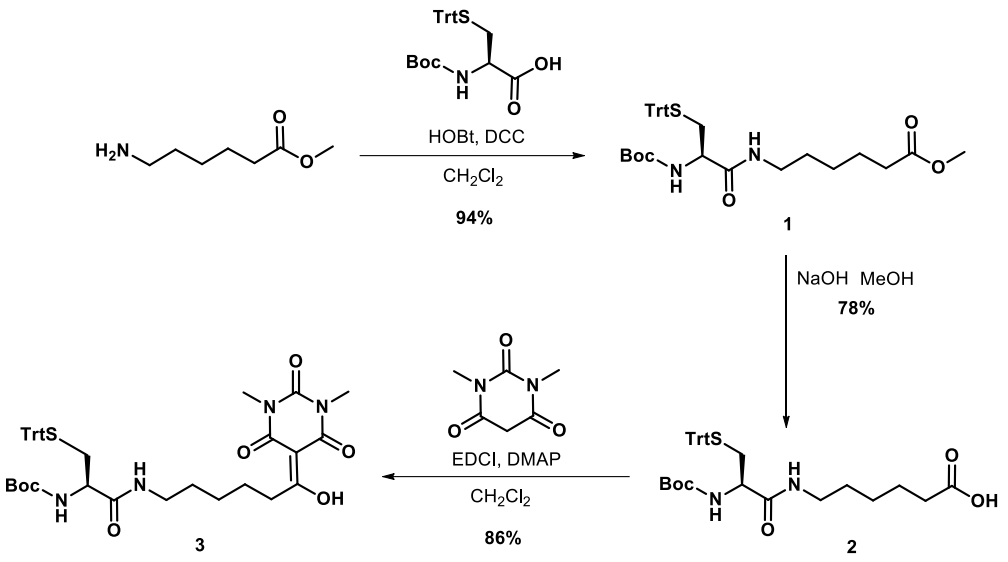


Figure 10. Synthesis of the linker precursor 3.

The *C*-acylation presumably goes through the reaction of a very reactive acyl pyridinium intermediate with the enolate formed by deprotonation of the dimethyl barbituric acid (Figure 11), to generate the ketone. Keto-enol equilibrium conducts to the formation of the enol **3**.

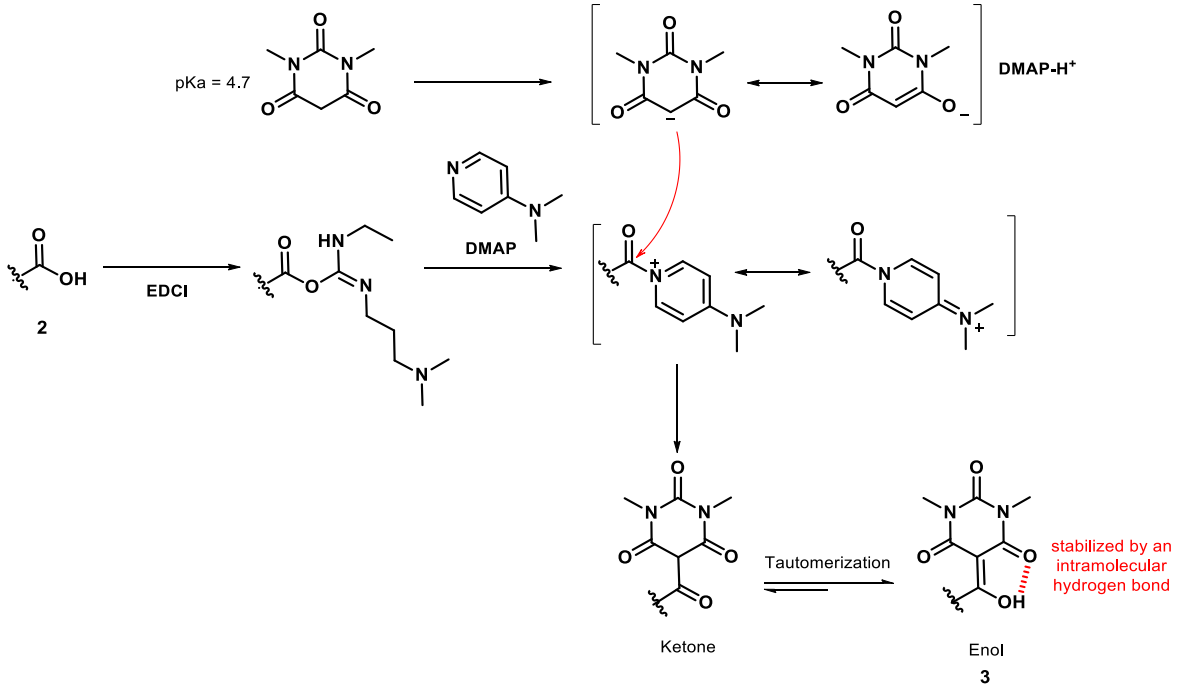


Figure 11. Detailed C-acylation reaction.

Enol tautomer is stabilized by conjugation of the alkene with the amide carbonyl groups, and probably by an intramolecular hydrogen bond as observed by El-Faham and collaborators in the case of 1,3-dimethyl-5-propionylpyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione (Figure 12).[37]

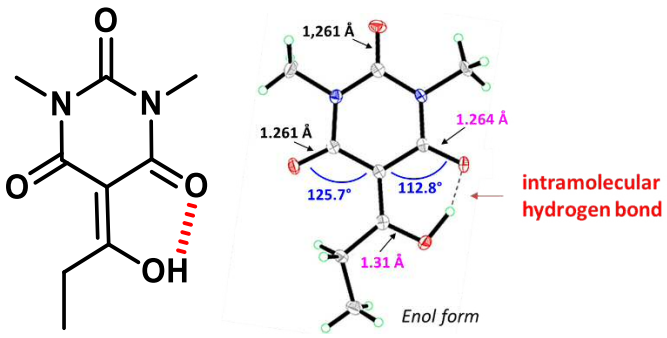


Figure 12. Chemical and crystal structure of 1,3-dimethyl-5-propionylpyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione (with selected bond lengths (Å) and angles (◦), reveals the presence of an intramolecular hydrogen bond which stabilizes the enol tautomer. X-ray structure adapted from ref 37.[37]

2.4. Synthesis of the Dtph-protected amino acids

Derivatization of an amine with reagent **3** can be achieved by simple incubation, not needing any coupling reagent. However, this reaction typically requires heating at high temperatures which are not compatible with a solid-supported peptide. Consequently, we decided to introduce the linker in solution as the amine protecting group of the N-terminal amino acid, and coupled then to the solid supported peptide under standard conditions.

We selected four AAs which do not contains any reactive function in their side chains, with different steric hindrance (Gly, Val, Leu, Ala) and one example of an AA with a reactive side chain (Ser) to generate a series of five *N*-protected N-term AAs, **4a**-**4e.**

The coupling was performed with an excess of amino acid (Gly, Val, Leu, Ala, and Ser(O*t*Bu)) in MeOH at 150°C in the presence of diisopropylethylamine (iPr$_2$EtN) under microwave irradiation for 10 min (Figure 13).

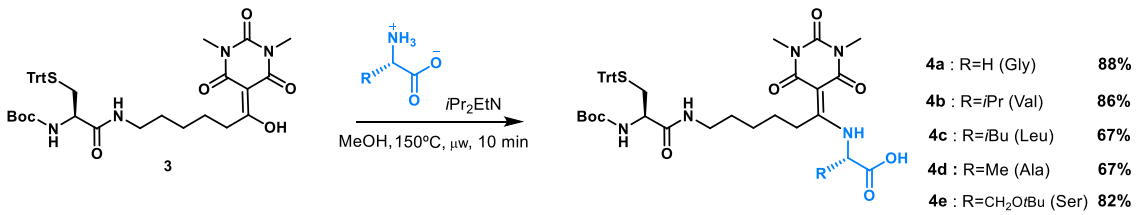


Figure 13. Synthesis of functionalized amino acid Boc-Cys(Trt)-Dtph-linkers. Note that for the application of our strategy, all peptides in this work have been synthesized without the last N-terminal amino acid.

Note that the conditions previously established in the laboratory for the synthesis of N_3-Dtpp-derivatized amino acids involved heating at reflux for 24 hours.[35] In comparison, MW heating dramatically reduces the reaction time while maintaining high yields.

The mechanism of that reaction likely involves a keto-enol equilibrium followed by the attack of the ketone by the nucleophilic amine of the amino acid to generate a carbinol intermediate, which conducts to the imine by loss of a water molecule. Tautomerization conducts to the formation of the enamine which is stabilized, similarly to enol **3**, by an intramolecular hydrogen bond and conjugation of the alkene with the two amide carbonyls (Figure 14).

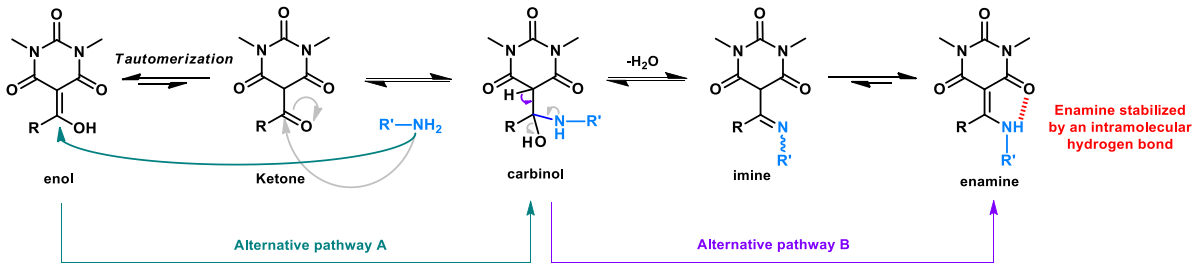


Figure 14. Mechanism of enamine formation.

Note that a direct transformation of the carbinol into the enamine could also occur through a retro-Michael reaction (Pathway B, Figure 14). Another alternative pathway could be proposed in which the amine reacts with the enol by a Michael addition to directly form the carbinol intermediate (Pathway A, Figure 14).

UV-VIS characterization of the Dtph-derivatized AA.

We noticed that our linker has a characteristic UV absorption profile, with a maximum at 303 nm (Figure 15) which could be very useful for the detection of the linker-peptide during HPLC. In addition, the specific absorption at this wavelength could be also used for the quantification of the amount of the Dtph-containing peptide at 303 nm during this work.

For that, the values of the molar extinction coefficient (ε) of the linker were measured. Note that we have supposed that there were no differences in the ε between Boc-Cys(Trt)-Dtph-Xaa-OH and H-Cys-Dtph-peptide.

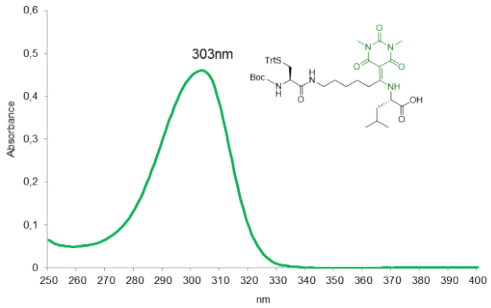


Figure 15. UV-profile of 4b.

Given that UV spectroscopy at 280 nm is the most common method for the quantification of peptides and proteins in solution (ε^{280} (Tyr) = 1490 M^{-1} cm^{-1}, and ε^{280} (Trp) = 5500 M^{-1} cm^{-1} in water)[38] we have also measured the molar extinction coefficient at this wavelength in order to take into account the contribution of the linker (Table 4).

Finally, the molar extinction coefficient was also measured at 320 nm which is specific for our linker (table 4).

Table 4 Molar extinction coefficient of 4c in a H_2O/MeCN/TFA (1:1:0.001) solution.

Wavelength (nm)	ε (M^{-1} cm^{-1})
280	1780
303 (λ max)	5980
320	1230

2.5. Preparation of a Dtph-free-cysteine model peptide though SPPS

A model peptide derived from human mucin 1 (MUC1) was chosen for the evaluation of the stability of our linker and optimization of its cleavage conditions. MUC1 is a transmembrane glycoprotein which is overexpressed in many cancers, and possesses a great importance for the development of cancer immunotherapy.[39] Its extracellular region is mainly constituted by a densely glycosylated tandem repeat sequence of 20 amino acids (VTSAPDTRPAPGSTAPPAHG), which is repeated between 50 and 150 times.[35] We chose the non-glycosylated form due to its easy synthesis and its remarkable solubility in water.

2.6. Synthesis of solid supported peptide 5

We synthesized a 20-mer MUC1-derived peptide sequence: H-**WTSAPDTRPAPGSTAPPAHG**, where the N-terminal valine was replaced by a tryptophan, to facilitate HPLC-detection and quantification by UV spectrophotometry ($\varepsilon^{280} = 5500\ M^{-1}\ cm^{-1}$).

Supported peptide **5** was obtained by automated Fmoc-SPPS starting from commercially available aminomethyl Tentagel resin using a Rink amide linker (methoxycarbonylamino]-2,4-dimethoxy benzyl}-phenoxyacetic acid), in order to generate a C-terminal amide after TFA-mediated cleavage of the peptidyl resin. The elongation yield (determined from the UV titration of the first and last Fmoc group deprotection) was 79%. An aliquot of the peptidyl resin **5** was cleaved from the resin to characterize the corresponding peptide **5´** (Figure 16).

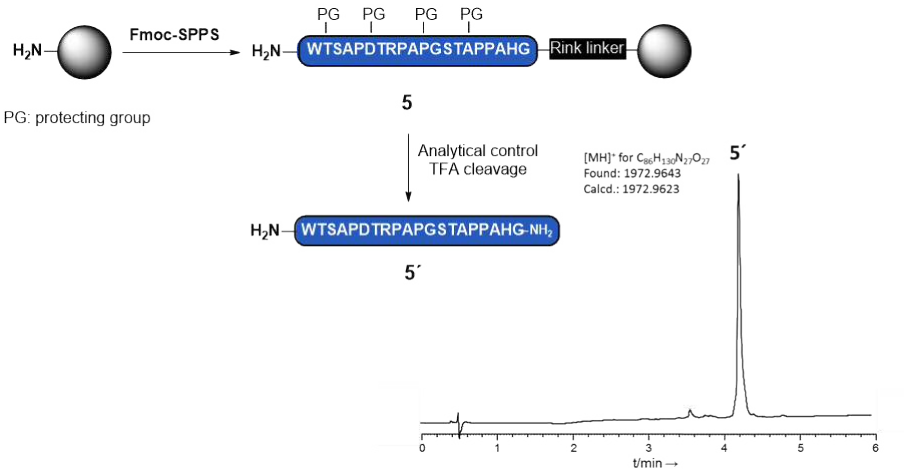


Figure 16. Fmoc SPPS synthesis of peptidyl-resin 5 and HPLC chromatogram of the crude 5´.

2.7. Synthesis of Dtph-containing peptides 6a-c

Boc-Cys(Trt)-Dtph amino acid **4a-c** were then coupled onto the N-terminal amino acid of the supported peptide **5**. In order to limit the number of equivalents of **4a-c**, we used PyAOP, a more powerful coupling agent than HCTU (routinely used for automated SPPS). The resulting peptides were cleaved and purified by RP-HPLC to give Cys-Dtph-peptides **6a-c** in high yields (Figure 17).

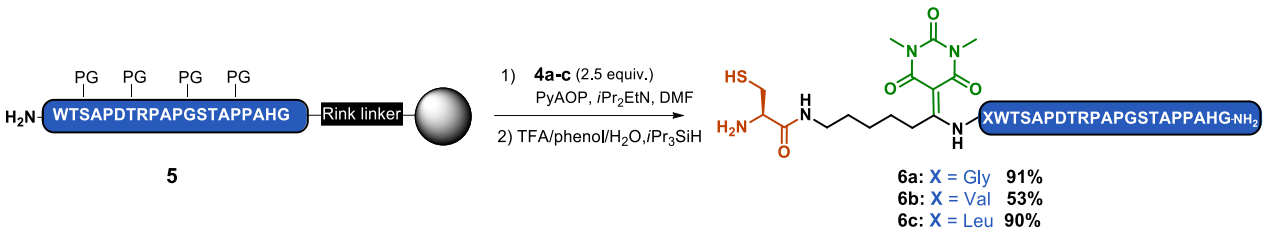


Figure 17. Synthesis of compound 6a-c.

Overall yields including coupling of the Dtph-derivatized AA, cleavage, and purification, were determined by UV titration of **6a-c** at 303 nm (ε^{303} (Dtph) = 5980 M^{-1} cm^{-1}) in a H_2O/MeCN/TFA (1:1:0.001) solution.

3. Preliminary studies of stability and cleavage kinetics of the Dtph Linker

3.1. Masking of the N-terminal cysteine

We have decided not to study the stability and cleavage kinetics of the Dtph linker using **6a-c**, which contains a reactive N-terminal cysteine. Instead of that, three Cys-Dtph-containing peptides **8a-c** have been synthesized through NCL between **6a-c** and thioester **16** (see chapter 3).

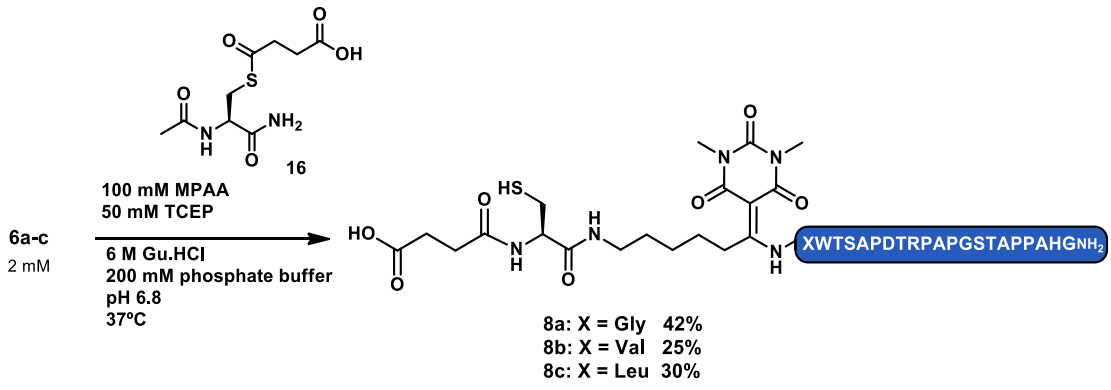


Figure 18. NCL between peptides 6a-c and thioester 16.

The NCL reaction was conducted at 2 mM peptide concentration, under classical NCL conditions in the presence of 4-mercaptophenylacetic acid (MPAA) and *tris*-carboxyethylphosphine (TCEP). An excess of thioester **16** was used (2.5 equivalents). The progress of the reaction was monitored by analytical HPLC. For peptides **6b** (N-terminal Val) and **6c** (Leu) we observed the apparition of one major peak, corresponding with the expected

peptides **8b** and **8c**, respectively (Figure 19). The reaction was completed over 1 hour. Prior to purification, the NCL mixtures were washed with diethyl ether in order to extract the MPAA. The resulting peptides were purified by RP-HPLC and the yields (corresponding to NCL and purification) were quantified by UV spectrometry a 303 nm: 25% for **8b** and 30% for **8c**.

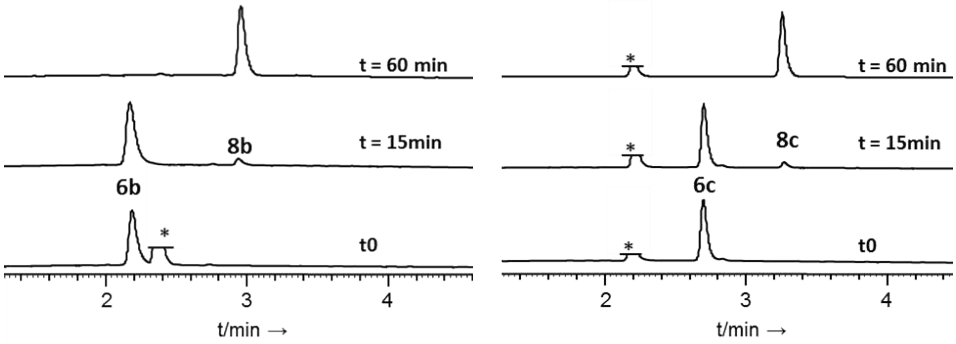


Figure 19. Analytical HPLC profiles of NCL for peptides 8b and 8c.*MPAA traces not completely remove during Et_2O extraction.

Unexpectedly, for peptide **6a** (N-terminal Gly), after one hour we detected another minor compound that was attributed to thioester **7** (Figure 20).

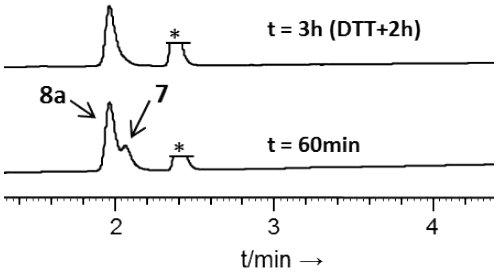


Figure 20. Analytical HPLC profil of NCL for peptides 8a. *MPAA traces not completely remove during Et_2O extraction.

Compound **7** could be transformed in **8a** by treatment with an excess of DL-dithiothreitol (DTT) (Figure 21) following conditions used in the literature for the hydrolysis of thioesters.[40–42] HPLC purification furnished **8a** in a 42% yield.

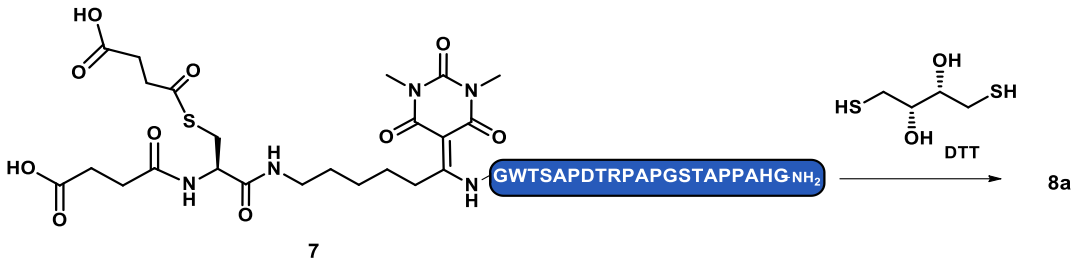


Figure 21. Supposed structure of 7.

A possible mechanism for this reaction may include a trans-thioesterification followed by a hydrolysis. This mechanism has been previously proposed for the hydrolysis of thioesters mediated by the 2-mercaptoethanol (Figure 22.[40–42]

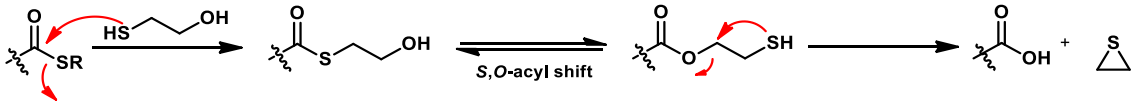


Figure 22. Mechanism of beta-mercaptoethanol-mediated hydrolysis of thioester proposed by Kent et al.[40,41]

For the three reactions, the yields were lower than expected and we believe that it may be due to a partial peptide precipitation during the washes with diethyl ether.

3.2. Study of the stability in solution of 8a-c at different pHs

In order to evaluate the stability of the linker and the robustness of the enamine bond between the linker and the peptide, peptides **8a**-**c** were subjected to a set of conditions with a pH range of 2 to 12, which are representative of conditions commonly used in peptide chemistry.[35] Three different buffers were used: 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), *tris*(hydroxymethyl)aminomethane (TRIS) and *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) (table 5).

Table 5. Useful pH ranges of selected biological buffers (25 °C, 0.1 M). Source: Sigma Aldrich home page.

Buffer	Useful pH range
HEPES	6.8 - 8.2
TRIS	7.5 - 9
CAPS	9.7 – 11.1

The solutions were stored at room temperature for 24 hours. The amounts of starting material and free peptide resulting from linker cleavage were measured by integration of the HPLC

peaks at λ = 214 nm (neglecting the differences in molar absorption). The results of these experiments are summarized in table 6.

Table 6. intact peptide linker after 24 hours of incubation.

		% intact linker		
Entry	**Conditions[a]**	**8a(Gly)**	**8b(Val)**	**8c(Leu)**
1	0.1% TFA in H_2O (pH 2)	100	100	100
2	100 mM HEPES buffer (pH 7)	100	100	100
3	100 mM HEPES buffer (pH 8)	100	100	99
4	1 M TRIS buffer (pH 8.5)	100	100	99
5	100 mM CAPS buffer (pH 9)	99	70	64
6	100 mM CAPS buffer (pH 12)	0	0	0

[a]: all reaction were performed at 0.1 mM peptide concentration.

The linker is completely stable from pH 2 to pH 8.5 (entries 1-4).

Unexpectedly at pH 9, the hydrolysis of **8a** which contains an N-terminal Gly was significantly slower than for **8b** and **8c** (entry 5).

At pH 12, quantitative hydrolysis of the linker was observed (entry 6).

Note that the enamine is stable in presence of a very large excess of an amine, as demonstrated by incubation with 1 M TRIS (entry 4).

3.3. Study of Dtph cleavage kinetics

Dimethylbarbiturate-derived enamines are known to be cleavable through treatment with an excess of hydrazine or hydroxylamine. These reagents are typical α-nucleophiles where a nucleophilic center is neighbored by a α-atom which bears a lone pairs of electrons.[35,43] The α-atom increases the nucleophily of the other atom through the so-called “alpha effect”.[44]

The cleavage is likely to occur through a 1,4-addition of the nucleophile (hydrazine or hydroxylamine) to the α,β-unsatured carbonyl moiety (Figure 23, pathway A). The formed *N*,*N*-ketal intermediate can undergo a retro-Michael addition leading to the liberation of either the nucleophile (reverse reaction) or the amine. In the presence of a large excess of nucleophile, the equilibrium can be displaced to exclusively lead to the liberation of the amine. Note that an alternative mechanism involving addition to the imine form could be also envisioned (figure 23, pathway B).

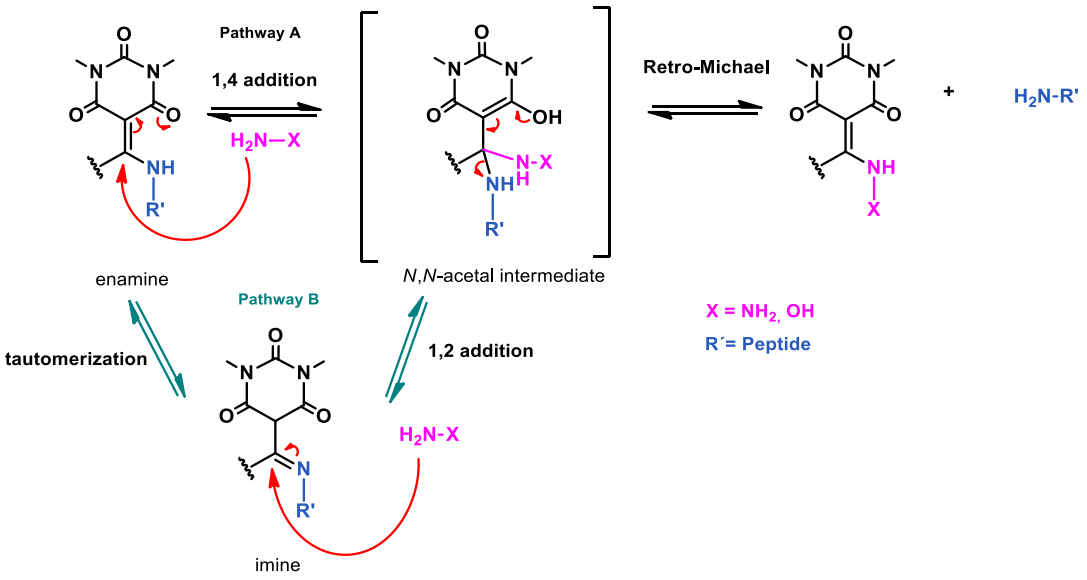


Figure 23. Proposed mechanism for the cleavage of dimethylbarbiturate-based enamine.

For this study, peptides **8a**-**c** (0.2 mM) were subjected to a set of aqueous solutions at different pH, using two different nucleophiles: hydroxylamine or hydrazine, at a 1 M concentration. TCEP was added to keep reduced the thiol group of the cysteine during the cleavage (Figure 24).

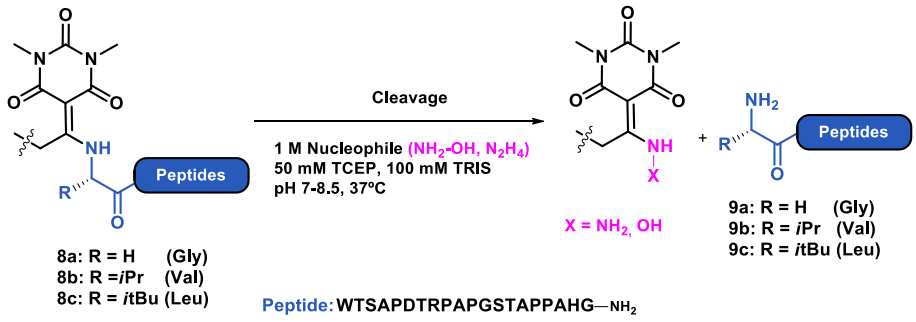


Figure 24. Cleavage of compound 8a-c.

The cleavage of the linker was followed by analytical HPLC similarly that for the stability studies.

The rate constant and half-life ($t_{1/2}$) were calculated by exponential fitting (Origin 9 software) using a pseudo-first order equation: , where A is the concentration of model linker-peptides (**8a**, **8b** and **8c**, at time t), A_0 is the concentration at time 0, k' is the pseudo-first order apparent rate constant (s^{-1}) and t is the time (minutes).

We considered a pseudo-first order reaction due to the large excess (5000 equiv.) of nucleophile, compared to the peptide.

Firstly, we evaluated the cleavage of compounds **8a-c**, under conditions that we have successfully used for the N_3-Dtph linker: hydroxylamine at pH 7.[35] The results are showed in table 7.

Table 7. Kinetic rate constant and half-life ($t_{1/2}$) of compound 8a-c upon 1 M NH_2-OH at pH 7 (See figure 27-29 in experimental section).

Entry	Compound	N-ter AA	$t_{1/2}$ (min)	k' (10^{-3} s^{-1})	Relative rate
1	**8a**	Gly	7.3	3.22	100
2	**8b**	Val	182	0.11	3.4
3	**8c**	Leu	53	0.27	8.3

Unexpectedly, we observed that the nature of the N-terminal AA has a strong influence on the cleavage kinetics. This aspect was not reported in any previous work; however it is of paramount importance for further applications of our linker. The bulkiness of the side chain group may limit the access of the nucleophile to the enamine and thus slow down the cleavage (Figure 25).

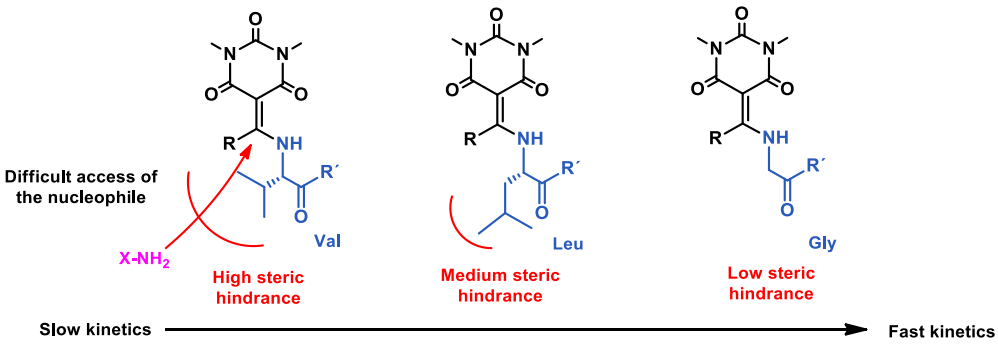


Figure 25. Influence of amino acids in the cleavage kinetics.

We can conclude that the cleavage conditions should be chosen in function of the N-terminal amino acid. Use of hydroxylamine at pH 7 is valid for glycine and leucine, but the reaction is too slow for the β-branched valine.

In order to accelerate the speed of the reaction, we first tried to increase the pH to 8.5 (See table 4 in experimental section), however no significant increment of the reaction rate was observed.

We next explored the use of hydrazine which is known to be a better nucleophile than hydroxylamine (Table 8).

Table 8. Rate constant and half-life ($t_{1/2}$) of compound 8a-c upon 1 M NH_2-NH_2 at pH 8.5 (See figure 27-29 in experimental section).

Entry	Compound	N-ter AA	$t_{1/2}$ (min)	k' (10^{-3} s^{-1})	Relative rate
1	**8a**	Gly	0.3	40	100
2	**8b**	Val	7.7	1.5	3.8
3	**8c**	Leu	1.7	7.0	17.4

At pH 8.5 the reaction was much faster than with hydroxylamine (around ten- to twelve-fold increase in rate constant), as we expected due to higher nucleophilicity (Table 8, figure 26). In all cases, cleavage was quantitative after 1 h of incubation and we selected these last conditions for all the applications of our linker described in this thesis.

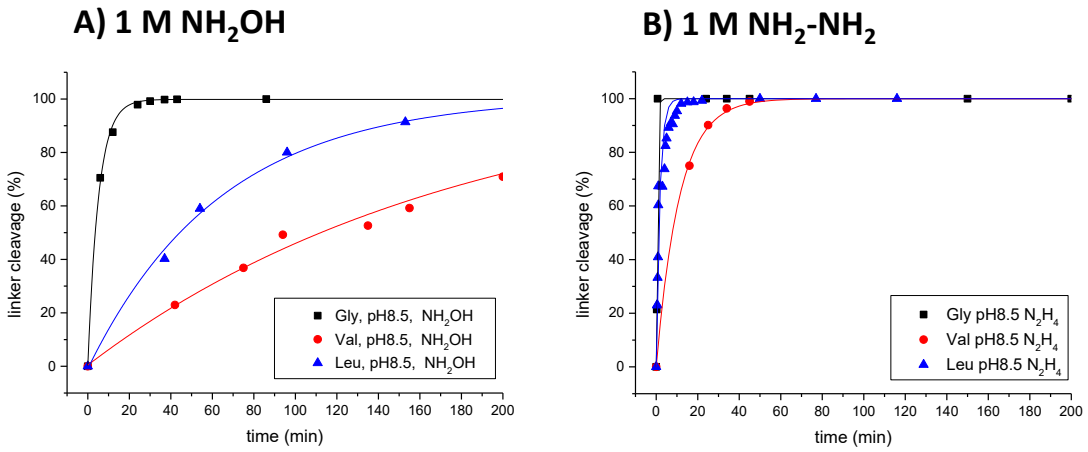


Figure 26. A) Representation of the reaction kinetics of 8a-c under 1 M NH_2OH at pH 8.5. B) Representation of the kinetics of the reaction of 8a-c under 1 M H_2N-NH_2 at pH 8.5.

We have also studied for **8b** the cleavage conditions at pH 7 with hydrazine (Table 9).

The large differences in the kinetic rates at pH 7 and 8.5 can be explained by the pKa of the conjugate acid of hydrazine (pKa = 8.1). At pH 8.5, hydrazine is mostly in its free-base reactive form (N_2H_4) whereas it is mostly in its non-nucleophilic conjugated acid form ($[N_2H_5]^+$) at pH 7.

Table 9. Rate constant and half-life($t_{1/2}$) of compound 8b upon 1 M NH_2-NH_2 at pH 8.5 and 7 (See figure 28 in experimental section).

Entry	N-ter AA	pH	$t_{1/2}$ (min)	k' (10^{-3} s^{-1})	Relative rate
1	Val	8.5	7.7	1.5	100
2	Val	7	1034	0.01	0.7

Interestingly, we noted that at pH 7, the cleavage is faster with hydroxylamine than with hydrazine (Figure 27).This result can be explained in regards to the pKa of the conjugate acid of hydroxylamine (5.95): hydroxylamine is mostly in its non-protonated nucleophilic form at both pH.

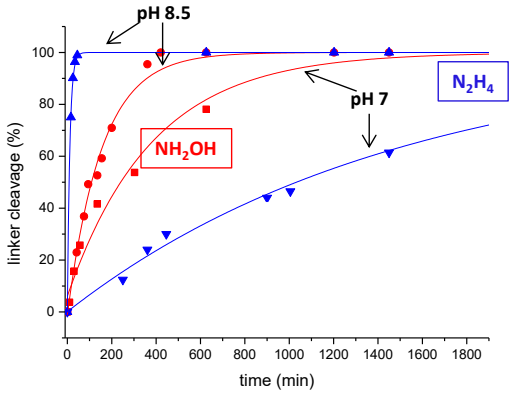


Figure 27. Reaction kinetics of compound 8b cleavage at different pH. Influence of the pka on the nucleophilic strength.

4. Conclusions

We conceived and synthesized a new N-terminal linker for application of the *catch*-and *release* purification and successive NCLs of unprotected cysteine-rich peptides. This new linker was selectively introduced in three model peptides, with different N-terminal amino acids. We demonstrated that the Cys-Dtph linker is stable under a large range of conditions and optimized suitable conditions for its cleavage. In addition, the capability of our linker to react by NCL and its stability under the ligation conditions were also demonstrated. Altogether these results suggested that Cys-Dtph-linker could be an excellent linker for both capture and release purification of medium-sized cysteine-rich peptides and the assembly of multiple fragments *via* successive solid-supported NCL in the N-to-C direction to produce longer ones.

5. References chapter 2

(1) Vincent, L.; Ryan, J. T.; Paul, F. A. Cysteine-Rich Mini-Proteins in Human Biology. *Curr. Top. Med. Chem.* **2012**, *12* (14), 1514.

(2) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014**, *114* (2), 901.

(3) Northfield, S. E.; Wang, C. K.; Schroeder, C. I.; Durek, T.; Kan, M. W.; Swedberg, J. E.; Craik, D. J. Disulfide-Rich Macrocyclic Peptides as Templates in Drug Design. *Eur. J. Org. Chem.* **2014**, *77*, 248.

(4) Ganz, T. Defensins: Antimicrobial Peptides of Innate Immunity. *Nat. Rev. Immunol.* **2003**, *3* (9), 710.

(5) Selsted, M. E.; Ouellette, A. J. Mammalian Defensins in the Antimicrobial Immune Response. *Nat. Immunol.* **2005**, *6* (6), 551.

(6) Shafee, T. M. A.; Lay, F. T.; Phan, T. K.; Anderson, M. A.; Hulett, M. D. Convergent Evolution of Defensin Sequence, Structure and Function. *Cell. Mol. Life Sci.* **2016**, *74* (4), 663.

(7) Shafee, T. M. A.; Lay, F. T.; Hulett, M. D.; Anderson, M. A. The Defensins Consist of Two Independent, Convergent Protein Superfamilies. *Mol. Biol. Evol.* **2016**, *33* (9), 2345.

(8) Derache, C.; Meudal, H.; Aucagne, V.; Mark, K. J.; Cadène, M.; Delmas, A. F.; Lalmanach, A.-C.; Landon, C. Initial Insights into Structure-Activity Relationships of Avian Defensins. *J. Biol. Chem.* **2012**, *287* (10), 7746.

(9) Sugiarto, H.; Yu, P. L. Avian Antimicrobial Peptides: The Defense Role of β-Defensins. *Biochem. Biophys. Res. Commun.* **2004**, *323* (3), 721.

(10) Reeks, T. A.; Fry, B. G.; Alewood, P. F. Privileged Frameworks from Snake Venom. *Cell. Mol. Life Sci.* **2015**, *72* (10), 1939.

(11) Martinez, G.; Hograindleur, J.-P.; Voisin, S.; Abi Nahed, R.; Abd El Aziz, T. M.; Escoffier, J.; Bessonnat, J.; Fovet, C.-M.; De Waard, M.; Hennebicq, S.; et al. Spermaurin, an La1-like Peptide from the Venom of the Scorpion Scorpio Maurus Palmatus , Improves Sperm Motility and Fertilization in Different Mammalian Species. *Mol. Hum. Reprod.* **2016**, *23* (2), 116.

(12) Terlau, H.; Olivera, B. M. Conus Venoms: A Rich Source of Novel Ion Channel-Targeted Peptides. *Physiol. Rev.* **2004**, *84* (1), 41.

(13) King, G. F. Venoms as a Platform for Human Drugs: Translating Toxins into Therapeutics. *Expert Opin. Biol. Ther.* **2011**, *11* (11), 1469.

(14) Lewis, R. J.; Garcia, M. L. Therapeutic Potential of Venom Peptides. *Nat. Rev. Drug Discov.* **2003**, *2* (10), 790.

(15) Mollay, C.; Wechselberger, C.; Mignogna, G.; Negri, L.; Melchiorri, P.; Donatell, B.; Kreil, G. Bv8, a Small Protein from Frog Skin and Its Homologue from Snake Venom Induce

Hyperalgesia in Rats. *Eur. J. Pharmacol.* **1999**, *374* (2), 189.

(16) Schmidtko, A.; Lötsch, J.; Freynhagen, R.; Geisslinger, G.; Spafford, J.; Zamponi, G.; Group, for the Z. 98-022 S. Ziconotide for Treatment of Severe Chronic Pain. *Lancet (London, England)* **2010**, *375* (9725), 1569.

(17) Craik, D. J.; Adams, D. J. Chemical Modification of Conotoxins to Improve Stability and Activity. *ACS Chem. Biol.* **2007**, *2* (7), 457.

(18) Conibear, A. C.; Craik, D. J. Chemical Synthesis of Naturally-Occurring Cyclic Mini-Proteins from Plants and Animals. *Isr. J. Chem.* **2011**, *51* (8-9), 908.

(19) Kourra, C. M. B. K.; Cramer, N.; Sato, A. K.; Viswanathan, M.; Kent, R. B.; Wood, C. R.; Goodwin, D.; Simerska, P.; Toth, I.; Vlieghe, P.; et al. Converting Disulfide Bridges in Native Peptides to Stable Methylene Thioacetals. *Chem. Sci.* **2016**, *7* (12), 7007.

(20) Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. Plant Cyclotides: A Unique Family of Cyclic and Knotted Proteins That Defines the Cyclic Cystine Knot Structural Motif. *J. Mol. Biol.* **1999**, *294* (5), 1327.

(21) Gruber, C. W.; Čemažar, M. M.; Anderson, M. A.; Craik, D. J. Insecticidal Plant Cyclotides and Related Cystine Knot Toxins. *Toxicon.* **2007**, *49* (4), 561.

(22) Ackerman, S. E.; Currier, N. V; Bergen, J. M.; Cochran, J. R. Cystine-Knot Peptides: Emerging Tools for Cancer Imaging and Therapy. *Expert Rev. Proteomics.* **2014**, *11* (5), 561.

(23) Terrett, N. Drugs in Middle Space. *Med. Chem. Comm.* **2013**, *4* (3), 474.

(24) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2015**, *20* (1), 122.

(25) Barkan, D. T.; Cheng, X.-L.; Celino, H.; Tran, T. T.; Bhandari, A.; Craik, C. S.; Sali, A.; Smythe, M. L. Clustering of Disulfide-Rich Peptides Provides Scaffolds for Hit Discovery by Phage Display: Application to Interleukin-23. *BMC Bioinformatics* **2016**, *17* (1), 481.

(26) Postma, T. M.; Albericio, F. Disulfide Formation Strategies in Peptide Synthesis. *European J. Org. Chem.* **2014**, *2014* (17), 3519.

(27) Prommer, E. Ziconotide: A New Option for Refractory Pain. *Drugs of Today* **2006**, *42* (6), 369.

(28) a) Lacy, B. E.; Levenick, J. M.; Crowell, M. D. Linaclotide: A Novel Therapy for Chronic Constipation and Constipation-Predominant Irritable Bowel Syndrome. *Gastroenterol. Hepatol. (N. Y).* **2012**, *8* (10), 653. b) Miner Jr, P. B.; Koltun, W. D.; Wiener, G. J.; De La Portilla, M.; Prieto, B.; Shailubhai, K.; Beth Layton, M.; Barrow, L.; Magnus, L.; Griffin, P. H. A Randomized Phase III Clinical Trial of Plecanatide, a Uroguanylin Analog, in Patients with Chronic Idiopathic Constipation. *Am. J. Gastroenterelol.* **2017**, *112*, 613.

(29) Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. The Future of Peptide-Based Drugs. *Chem. Biol. Drug. Des.* **2013**, *81* (1), 136.

(30) Wong, C. T. T.; Rowlands, D. K.; Wong, C.-H.; Lo, T. W. C.; Nguyen, G. K. T.; Li, H.-Y.; Tam, J. P. Orally Active Peptidic Bradykinin B1 Receptor Antagonists Engineered from a Cyclotide Scaffold for Inflammatory Pain Treatment. *Angew. Chem. Int. Ed.* **2012**, *51* (23), 5620.

(31) Canne, L.; Kent, S. B. H.; Reyna, S. Solid Phase Native Chemical Ligation of Unprotected or N-Terminal Cysteine Protected Peptides in Aqueous Solution. U. S. Patent. 6,326,468. Dec. 4, 2001, 1998.

(32) Galande, A. K.; Trent, J. O.; Spatola, A. F. Understanding Base-Assisted Desulfurization Using a Variety of Disulfide-Bridged Peptides. *Biopolymers* **2003**, *71* (5), 534.

(33) Hara, T.; Tainosho, A.; Nakamura, K.; Sato, T.; Kawakami, T.; Aimoto, S. Peptide Purification by Affinity Chromatography Based on Alpha-Ketoacyl Group Chemistry. *J. Pept. Sci.* **2009**, *15* (5), 369.

(34) Kellam, B.; Chan, W. C.; Chhabra, S. R.; Bycroft, B. W. Transient Affinity Tags Based on the Dde Protection/deprotection Strategy: Synthesis and Application of 2-Biotinyl-and 2-Hexanoyldimedone. *Tetrahedron Lett.* **1997**, *38* (30), 5391.

(35) 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*, 3617.

(36) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of Proteins by Native Chemical Ligation. *Science.* **1994**, *266*, 776.

(37) Sharma, A.; Jad, Y.; Ghabbour, H.; de la Torre, B.; Kruger, H.; Albericio, F.; El-Faham, A. Synthesis, Crystal Structure and DFT Studies of 1,3-Dimethyl-5-Propionylpyrimidine-2,4,6(1H,3H,5H)-Trione. *Crystals* **2017**, *7* (1), 31.

(38) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to Measure and Predict the Molar Absorption Coefficient of a Protein. *Protein Sci.* **1995**, *4* (11), 2411.

(39) Nath, S.; Mukherjee, P. MUC1: A Multifaceted Oncoprotein with a Key Role in Cancer Progression. *Trends Mol. Med.* **2014**, *20* (6), 332.

(40) Harding, J. S.; Owen, L. N. Dithiols. Part XIII. The Alkaline Hydrolysis of Acetylated Vicinal Hydroxy-Thiols. *J. Chem. Soc.* **1954**, *0* (0), 1528.

(41) Gates, Z. P.; Stephan, J. R.; Lee, D. J.; Kent, S. B. H.; Li, J. Rapid Formal Hydrolysis of Peptide- $^{\alpha}$ Thioesters. *Chem. Commun.* **2013**, *49* (8), 786.

(42) Lechner, C. C.; Becker, C. F. W.; Shiba, K.; Dong, E.; Yang, X.; Liu, G.; Deng, X.; Wang, Y.; Yang, S. T.; Wang, F.; et al. Immobilising Proteins on Silica with Site-Specifically Attached Modified Silaffin Peptides. *Biomater. Sci.* **2015**, *3* (2), 288.

(43) Edwards, J. O.; Pearson, R. G. The Factors Determining Nucleophilic Reactivities. *J. Am. Chem. Soc.* **1962**, *84* (1), 16.

(44) Jencks, W. P.; Carriuolo, J. Reactivity of Nucleophilic Reagents toward Esters. *J. Am. Chem. Soc.* **1960**, *82* (7), 1778.

Chapter 3

Chapter 3. A new water-compatible thioester-functionalized solid support

Having in hands the cysteine-functionalized N-terminal linker described in chapter 2, for the *catch*-and-*release* purifications of cysteine-rich peptides and successive solid supported NCL, the second key element towards the realization of my thesis project was a thioester-functionalized solid support. To our knowledge, no such support is neither currently commercially available not have been described in the literature. We hence synthesized and evaluated a series of thioester-functionalized solid supports potentially adapted to our goals.

1. Selection of a solid support adapted to the NCL

From a physicochemical point of view, an optimal solid support must be compatible with the aqueous conditions required for the solvation of unprotected peptides, thus excluding the polystyrene-based and other related hydrophobic resins used for SPPS. In addition, the solid support has to be chemically-compatible with the conditions used for the capture and release steps but also to the successive ligations and washing steps.

Until now, only a few types of solid supports have been shown to be compatible with solid supported NCLs. In its pioneering work,[1] Kent used Spherilose, a cellulose-based resin previously developed for the affinity purification of proteins. Similarly, Dawson then reported the use of a cross-linked agarose resin (Sepharose, also widely used for affinity purification).[2] Seitz used the same type of resin in a very recent work.[3] While being well adapted to aqueous conditions, these two types of solid supports are associated with limited chemical and mechanical stability. Kent later made use of a polyethylene glycol (PEG)-poly(oxetane)copolymer, SPOCC, a polar and highly permeable solid support developed for combinatorial chemistry and enzymatic reactions.[4,5] In more recent works another type of PEG-based resin have been used: PEGA.[6–10]

Based on our previous experience we have selected two different types of solid support: PEGA resin and control porous glass (CPG).[9,10]

Amino-PEGA resins are hydrophilic PEG-polyacrylamide co-polymers (Figure 1).[11]

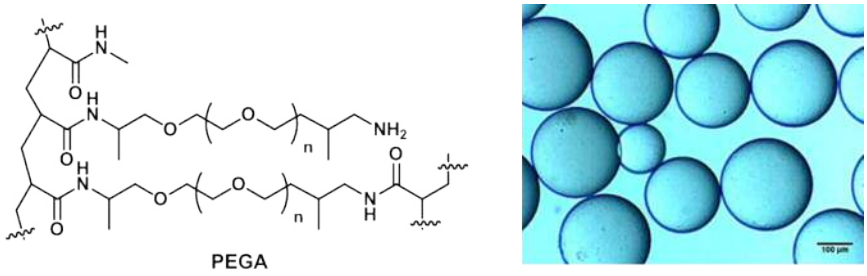


Figure 1. Chemical structure of amino-PEGA resins and microscopy view of $PEGA_{800}$ beads.

PEGA resins are especially well swollen in water, but also in a wide range of organic solvents such as DMF and methanol.[12] Several PEGA resins have been developed, differing in the length of the PEG chain. Two different ones have been previously employed for SPCL: $PEGA_{800}$ (800 Da PEG)[6–8] and $PEGA_{1900}$ (1900 Da).[9] The latter is no longer commercially available, and we thus selected $PEGA_{800}$ for this work (loading: 0.21 mmol/g).

Controlled pore glass CPG is constituted of silica glass forming porous beads with pores of 75 to 3000 angstroms (Å) in diameter, used for solid-phase oligonucleotide synthesis. CPG are chemically stable and are compatible with both aqueous and organic conditions. It does not swell or shrink like polymeric resins.

In a recent work in the team, the use of CPG was proposed for the SPCLs *via* triazole ligations for the synthesis of a very long peptide of 160 amino acid residues.[10] Similar to this last work, for this thesis, we have selected a commercially-available aminopropyl-CPG (Biosearch Technologies), with a pore size of 1000 Å and a loading of 62 µmol/g (Figure 2).

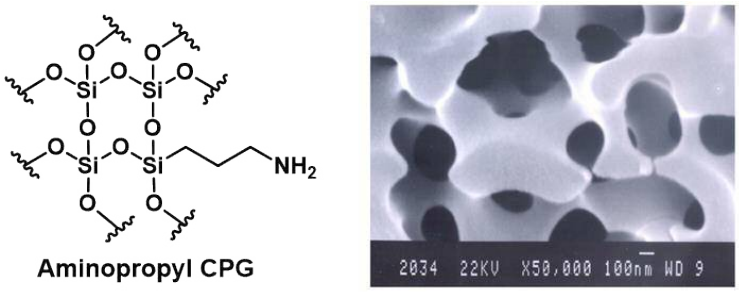


Figure 2. Left, chemical structure of aminopropyl CPG. Right, scanning electron microscope photograph of CPG matrix (image extracted from publication 13).[13]

2. Functionalization of the solid support with a thioester moiety

We have functionalized the PEGA and CPG solid supports with a thioester moiety (Figure 3). Strategically, we decided to introduce an orthogonally cleavable linker between the solid support and the thioester, which will be used to quantify the further ligations. The linker of choice was the acid-cleavable PAL linker (5-[3,5-dimethoxy-4-(aminomethyl) phenoxy]pentanoic acid).

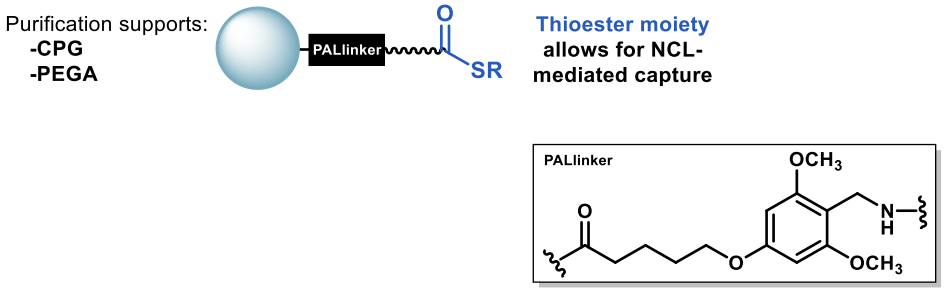


Figure 3. Proposed thioester functionalized solid support.

We investigated two different strategies to prepare the thioester-functionalized solid support based either on a solid phase thioesterification or on a thioesterification in solution.

2.1. Solid phase thioesterification

We first explore the solid phase thioesterification strategy that we thought to be simpler to implement due to the facility to remove the excess of reagents.

TFA-labile Pal linker was first introduced onto the aminopropyl-CPG resin. The resulting Pal-CPG **10**, was then coupled with succinic anhydride in the presence of *i*Pr$_2$EtN in DMF to yield **11**. Analytical TFA-cleavage of few beads of **11** yields the expected compound **11´.**

For the thioesterification, we chose sodium 2-mercaptoethanesulfonate (MesNa), a cheap and commercially available, water-soluble odorless thiol which is commonly used as thiol additive in the NCL reaction.[14] MesNa was coupled to **11** using PyAOP. TFA-cleavage of a few beads of the resulting support **12** gave the expected **12´**, which was contaminated with many impurities, where **11´** was the major product. Despite all our efforts, we did not find conditions for efficient solid phase thioesterification. Thus we decided to perform the thioesterification in solution.

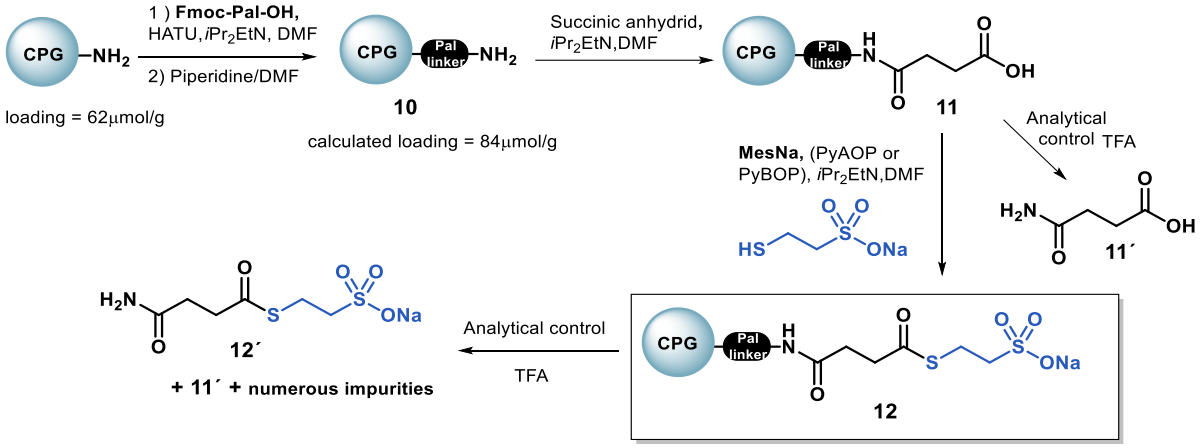


Figure 4. Attempted synthesis of solid support 12.

2.2. Thioesterification in solution

We decided to synthesize a thioester/carboxylic acid hetero-bifunctional linker in solution and then to introduced it into the solid support through standard amide coupling (Figure 5).

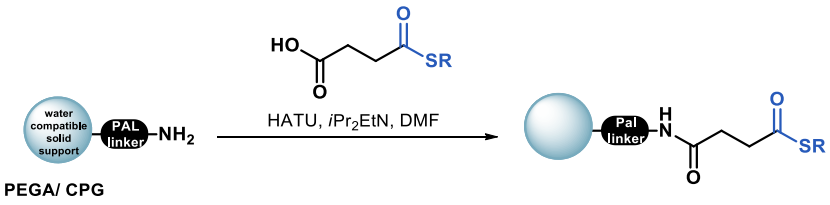


Figure 5. Strategy followed for the synthesis of the thioester-functionalized solid support through thioesterification in solution.

We first tried to synthesize the heterobifunctional linker **13** (Figure 6), by reacting MesNa and succinic anhydride in DMSO. Unfortunately, the resulting thioester was highly contaminated with succinic acid and MesNa. Due to its high hydrophilicity and instability in aqueous media its purification by either normal or reverse-phase column chromatography was impossible.

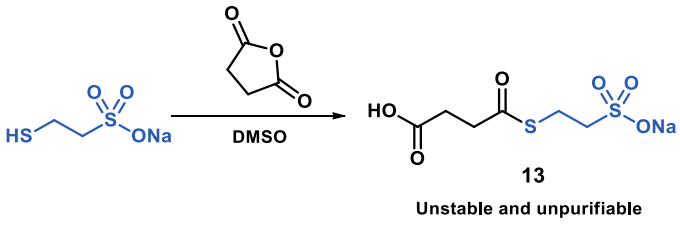


Figure 6. Condensation of sodium 2-mercaptoethanesulfonate and succinic anhydride.

Hence, we decided to synthesize a benzyl thioester. These compounds are known to be stable. Following a literature procedure, *S*-benzyl-thiosuccinic acid **14** was successfully obtained by reaction of benzyl mercaptan with succinic anhydride, and was easily purified by flash chromatography (Figure 7).[15]

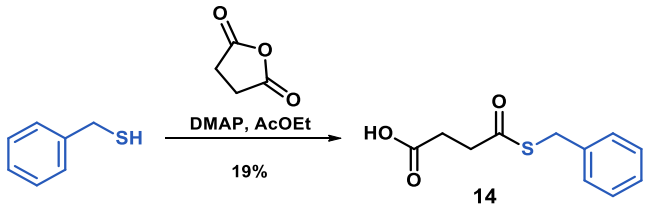


Figure 7. Synthesis of *S*-benzyl-thiosuccinic acid 14.

However, due to the strong odor of benzyl mercaptan which complicated its handling, we decided to synthesize a new thioester, *S*-succinidyl-*N*-acetyl-L-cysteinamide **16** from a non-malodorous thiol (Figure 8).[16]

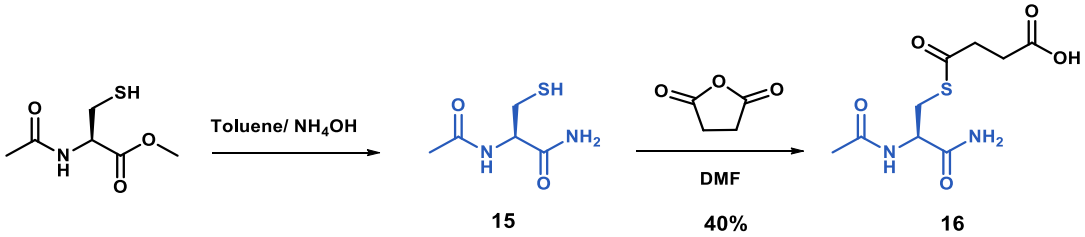


Figure 8. Synthesis of compound 16.

Thioester **16** was obtained through a two-step synthesis starting from commercially available methyl *N*-acetyl L-cysteinate. The first step consisted in the ammoniolysis of the ester, to yield amide **15**, which was used without further purifications.[17,18] Then the nucleophilic attack of thiol **15** on succinic anhydride led to thioester **16,** which was purified by reverse phase chromatography, in a 40% overall yield.

2.3. Synthesis of functionalized solid support 17

Thioester **16** was first coupled to CPG resin **10**. Analytical TFA cleavage yielded the expected compound **17'** (Figure 9).

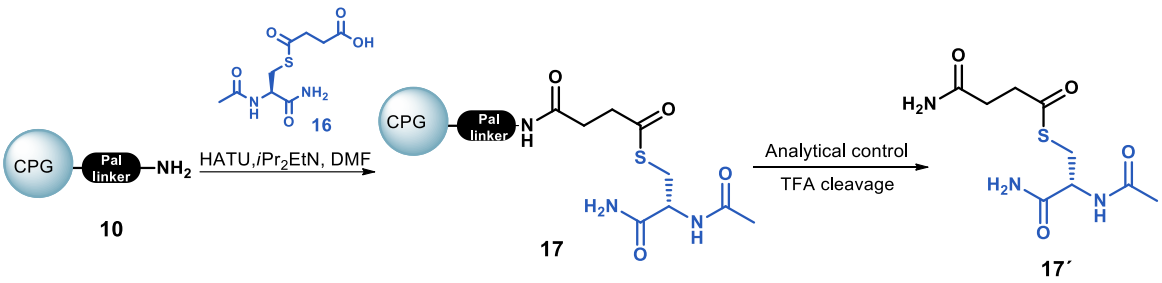


Figure 9. Synthesis of support 17.

Importantly, **17** is completely stable upon storage in NMP: TFA cleavage after six months resulted in pure **17´**, demonstrating the high stability of this thioester.

3. Synthesis of a set of functionalized solid support with different spacers

Following the same reaction protocol, we generated a set of six different thioester solid supports (Figure 10). In two of them, the PAL-linker was substituted by an hydrophilic PEG-based spacer between the thioester and CPG (either 4,7,10-trioxatridecan-succinamic acid (TTDS) or a 3000 Da polydisperse PEG) (Table 1, entries 2 and 3), in order to modify the adsorption properties of the support (see chapter 4) CPG and amino PEGA resins were also coupled directly to **16** to generate **20** and **21** (Table 1, entries 4 and 5).

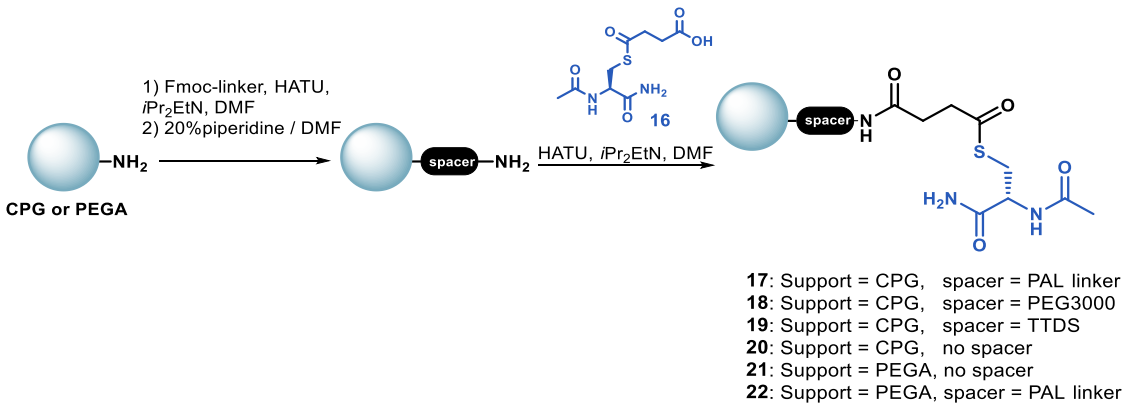
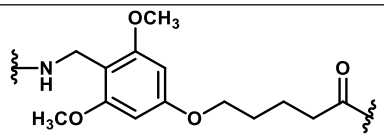
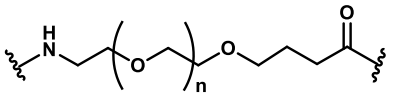
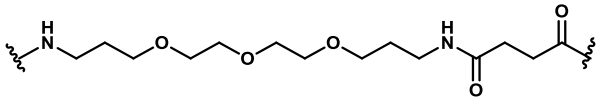


Figure 10. Preparation of the thioester-functionalized capture solid support.

Table 1. Loading of the functionalized CPG and amino PEGA resin.

Entry	Compound	Solid Support	Spacers
1	**17**	CPG	 PAL
2	**18**	CPG	 PEG (3000 Da)
3	**19**	CPG	 TTDS (~ 300 Da)
4	**20**	CPG	-
5	**21**	PEGA	-
6	**22**	PEGA	PAL

4. Conclusion

We have developed a simple strategy for the synthesis of thioester-functionalized water-compatible solid supports. This strategy consisted in the synthesis in solution of the carboxylic acid/thioester heterobifunctional linker **16**. Conveniently, **16** is stable, non-malodorous and easy to purify. This thioester was then coupled into CPG and PEGA solid supports, generating a set of six thioester-functionalized solid supports, using three different spacers (PAL, PEG and TTDS). These will serve to evaluate the influence of the spacers or linkers in the capture of the peptide.

5. References chapter 3

(1) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121* (38), 8720.

(2) Brik, A.; Keinan, E.; Dawson, P. E. Protein Synthesis by Solid-Phase Chemical Ligation Using a Safety Catch Linker. *J. Org. Chem.* **2000**, *65* (12), 3829.

(3) Loibl, S. F.; Harpaz, Z.; Zitterbart, R.; Seitz, O. Total Chemical Synthesis of Proteins without HPLC Purification. *Chem. Sci.* **2016**, *7* (11), 6753.

(4) Johnson, E. C. B.; Durek, T.; Kent, S. B. H. Total Chemical Synthesis, Folding, and Assay of a Small Protein on a Water-Compatible Solid Support. *Angew. Chem. Int. Ed.* **2006**, *45* (20), 3283.

(5) Rademann, J.; Grøtli, M.; Meldal, M.; Bock, K. SPOCC: A Resin for Solid-Phase Organic Chemistry and Enzymatic Reactions on Solid Phase. *J .Am. Chem. Soc.* **1999**, *121* (23), 5459.

(6) Raibaut, L.; Adihou, H.; Desmet, R.; Delmas, A. F.; Aucagne, V.; Melnyk, O. Highly Efficient Solid Phase Synthesis of Large Polypeptides by Iterative Ligations of bis(2-Sulfanylethyl)amido (SEA) Peptide Segments. *Chem. Sci.* **2013**, *4* (10), 4061.

(7) Jbara, M.; Seenaiah, M.; Brik, A. Solid Phase Chemical Ligation Employing a Rink Amide Linker for the Synthesis of Histone H2B Protein. *Chem. Commun.* **2014**, *50* (83), 12534.

(8) Ollivier, N.; Desmet, R.; Drobecq, H.; Blanpain, A.; Boll, E.; Leclercq, B.; Mougel, A.; Vicogne, J.; Melnyk, O.; Todeschini, R.; et al. A Simple and Traceless Solid Phase Method Simplifies the Assembly of Large Peptides and the Access to Challenging Proteins. *Chem. Sci.* **2017**, *11*, 135.

(9) 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.

(10) Galibert, M.; Piller, V.; Piller, F.; Aucagne, V.; Delmas, A. F.; Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; et al. Combining Triazole Ligation and Enzymatic Glycosylation on Solid Phase Simplifies the Synthesis of Very Long Glycoprotein Analogues. *Chem. Sci.* **2015**, *6* (6), 3617.

(11) Meldal, M. Pega: A Flow Stable Polyethylene Glycol Dimethyl Acrylamide Copolymer for Solid Phase Synthesis. *Tetrahedron Lett.* **1992**, *33* (21), 3077.

(12) Renil, M.; Ferreras, M.; Delaisse, J. M.; Foged, N. T.; Meldal, M. PEGA Supports for Combinatorial Peptide Synthesis and Solid-Phase Enzymatic Library Assays. *J. Pept. Sci.* **1998**, *4* (3), 195.

(13) Kutnjak, Z.; Kralj, S.; Lahajnar, G.; Žumer, S. Influence of Finite Size and Wetting on Nematic and Smectic Phase Behavior of Liquid Crystal Confined to Controlled-Pore Matrices. *Phys. Rev. E* **2004**, *70* (5), 51703.

(14) Kang, J.; Reynolds, N. L.; Tyrrell, C.; Dorin, J. R.; Macmillan, D. Peptide Thioester Synthesis through N-->S Acyl-Transfer: Application to the Synthesis of a Beta-Defensin. *Org. Biomol. Chem.* **2009**, *7* (23), 4918.

(15) Stetsenko, D. A.; Gait, M. J. Efficient Conjugation of Peptides to Oligonucleotides by "Native Ligation." *J. Org. Chem.* **2000**, *65*, 4900.

(16) Tsuda, S.; Yoshiya, T.; Mochizuki, M.; Nishiuchi, Y. Synthesis of Cysteine-Rich Peptides by Native Chemical Ligation without Use of Exogenous Thiols. *Org. Lett.* **2015**, *17* (7), 1806.

(17) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. From Disulfide-to Thioether-Linked Glycoproteins. *Angew. Chem. Int. Ed. Engl.* **2008**, *47* (12), 2244.

(18) Danheiser, R. L.; Miller, R. F.; Brisbois, R. G.; Park, S. Z. An Improved Method for the Synthesis of .alpha.-Diazo Ketones. *J. Org. Chem.* **1990**, *55* (6), 1959.

Chapter 4

Chapter 4. Application of the *catch*-and-*release* strategy to the purification of a β-defensin difficult peptide

1. Introduction

In the two previous chapters, we developed a useful N-terminal linker adapted to the synthesis of DRPs, and a thioester solid support compatible with the NCL conditions. Next step was to apply them to the *catch*-and *release* purification of a peptide which had been previously synthesized in the team: the avian-β-defensin 2 (AvBD2). Its synthesis by Fmoc-SPPS results in the formation of a high number of truncated peptides. Hence this peptide constitutes a very good example for the application of our purification methodology.

2. Presentation of the target: Avian β-Defensin 2 (AvBD2)

2.1. Avian β defensins

Beta defensins are the only family of defensins found in avian species.[1] Most of known avian β-defensins (AvBD) have been isolated from domestic birds: chicken, turkeys, ostriches, and ducks. However little is known about defensins in wild bird.[1,2]

In general they are mainly expressed in bone marrow, but they have been also found in skin, egg, respiratory, digestive and urogenital tract.[1] Gene expression in others tissues can be induced by infections and varies between species.[1]

Like other β-defensins, AvBDs contain six cysteines forming 3 disulfide bridges (C1-C5, C2-C4, C3-C6 connectivity), and their three-dimensional structures include three antiparallel β-sheets. In order to illustrate the similarities between avian and mammalian β-defensins, some examples of amino acid sequences (consensus sequence: x_n-C-x_{2-4}-G-x_{1-2}-C-X_{3-5}-C-X_{9-10}-C-X_{5-6}-CC-X_n) are shown in table 1.[1,3]

Table 1. Amino acid sequences of some avian beta defensins.[3]

Source	Peptide	Amino acid sequence
Consensus		**-----C----G-C----C---------C------CC-----**
Human	hBD-1	DHYN**C**VSSG**G**Q**C**LYSA**C**PIFTKIQGT**C**VGRAVK**CC**RKK
Bovine	TAP	NPVS**C**VRNK**G**I**C**VPIR**C**PGSMKQIGT**C**YRGKAK**CC**K
Chicken	AvBD1	GRKSD**C**FRKS**G**F**C**AFLK**C**PSLTLISGK**C**SRFYL-**CC**KRIW
	AvBD1α	GRKSD**C**FRKN**G**F**C**AFLK**C**PYLTLISGL**C**SRFHL-**CC**KRIW
	AvBD2	LF**C**--KG**G**S**C**HFGG**C**PSHLIKVGS**C**FGFRS-**CC**KWPWNA
	AvBD3	GTATQ**C**RIRG**G**F**C**RVGS**C**RFPHIAIGK**C**A-TFIS**CC**GRAY
Turkey	AvBD1	GKREK**C**LRRN**G**F**C**AFLK**C**PTLSVISGT**C**SRFQV-**CC**
	AvBD2	LF**C**--KR**G**T**C**HFGR**C**PSHLIKVGS**C**FGFRS-**CC**KWPWDA
Ostrich	AvBD2	LF**C**--RK**G**T**C**HFGG**C**PAHLVKVGS**C**FGFRA-**CC**KWPWDV

hBD-1: human beta-Defensin-1, TAP: tracheal antimicrobial peptide.

2.1.1. Chicken Avian β-defensin 2, AvBD2

Chicken β defensins (formerly known as gallinacins)[3] are the most studied bird β-defensins. Fourteen different genes have been found in the chicken genome,[4] encoding fourteen different defensins, named from AvBD1 to AvBD14.[1]

Our research group has a long-standing interest in AvBD2, the most abundant defensin in chicken bone marrow.[5] Its extraction and characterization in 2009, was reported in the team in collaboration with Lalmanach´s team.[5] In collaboration with Landon´s NMR group at CBM it was also reported in 2012 its synthesis, three-dimensional structure and insight into structure-activity (Figure 1).[6]

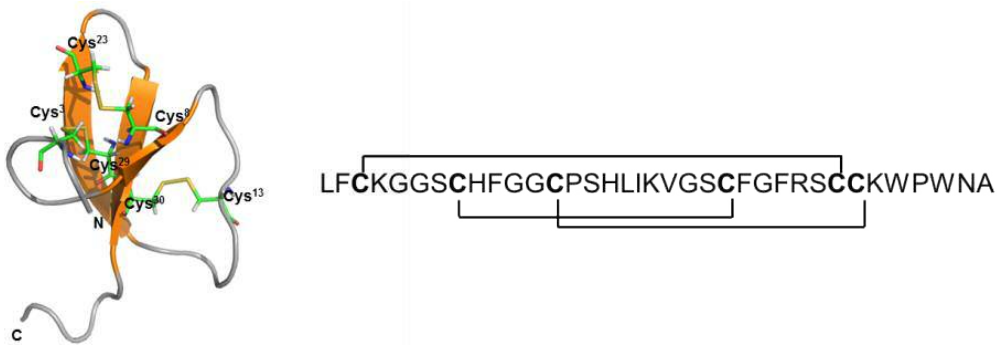


Figure 1. AvBD2 3D-structure.

Solid phase synthesis and purification of AvBD2 have been studied in detail in the team. The crude peptide is highly contaminated with a large number of truncated peptides and is difficult to purify by HPLC due to the presence of some co-eluted truncated peptides.

We thus thought that AvBD2 was an excellent target for the application of the *catch*-and-*release* purification strategy making use of our Cys-Dtph linker.

3. Synthesis of Cys-Dtph-AvBD2 by Fmoc-SPPS

3.1. Synthesis of AvBD2[2-36]

We synthesized the AvBD2[2-36] peptide (sequence: ^{2}FCKGGSCHFGGCPSHLIKVGSCFG FRSCCKWPWNA^{36}). As mentioned, in our strategy the peptide is synthesized without the last N-terminal residue which is introduced at the end of the elongation protected with the Boc-Cys(Trt)-Dtph linker. Peptidyl resin **23** was obtained in a 24% elongation yield by automated Fmoc-SPPS starting from commercially available aminomethyl Tentagel R, using a methylphenoxypropionic acid linker (MPPA) generating a C-terminal carboxylic acid peptide upon TFA-mediated cleavage of the peptidyl resin.

An aliquot of resin **23** was cleaved to characterize the corresponding peptide **23´** (Figure 2).

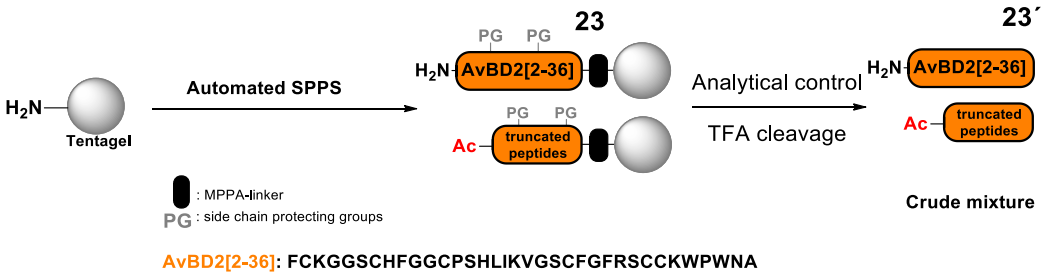


Figure 2. Synthesis and analysis of the crude AvBD2 23 peptide.

As we expected, the crude mixture presented in addition to the target peptide **23´** a high number of other compounds whose mass spectra were consistent with acetylated truncated peptides. The most abundant were Ac[18-36] (sequence: Ac-^{18}IKVGSCFGFRSCCKWPWNA^{36}), Ac[17-36] and Ac[16-36] (Figure 3). Unexpectedly, we also observed a peak whose MS data and UV spectra were consistent with the truncated Fmoc[18-36] (sequence: Fmoc^{18}IKVGSCFGFRSCCKWPWNA^{36}) resulting from the incomplete removal of the N$^{\alpha}$-Fmoc protecting group of Ile18.

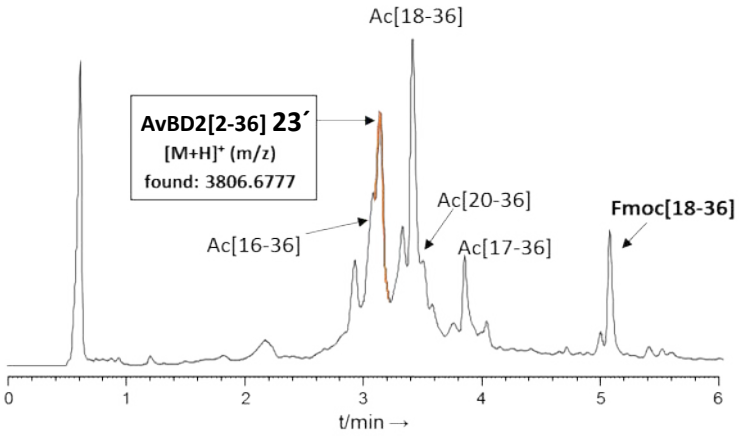


Figure 3. Chromatogram of the crude peptide 23´after TFA release.

3.2. Preparation of Cys-Dtph-AvBD2 24

Boc-Cys(Trt)-Dtph-Leu-OH **4c,** was then coupled to supported peptide **23** to yield Cys-Dtph-peptide **24** using the same conditions that for model peptide **5** in chapter 2 (Figure 4).

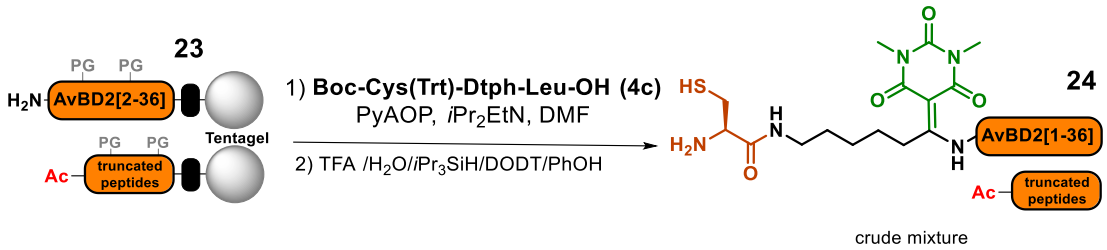


Figure 4. Coupling of Boc-Cys(Trt)-Dtph linker into the peptidyl-resin.

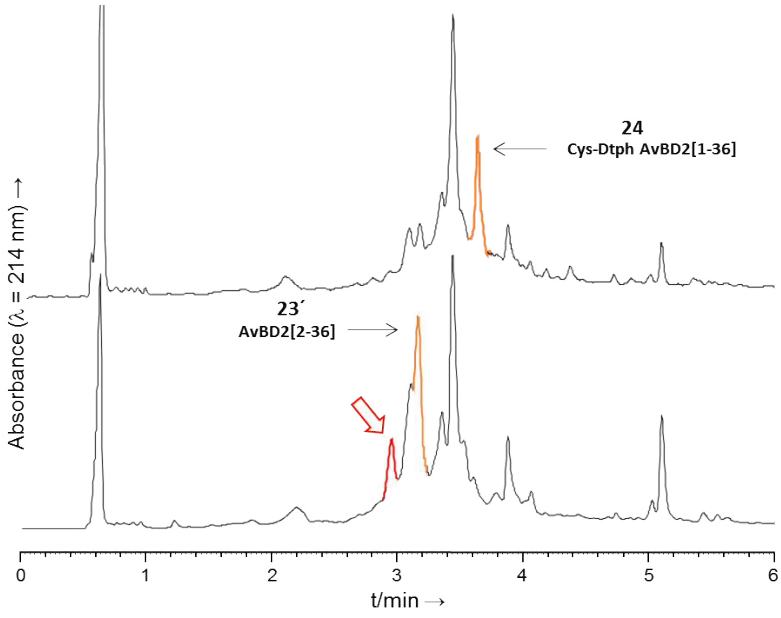


Figure 5 Chromatogram of crude 23´and crude Cys-Dtph-peptide 24 after TFA release.

As expected, we observed the disappearance of the peak which corresponds to peptide **23´**and the apparition of a new peak, whose MS data was consistent with the Cys-Dtph-containing peptide **24** (Figure 5, orange peaks). Unexpectedly, we also observed the disappearance of another peak (Figure 5, red peak) which could indicate that another compound had incorporated our linker. Observation of the chromatogram at λ = 320 nm (specific for our linker) confirmed the hypothesis (Figure 6): the major product **24** is accompanied by a minor peak (red arrow).

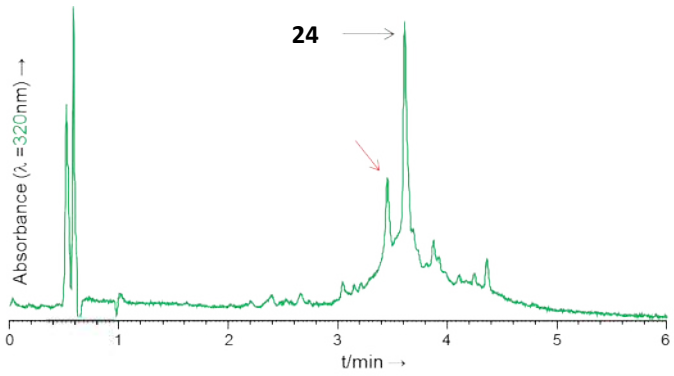


Figure 6. Chromatogram of crude Cys-Dtph-peptide 24 after TFA release (UV absorbance at 320 nm).

LC-MS analysis indicated a difference in mass of 131 Da which would be consistent with the lack of a Leu or Ile residue. Considering the observation of an incomplete removal of the Fmoc group of Ile18, we believe that the minor compound corresponds to a deletion of the Leu17 residue.

The formation of such a **deleted peptide** can result either from incomplete Fmoc deprotection or inefficient capping, and is characteristic of the Fmoc-SPPS of the so-called "difficult sequence" (Table 2).[7] This term makes reference to those peptides which contain a region (consisting in general of four or five amino acids) where the kinetics of coupling and deprotection dramatically slows down, mainly due to aggregation of the growing peptide chains through formation of an inter-chain hydrogen bond network.[7–9]

Table 2. Characteristics of difficult sequences in Fmoc-SPPS. Adapted from Kent[7] and Milton.[9]

1	Repetitive incomplete coupling
2	Limited improvement with recoupling or capping
3	Incomplete removal of the N^{α}-Fmoc protecting group
4	Resin-dependent: aggravated by high resin loading
5	Sequence-dependent: aggravated by sterically hindered AAs
6	Mechanism: intermolecular aggregation of the protected peptide

We carefully analyzed the LC-MS of crude **23'** and we characterized other peaks which masses were consistent with other deleted sequences in the same region (Figure 7).

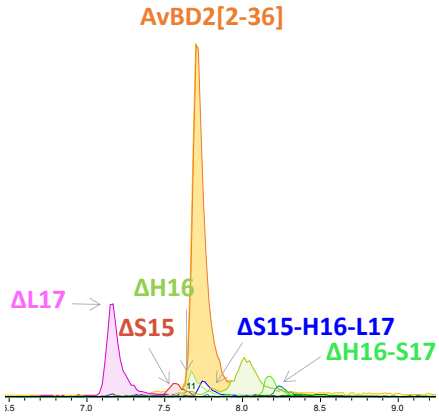


Figure 7. Extracted Ion Chromatograms (EIC) of target AvBD2[2-36] 23´and main deleted peptides.

Most of the deletions occur from Ile18 to Ser15 residues which correspond to the difficult sequence (Figure 8). The presence of deleted peptides missing two and three consecutive amino acids was also observed (Figure 7).

Difficult sequence

H_2N-FCKGGSCHFGGC^{14}P^{15}S^{16}H^{17}LIKVGSCFGFRSCCKWPWNA

Figure 8. Difficult sequence in AvBD2 [2-36] peptide.

The presence of deleted peptides limits the purification by *catch*-and-*release* since at the end of the SPPS they also present a free N-terminal amine.

3.3. Optimization of the SPPS of Cys-Dtph-AvBD2

There are different factors that can contribute to a better solvation of the peptides and thus to reduce their tendency to aggregation such as increase of the temperature, use of different solvents, solid supports or protecting groups.[8–10]

High temperatures increase reaction kinetics and disfavor hydrogen bond-mediated aggregates. However, in some cases these temperatures can conduce to epimerization.

The introduction of backbone amide protecting groups results in the masking of the hydrogen bond donating properties of the peptide bond, thus preventing the formation of β-sheets and aggregates through hydrogen bond interactions.[11,12] Four examples of these protecting groups are *N*-2,4-dimethoxybenzyl (Dmb),[13,14] *N*-2-hydroxy-4-methoxybenzyl (Hmb)[15] and its analog *N*-2-hydroxy-4-nitrobenzyl (2,4-Hnb),[16] or pseudoprolines. The latter are used for Ser and Thr residues, and are based on a *N*,*O* ketal protecting both the side chain alcohol and the backbone amide.[8,17–20]

3.3.1. Influence of the SPPS resin

The nature of the resin influences the success of the SPPS.[21] If polystyrene-based resins solid supports have been widely used for decades, [22] PEG-based or other polar resins seem to be most adapted to the synthesis of difficult peptides.[22,23]

In our first approach, we used Tentagel, a polystyrene matrix on which PEG chains are grafted.[24] To try to suppress the apparition of these deleted sequences, we have decided to repeat the synthesis of AvBD2 using two more polar resins: **ChemMatrix**[22,25,26] **and Spheritide**.[27] ChemMatrix is a 100% PEG-based resin and Spheritide is a novel hydrophilic resin, constituted by poly-lysines cross-linked with dicarboxylic acids (Figure 9).

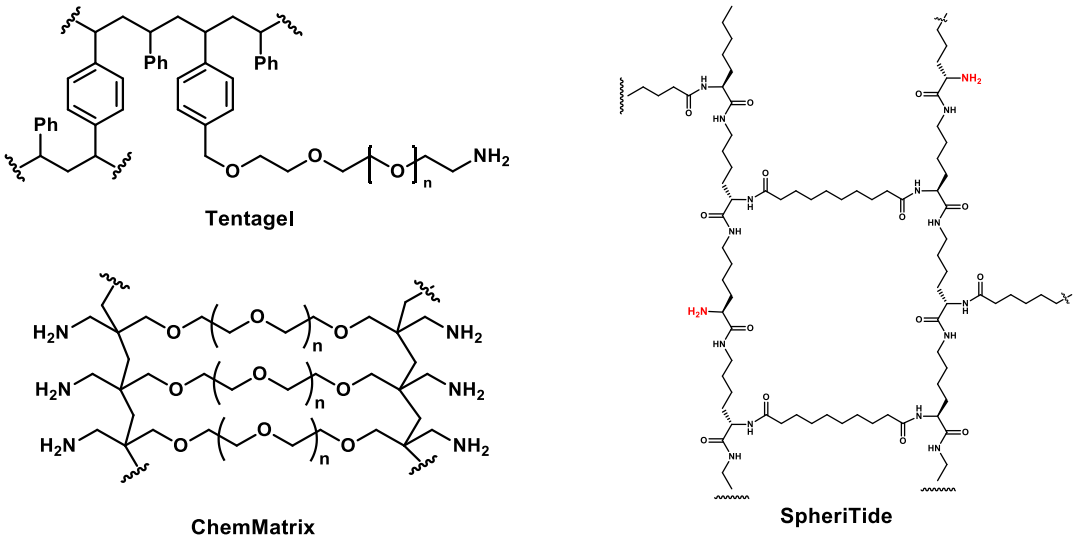


Figure 9. Structure of Tentagel, ChemMatrix and Spheritide resins.

To facilitate the description of these results, we have established a special nomenclature based on the assignation of a capital letter code (A, B, C) to each of the different synthesis of AvBD2: A for the previously reported synthesis using Tentagel, and B and C for the new syntheses using ChemMatrix and Spheritide, respectively.

The elongation yields were 34% for **23B** (ChemMatrix) and 36% for **23C** (Spheritide), significanty higher than the 24% obtained for **23A** (Tentagel).

The crude peptide mixtures after SPPS (**23'A-C**) and after the introduction of our linker (**24A-C**) were studied in parallel (Figure 10).

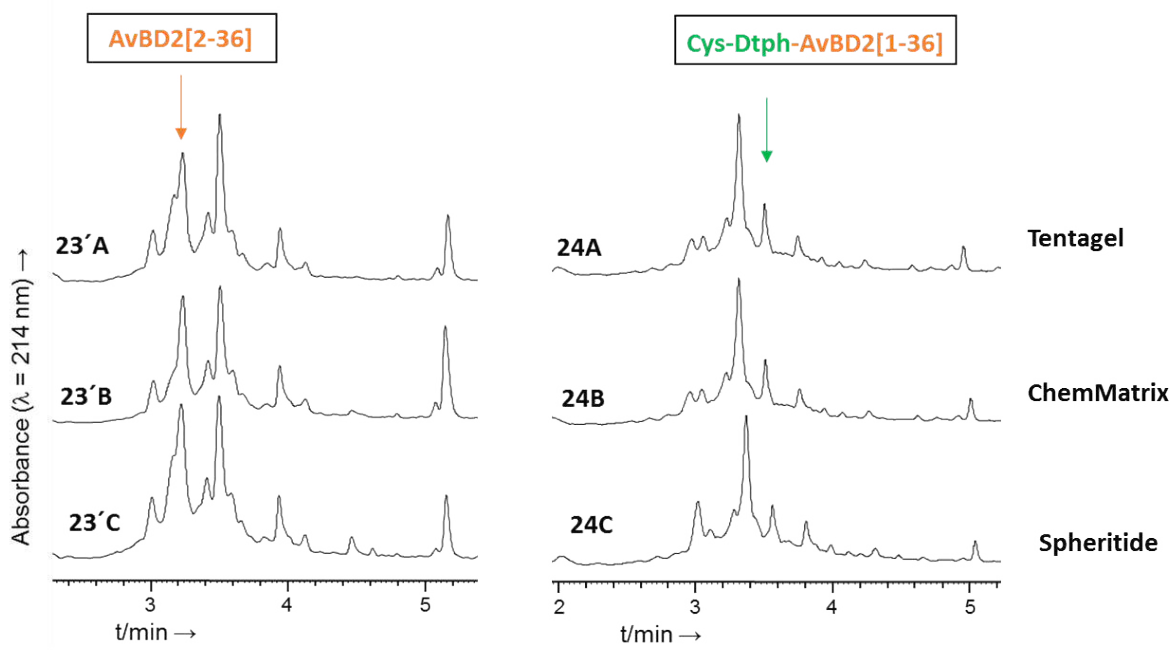


Figure 10. Left, HPLC chromatogram of AvBD2[2-36]crude mixtures 23´A-C. Right, HPLC chromatogram of Cys-Dtph-AvBD2[1-36] crude mixtures 24A-c.

HPLC chromatograms of **23´A-C** were very similar.

In order to detect differences between the Cys-Dtph-containing peptides **24A-C**, we also observed the chromatograms at λ = 320nm (specific for our linker) (Figure 11) but disappointingly, we could not appreciate any significant difference.

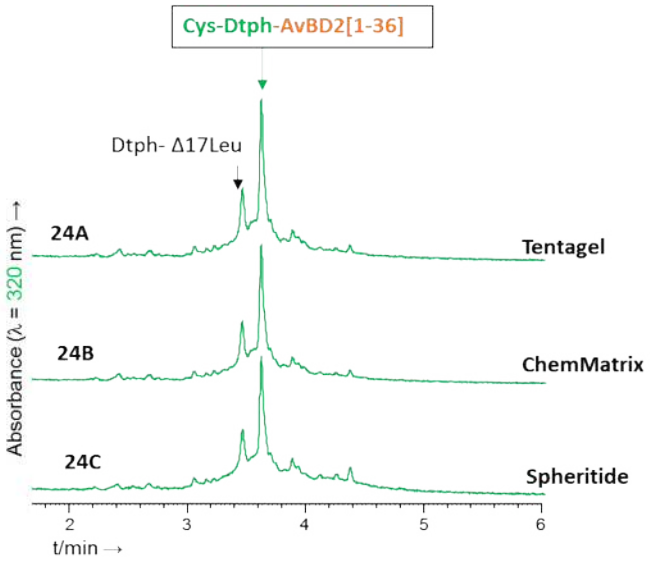


Figure 11. HPLC profile of crude Dtph peptides 24A-C at 320 nm.

The best results were obtained with ChemMatrix and Spheritide. We have selected ChemMatrix for our next studies.

3.3.2. Optimization of the Fmoc deprotection conditions

Given that deleted peptides come from an incomplete Fmoc deprotection, we next try to optimize the Fmoc removal conditions in the region which corresponds to the difficult sequence (from Ile18 to Pro15). Thus we have repeated the synthesis but instead of the classical Fmoc deprotection protocol which consist of three consecutive treatments with a 20% solution of piperidine in DMF for three minutes (table 3, entry 1), we increased the reaction time and/or the temperature (Table 3, entries 2-4). We also tried to use a deprotection solutions containing 2% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), a stronger base than piperidine[28] previously used for the synthesis of difficult peptides (Table 3, entry 5).[28–30] Following the nomenclature introduced earlier, capital letter code D to G were assigned to these 4 new syntheses.

Table 3 . Fmoc deprotection conditions.

Entry	peptidyl-resin	Fmoc deprotection reagent	Treatment 1		Treatments 2-3		Elongation yield
			t (min)	T (°C)	t (min)	T (°C)	
1	**23B**	20% piperidine/DMF	3	20	3	20	34
2	**23D**	20% piperidine/DMF	3	20	30	20	33
3	**23E**	20% piperidine/DMF	3	20	3	60	34
4	**23F**	20% piperidine/DMF	10	20	10	60	30
5	**23G**	20% piperidine/2% DBU/DMF	3	20	3	20	25

The HPLC profiles of the crude mixtures of AvBD2[2-36] (**23´B,D-G**) and of the Cys-Dtph-AvBD2[1-36] (**24B,D-G**) presented important differences (Figure 12).

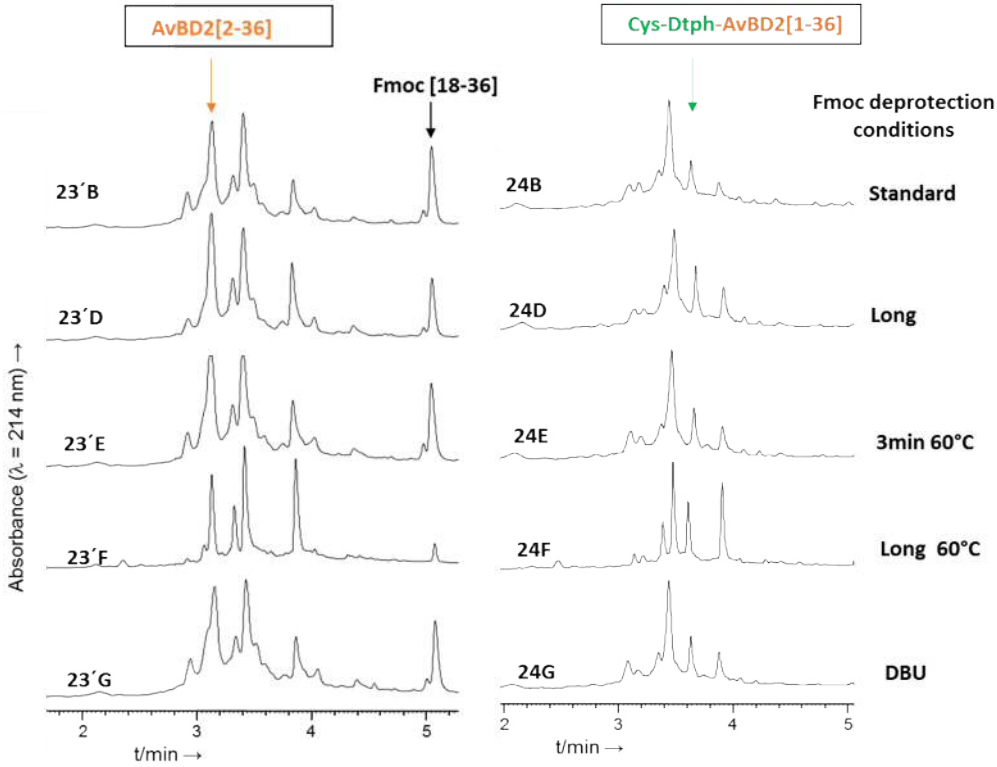


Figure 12. Left, HPLC chromatogram of AvBD2[2-36]crude mixtures 23´B,D-G. Right, HPLC chromatogram of Cys-Dtph-AvBD2[1-36] crude mixtures 24B,D-G. Compound 24F was analyzed using a C4 column, while all the other compounds were analyzed with a C18 column. The time scale for 24F is different (see experimental section).

We observed a great diminution of the Fmoc[18-36] peptide peak in the case of compound **23´F** (Figure 12) which corresponds to a long treatment at 60°C (Table 3, entry 4), indicating a much more efficient Fmoc deprotection under these conditions.

In accordance with this observation, careful LC-MS analysis of **23´F** revealed only very low amounts of deleted peptides, and chromatogram of Cys-Dtph-containing AvBD2 (**24F**) observed at λ = 320 nm is much cleaner than the one of **24B** corresponding to standard deprotection conditions (Figure 13).

Note that the increase of the deprotection time (**24D**) and addition of DBU in the deprotection mixture (**24G**) (table 3, entries 2 and 5 respectively) also greatly improved the quality of the 320 nm chromatogram, but significant amounts of deleted sequences are still observed (Figure 13).

We have chosen **24F** for the application to the *catch*-and-*release* purification.

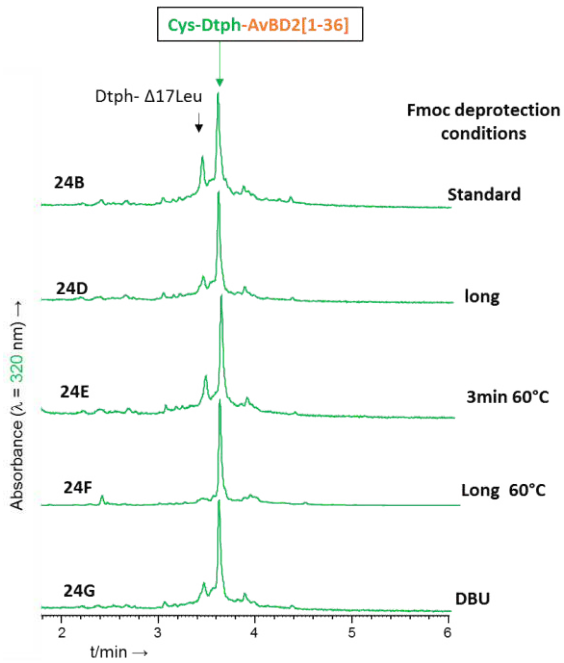


Figure 13. HPLC profile of crude Dtph-peptides 24B, D-G at 320 nm.

3.3.3. Optimization of the TFA cleavage conditions

During the LC-MS analysis of **24A-F**, we also found a peak showing a mass increase of 56 Da compared to the target compound, consistent with a *tert*-butyl group.[31,32] Increase of the cleavage time did not reduce the amount of this side product, suggesting that it does not come from an incomplete deprotection of the *tert*-butyl ethers and esters of Ser, Thr, Asp and Glu, but rather from the re-alkylation of an unprotected cysteines by reaction with the *tert*-butyl cation during the cleavage.[33,34] This side reaction is quite common during the synthesis of cysteine-rich peptide.

To try to reduce the amount of this side product, we substituted DODT in the cleavage cocktail for ethanedithiol (EDT), which is known as a very effective *tert*-butyl cation scavenger.

Unfortunately, the use of EDT does not avoid this *S*-alkylation (Figure 14). We have thus decided to continue using DODT instead of EDT which has a strong and very unpleasant odor and is thus difficult to handle.[35,36]

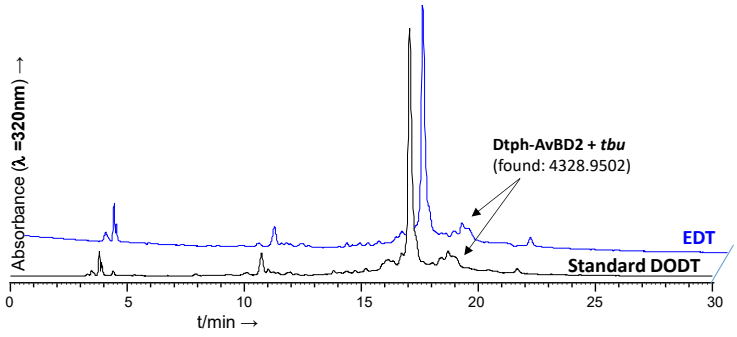


Figure 14. HPLC chromatogram of crude Cys-Dtph-peptide 24F at 320 nm.

4. *Catch*-and-*release* purification of AvBD2

4.1. Preliminary studies of the *catch*-and-*release* purification

In order to identify the best solid support for the realization of our purification strategy, six preliminary *catch*-and-*release* purifications were done using the functionalized solid supports **18-22** described in chapter 3 (Table 4) . All these studies have been performed with the non-optimized Cys-Dtph-containing peptides **24B** (ChemMatrix, standard Fmoc deprotection conditions).

Table 4. Functionalized solid support used in this study (see chapter 3).

Entry	Compound	Solid Support	linkers/spacer
1	**20**	CPG	-
2	**18**	CPG	PEG3000
3	**19**	CPG	TTDS
4	**21**	PEGA	-
5	**22**	PEGA	PAL

4.1.1. *Catch*-and-*release* purification with a thioester functionalized-CPG

The NCL-mediated immobilization reaction was conducted under standard NCL conditions using an excess (10 equivalents) of a thioester-functionalized CPG **20** (without spacer). The progress of the reaction was monitored by HPLC analysis of the supernatant (Figure 15).

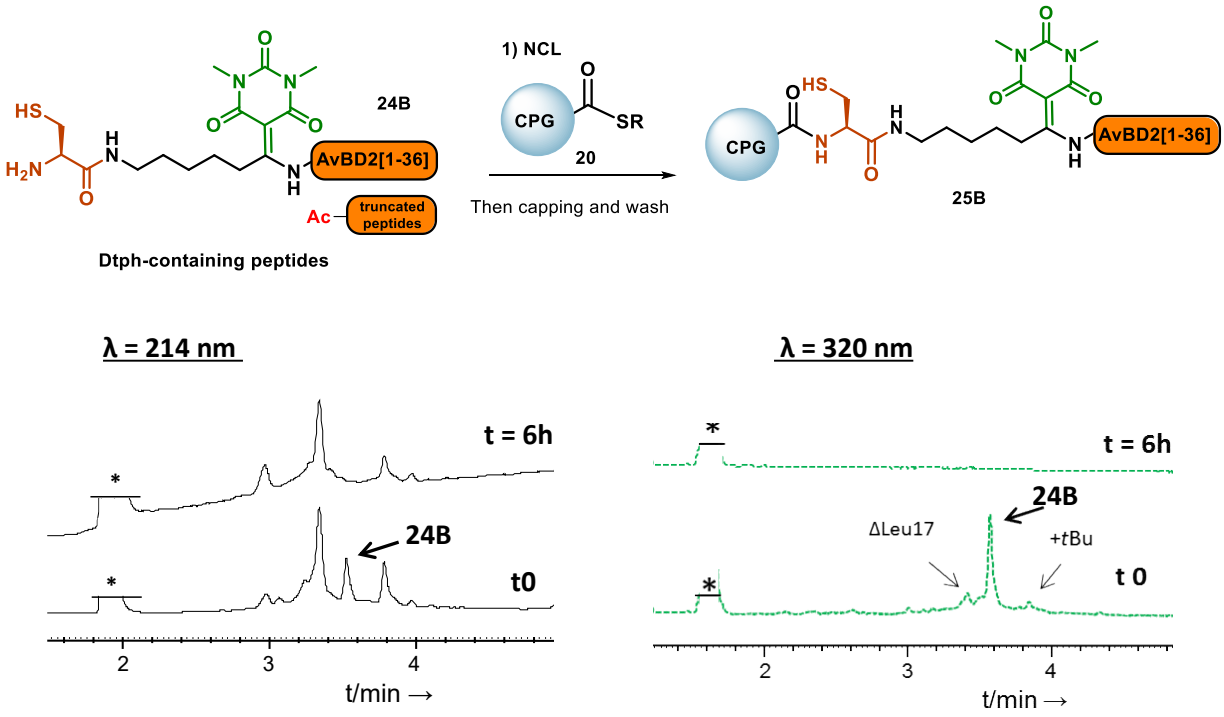


Figure 15. HPLC profile of the supernatant of 24B at 214 and 320 nm during immobilization onto the functionalized-CPG 20.*MPAA traces not removed during Et_2O extraction.

After 6 hours, we observed the complete disappearance of all the tagged peptides from the supernatant, thus we considered that the capture was complete (no trace of any peak at 320 nm confirmed that all the peptides containing the Cys-Dtph linker were immobilized onto the solid support) (Figure 15).

In order to quench the remaining thioesters, an excess of cysteine was then added and the reaction was stirred for an additional hour. The resulting immobilized peptide **25B** was extensively washed in order to remove the acetylated peptides and the excess of reagents.

For the release of the peptide in solution, we used the conditions previously established (1 M aqueous hydrazine solution containing 100 mM TCEP and 100 mM TRIS, pH 8.5) for 2 hours at 37°C. HPLC and LC-MS analysis showed a major peak with a mass consistent with the expected AvBD2[1-36]peptide **26B** (Figure 16). Disappointingly, we also observed the presence of many impurities, including significant amounts of acetylated truncated peptides and MPPA, suggesting an inefficient washing.

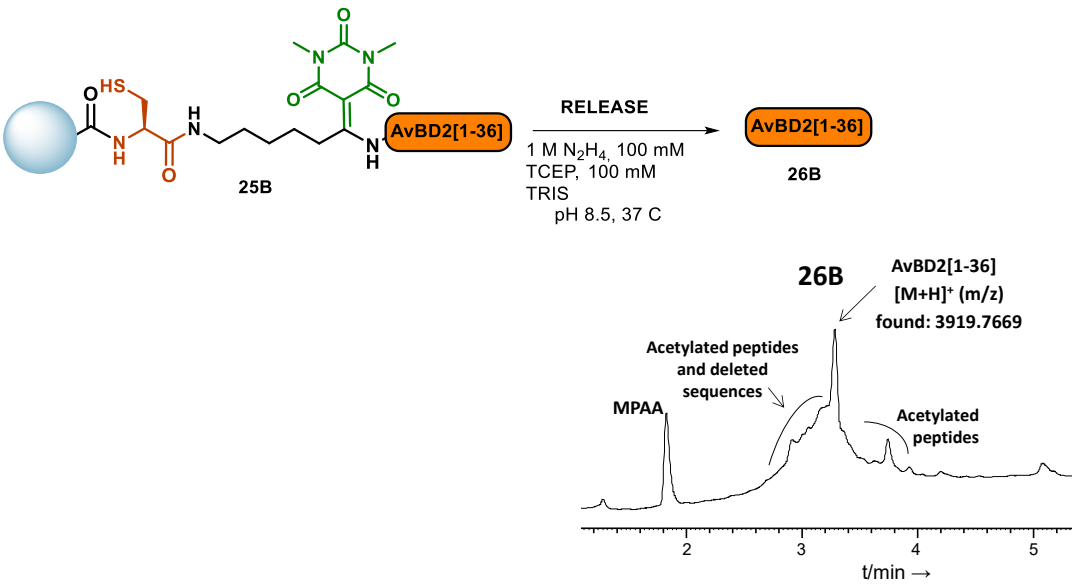


Figure 16 . HPLC profile of the crude mixture 26B after release.

From these preliminary results, we concluded that the peptide containing our Cys-Dtph linker can be efficiently immobilized to the solid support *via* NCL. In addition, the Cys-Dtph linker can be easily cleaved using our optimized nucleophilic conditions to liberate the target AvBD2 peptide into solution.

As expected, deleted peptides generated during the SPPS and peptides containing a *t*Bu moiety were able to incorporate our Cys-Dtph linker. Hence they were also captured on the solid support and then released into solution.

The presence of acetylated peptides and MPPA in the mixture after the cleavage of our linker could be due to absorption at the surface of CPG which complicates their elimination by washing.

To try to limit this problem by modifying the non-specific absorption properties of the CPG surface, we incorporated two types of PEG-based hydrophilic spacers: PEG300 and TTDS (solid supports **18** and **19** respectively, see chapter 3).

The capture and release reactions were repeated under the same conditions as for support **20**. If the NCL-mediated immobilization proceeded well (Figure 17), we inexplicably did not observe any trace of the expected compound after Dtph cleavage, and we did not pursued further studies using these solid supports.

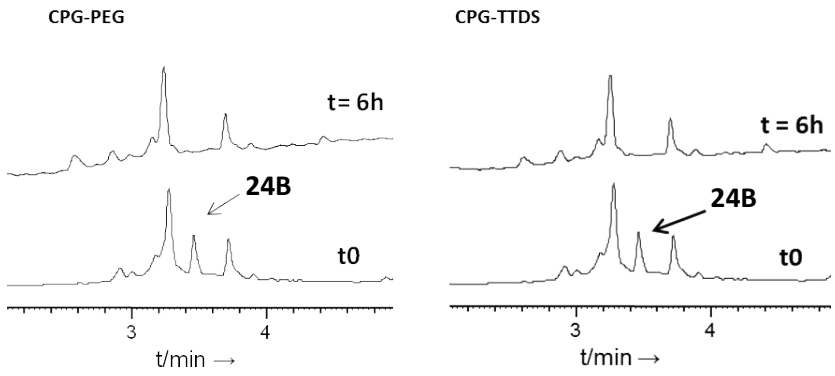


Figure 17. HPLC profiles at 240 nm of the supernatant of 24B during immobilization onto the functionalized-CPG 18 and 19.

4.1.2. *Catch*-and-*release* purification using PEGA

To try to improve the *catch*-and-*release* process by optimizing the washing step, we have then carried out the reactions using the PEGA resin.

Immobilization, washing and release steps were repeated under the same conditions as for CPG, using solid support **22** incorporating a TFA-labile PAL linker. The capture step proceeded similarly. Very disappointingly, hydrazine-mediated Dtph cleavage resulted on the release of much lower amounts of peptide as we expected. Moreover the expected target AvBD2 was highly contaminated (Figure 18).

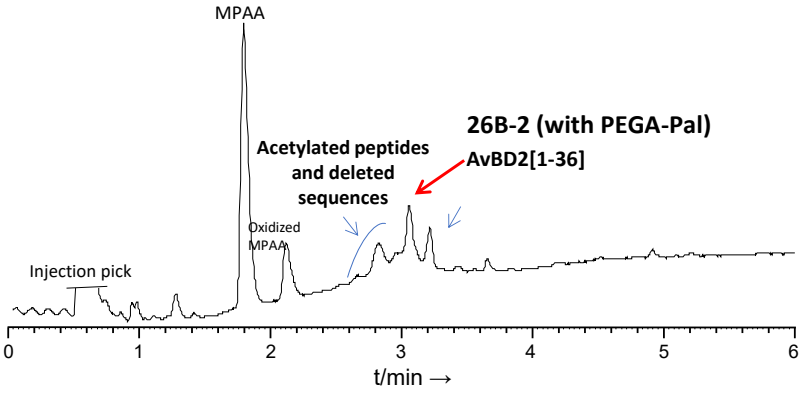


Figure 18. HPLC profile at 240 nm and MALDI-TOF result of release peptide 26B-2.

Given that we used mildly acidic conditions during the washing (0.1% TFA in water and acetonitrile), we suspected an unexpected cleavage of the Pal linker as the cause of the low yield.

Thus we repeated the *catch*-and-*release* process using the thioester functionalized PEGA **21** (not incorporating any spacer or linker). Contrasting to all the previous assays, the HPLC profile of the release peptide **26B-3** was very satisfying, showing a major peak which mass was consistent with the target AvBD2 and two minor peaks corresponding to the expected Leu17-deleted and *S*-*tert*-butylated peptides (Figure 19). Reducing the amount of thioester solid support to 5 or 2 equivalents gave similar results.

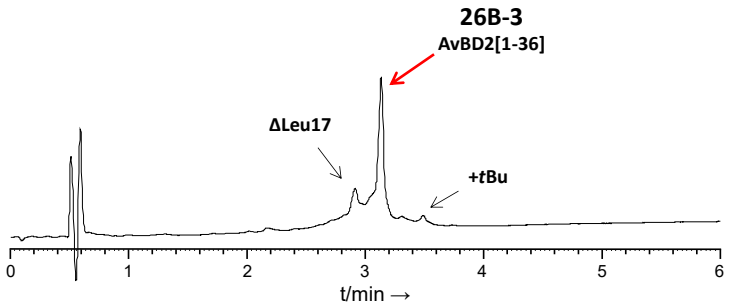


Figure 19. HPLC profile at 240 nm of released peptide 26B-3.

These very promising results allowed confirming the usefulness of our linker for the purification of cysteine-rich difficult peptides *via* a *catch*-and-*release* purification process. The PEGA-based thioester-functionalized solid support **21** is the most adapted for this purpose, and it has been used for all the next applications.

4.2. *Catch*-and-*release* purification of the optimized Cys-Dtph-AvBD2 24F.

The *catch*-and-*release* purification was repeated using the same conditions with the optimized Cys-Dtph-containing peptides **24F** and the thioester functionalized PEGA **21** (Figure 20).

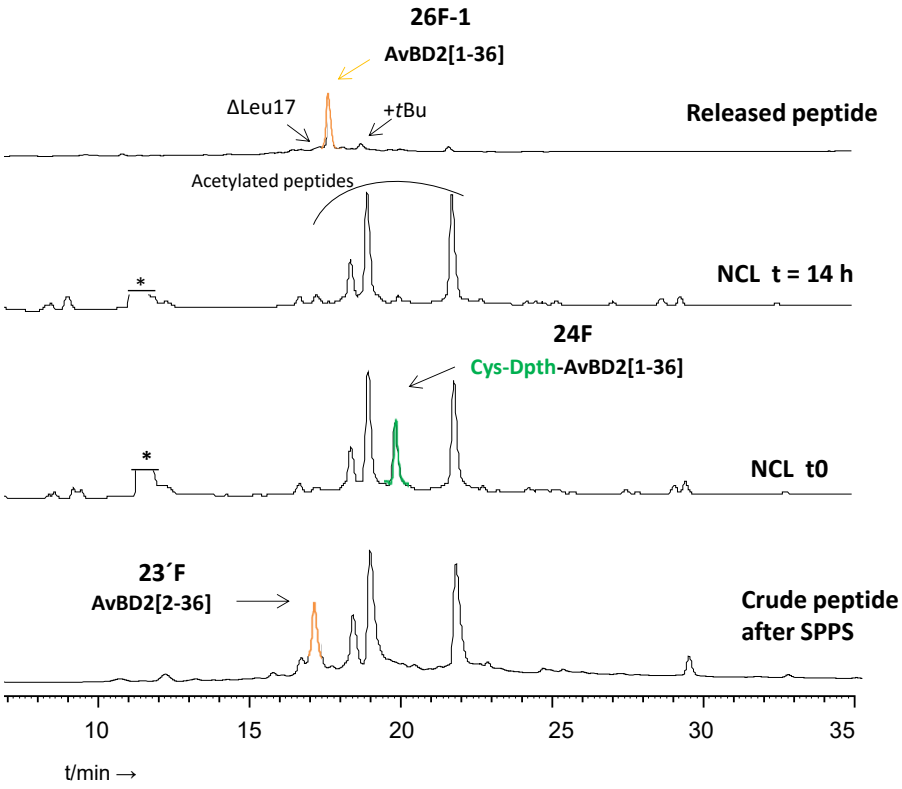


Figure 20. HPLC profiles at 240 nm of the synthesis and *catch*-and-*release* purification of AvBD2. *MPPA.

The amount of peptide after the release was determined by UV spectrophotometry (ε^{280} = 11000 M^{-1} cm^{-1}). Comparison with the starting amount of Cys-Dtph-containing peptide (also determined by UV, but at 320 nm) allowed to estimate the yield of the *catch*-and-*release* process to 95%.

Purity of the released compound was estimated to 85% by integration of the chromatogram at 280 nm. The corrected overall yield* of the entire process: Fmoc-SPPS, Boc-Cys(Trt)-Dtph linker introduction, TFA cleavage, catch and release was 23%.

In order to obtain a very pure peptide, we carried out a "polishing" HPLC which allowed increasing the purity of AvBD2 to more than 98% (Figure 21). However, we observed severe material loss during the HPLC purification, probably due to the low amounts of peptide and non-specific absorption. Deceptively, the corrected yield was only of 3.3%.

*This yield corresponds only to the amount of AvBD2[1-36] in the purified peptide, not taking the 15% impurities into accounts .

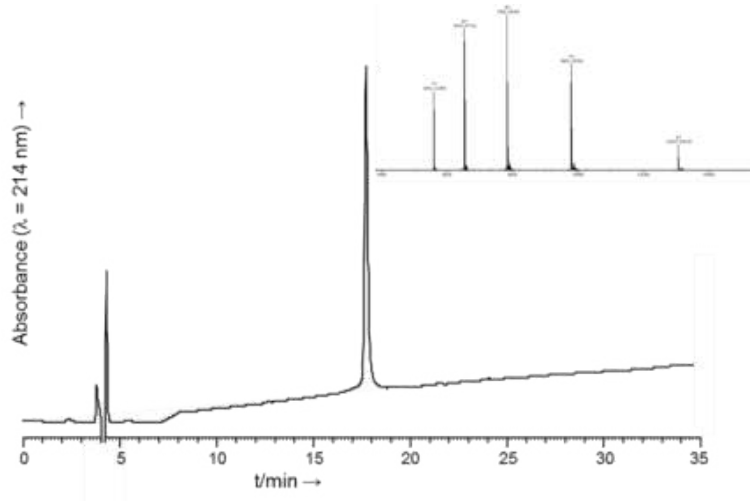


Figure 21. HPLC chromatogram and ESI-HRMS analysis of AvBD2 purified by *catch*-and-*release* followed by HPLC.

In order to rigorously evaluate the benefits of our methodology compared to classical HPLC purification, we synthesized AvBD2[1-36] by Fmoc-SPPS, using the Fmoc deprotection conditions previously optimized, and purified the crude mixture by semi-preparative HPLC.

Due to the close elution with truncated acetylated peptides, the purification yield was very low: 2.5%. If the HPLC analysis of the purified compound showed a single peak (Chromatogram B, Figure 22), its dissymmetry suggested a low purity. Use of a high resolution analytical column, combined with a different mobile phase (formic acid instead of TFA) showed that HPLC-purified AvBD2[1-36] was still contaminated with high amounts of truncated acetylated sequences that co-eluted during HPLC purification and analysis under standard conditions (Chromatogram C, Figure 22). Purity was estimated to only 56%, resulting in a corrected overall yield of only 1.5%.

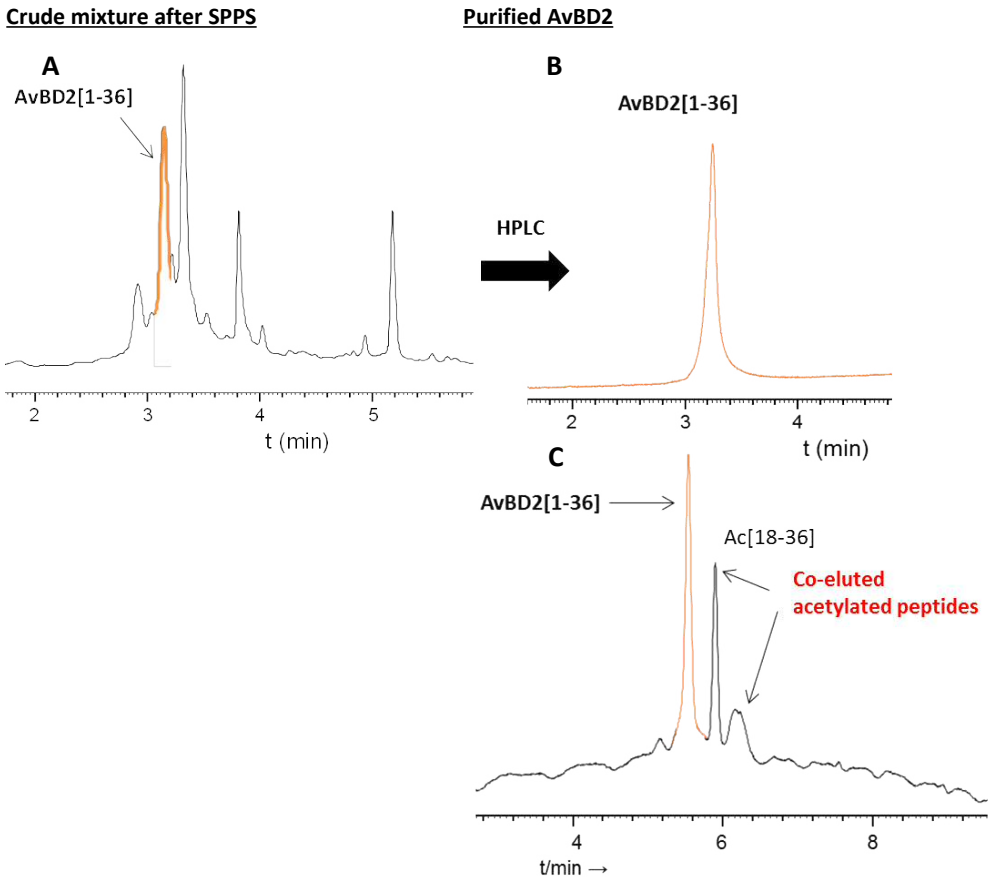


Figure 22. HPLC chromatogram of: A) crude mixture AvBD2 after Fmoc-SPPS (Jupiter C4 (300 Å, 5 µm, 250 × 10 mm), B) Purified AvBD2 after semi-preparative HPLC (Nucleosil C18 (300 Å, 5 µm, 250 × 10 mm) column) and C) Purified AvBD2 analyzed with more resolutive conditions (Aeris Widepore XB-C18 2 3.6 µm, 150 × 2.1 mm) column.

Compared to the standard HPLC purification, the *catch*-and-*release* purification of AvBD2 not only results in considerably higher yields (from 1.5% to 23%) but also allows to highly improve the purity of the peptide (56 % *vs* 85%) (Table 5).

Table 5. Comparison of the yield and purity of AvBD2 using the standard HPLC purification and the *catch*-and-*release* method. [a]: yield corresponding only to AvBD2[1-36] in the purified peptide, not taking impurities into account.

Purification method	Purity	Corrected overall yield[a]
HPLC	56	1.5%
Cacth-and-*release*	85	23%
Cacth-and-*release* + HPLC	98	3.3%

We can confirm that the *catch*-and-*release* approach is an excellent method for the purification of difficult cysteine-rich peptides such AvBD2.

5. Oxidative folding of immobilized AvBD2

5.1. Disulfide bond formation in peptide synthesis

The disulfide bonds are crucial for the biological activities of disulfide-rich peptides.[37,38] Their formation can be achieved by three main different strategies:

a) Air mediated oxidation: Thiolates spontaneously get oxidized into disulfides in the presence of oxygen. Thus, simple incubation of a cysteine-rich peptide at neutral or slightly basic pH leads to the formation of disulfide bonds. However, in most cases the reaction cannot be controlled and results in multiple isomers with different connectivities between cysteines. In addition, oxygen can also lead to an irreversible oxidation of the thiols into sulfonic acids.[37]

b) Orthogonal cysteine protections: protecting each pairs of two cysteines involved in a disulfide bond with thiol protective groups removable under orthogonal conditions allows the stepwise formation of the disulfides, thus preventing the formation of isomers with wrong connectivities.[39] However, it is a lengthy process, limited by the numerous intermediate HPLC purifications needed.

c) Redox-controlled oxidative folding: The most common conditions make use of a combination of a large excess (100 equivalents) of the cysteine-containing tripeptide glutathione (GSH) and a smaller excess (10 equivalents) of its oxidized form (GSSG), at slightly basic pH. The oxidative folding proceeds through a series of intra- and intermolecular thiolate/disulfide exchanges between GSH, GSSG and the thiols and disulfide bond of the peptide. The equilibrium is progressively displaced towards the formation of the correct disulfide pairing.[40] This process is thermodynamically-controlled, the native compound usually being much more stable than its isomers with a unnatural connectivity (Figure 23).[41]

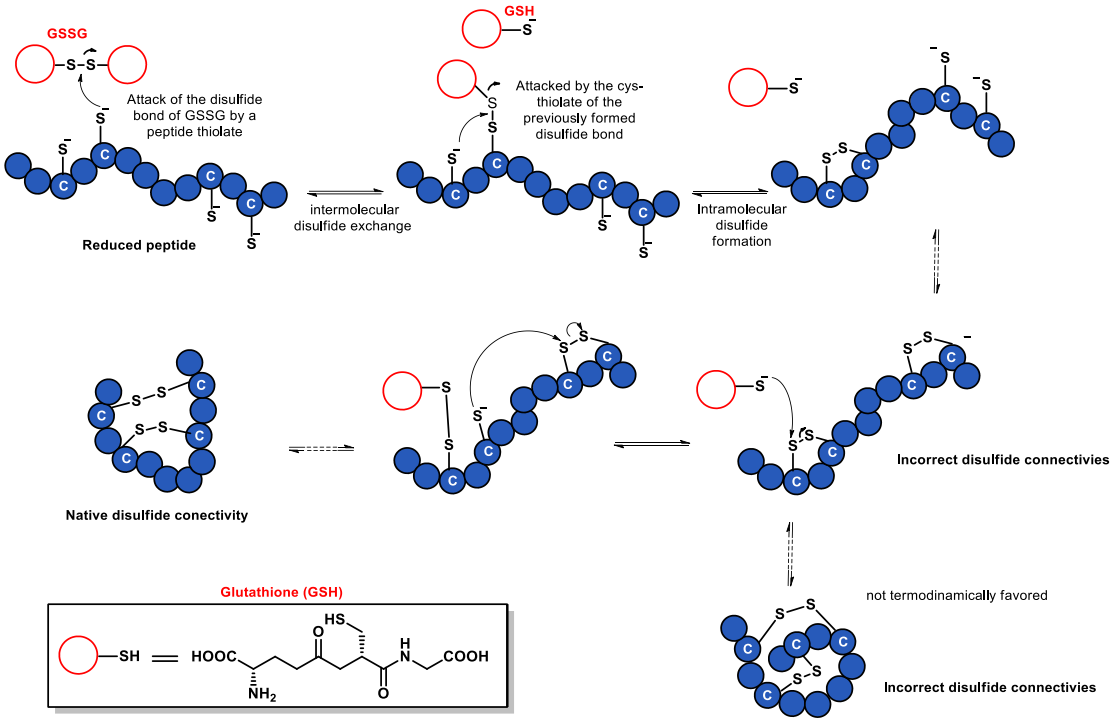


Figure 23 . Peptide oxidative folding mediated by the GSH/GSSG redox system. Image adapted from reference 40.[40]

5.2. Solid phase oxidative folding

The main limitations of the oxidative folding are associated with the highly diluted conditions (typically 10 µM) that are required to avoid the formation of intermolecular disulfide bonds. Very large reaction volumes are thus needed, complicating reaction work-up, and sometimes resulting in low recoveries due to material loss. Kent reported the oxidative folding of a small DRP, EETI-II (28 amino acids, three disulfide bond) immobilized on a water-compatible solid support,[42] exploiting the "pseudo-dilution effect"[43] which kinetically disfavor intermolecular reactions of solid-supported substrates. We decided to take advantage of the immobilization of AvBD2 to try to conduct its oxidative folding in solid phase, even if we were conscious that the cysteine contained in our linker could interfere with the process (Figure 24).

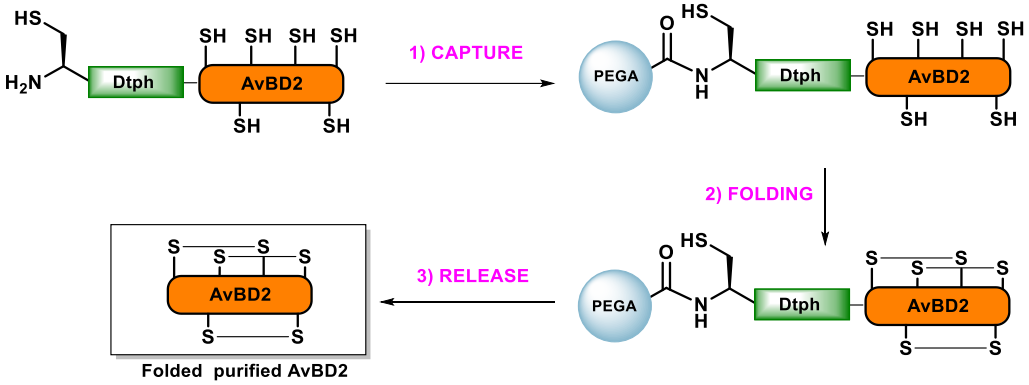


Figure 24. Proposed "*catch*-folding and-*release* purification" strategy. Acetylated peptides and other impurities are not represented here.

The process was done with the Cys-Dtph-containing peptide **24A** (*Tentagel, standard Fmoc-deprotections*) and the thioester-functionalized PEGA **21.** The immobilized peptide was incubated under classical redox couple-mediated oxidative folding conditions, followed by hydrazine-mediated cleavage. Disappointingly, we did not observe any release of the expected peptide.

We suspected the formation of a disulfide bond between the linker and the peptide (Figure 25). Treatment with an aqueous TCEP solution after the cleavage released the reduced form of AvBD2, constituting a strong evidence in favor of this hypothesis (Figure 26).

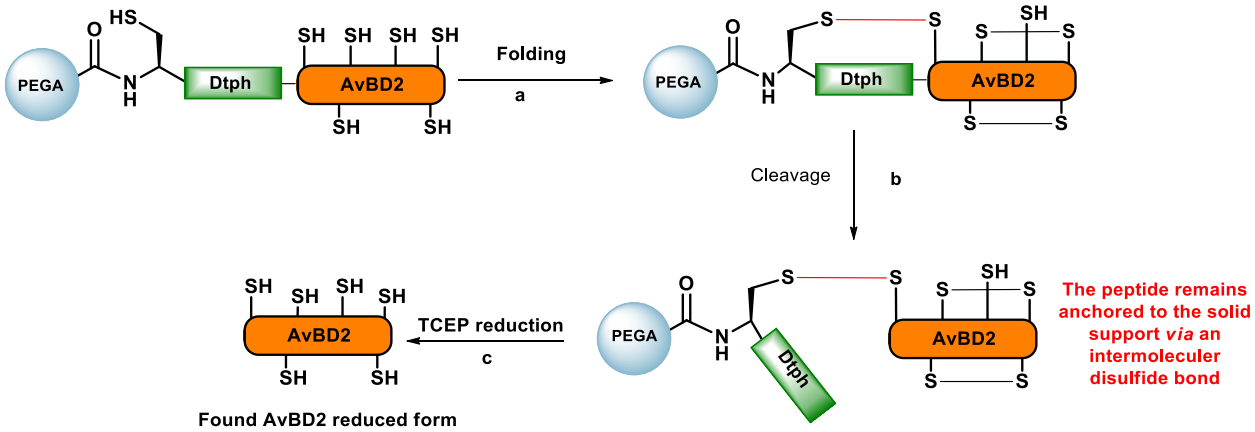


Figure 25. Failed attempt for AvBD2 solid-phase oxidative folding. Reaction conditions: a: peptide/GSH/GSSG (1:100:10), 1 mM EDTA, 100 mM TRIS, MeCN/H_2O (2:8), pH 8.5. b: 1 M N_2H_4, 100 mM TRIS, pH 8.5, c: 100 mM TCEP, HEPES, pH 7.

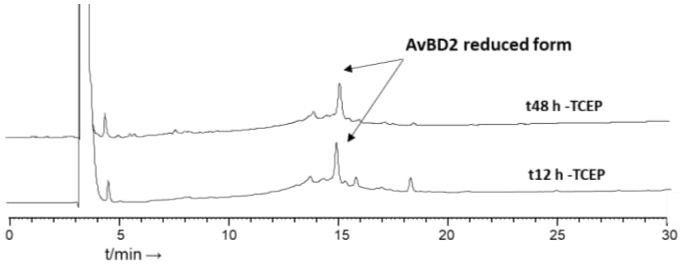


Figure 26. HPLC chromatogram of the released crude AvBD2 after TCEP reduction.

This method is not adapted to the solid phase oxidative folding of AvBD2.

6. Conclusions

We have synthesized by Fmoc-SPPS the reduced form of the avian defensin AvBD2, as a model for the application of a catch-and-release purification process using our Cys-Dtph linker. Identification of large amounts of deleted peptides prompted us to optimize the synthesis using different Fmoc deprotection conditions.

Application of optimized NCL-mediated immobilization and Dtph cleavage conditions to AvBD2 furnished a peptide of good purity (85%) in very high yield (23% corrected overall yield) without doing any HPLC purification. Comparatively, purification using standard semi-preparative HPLC gave dramatically lower yield (1.5%) and purity (56%), demonstrating the benefits of our approach and the validity of our strategy for the synthesis of very difficult cysteine-rich-peptides.

The non-chromatographic purification of AvBD2 constitutes the first example of the application of the *catch*-and-*release* purification to cysteine-rich peptides.

7. References chapter 4

(1) van Dijk, A.; Veldhuizen, E. J. A.; Haagsman, H. P. Avian Defensins. *Vet. Immunol. Immunopathol.* **2008**, *124* (1-2), 1.

(2) Ma, D.; Wang, R.; Liao, W.; Han, Z.; Liu, S. Identification and Characterization of a Novel Antibacterial Peptide, Avian β-Defensin 2 from Ducks. *J. Microbiol.* **2009**, *47* (5), 610.

(3) Sugiarto, H.; Yu, P. L. Avian Antimicrobial Peptides: The Defense Role of β-Defensins. *Biochem. Biophys. Res. Commun.* **2004**, *323* (3), 721.

(4) Yang, M.; Zhang, C.; Zhang, X.; Zhang, M. Z.; Rottinghaus, G. E.; Zhang, S. Structure-Function Analysis of Avian β-Defensin-6 and β-Defensin-12: Role of Charge and Disulfide Bridges. *BMC Microbiol.* **2016**, *16* (1), 210.

(5) Derache, C.; Labas, V.; Aucagne, V.; Meudal, H.; Landon, C.; Delmas, A. F.; Magallon, T.; Lalmanach, A.-C. Primary Structure and Antibacterial Activity of Chicken Bone Marrow-Derived Beta-Defensins. *Antimicrob. Agents Chemother.* **2009**, *53* (11), 4647.

(6) Derache, C.; Meudal, H.; Aucagne, V.; Mark, K. J.; Cadène, M.; Delmas, A. F.; Lalmanach, A.-C.; Landon, C. Initial Insights into Structure-Activity Relationships of Avian Defensins. *J. Biol. Chem.* **2012**, *287* (10), 7746.

(7) Kent, S. B. H. Chemical Synthesis of Peptides and Proteins. *Annu. Rev. Biochem.* **1988**, *57* (1), 957.

(8) Paradís-Bas, M.; Tulla-Puche, J.; Albericio, F. The Road to the Synthesis of 'difficult Peptides'. *Chem. Soc. Rev.* **2016**, *45* (3), 631.

(9) Milton, R. C. de L.; Milton, S. C. F.; Adams, P. A. Prediction of Difficult Sequences in Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* **1990**, *112* (16), 6039.

(10) Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F. Handles for Fmoc Solid-Phase Synthesis of Protected Peptides. *ACS Comb. Sci.* **2013**, *15* (5), 217.

(11) Sampson, W. R.; Patsiouras, H.; Ede, N. J. The Synthesis of 'difficult' Peptides Using 2-Hydroxy-4-Methoxybenzyl or Pseudoproline Amino Acid Building Blocks: A Comparative Study. *J. Pept. Sci.* **1999**, *5* (9), 403.

(12) Abdel-Aal, A.-B. 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.

(13) Weygand, F.; Steglich, W.; Bjarnason, J. Leicht Abspaltbare Schutzgruppen Für Die Säureamidfunktion, 3. Derivate Des Asparagins Und Glutamins Mit 2.4-Dimethoxy-Benzyl-Und 2.4.6-Trimethoxy-Benzyl-Geschützten Amidgruppen. *Chem. Ber.* **1968**, *101* (10), 3642.

(14) Zahariev, S.; Guarnaccia, C.; Zanuttin, F.; Pintar, A.; Esposito, G.; Maravi, G.; Krust, B.; Hovanessian, A. G.; Pongor, S. Efficient Synthesis and Comparative Studies of the

Arginine and Nω,Nω-Dimethylarginine Forms of the Human Nucleolin Glycine/arginine Rich Domain. *J. Pept. Sci.* **2005**, *11* (1), 17.

(15) Hyde, C.; Johnson, T.; Owen, D.; Quibell, M.; Sheppard, R. C. Some 'Difficult Sequences' Made Easy. A Study of Interchain Association in Solid-Phase Peptide Synthesis. *Int. J. Pept. Protein Res.* **1994**, *43* (5), 431.

(16) Meutermans, W. D. F.; Golding, S. W.; Bourne, G. T.; Les P., M.; Dooley, M. J.; Alewood, P. F.; Smythe, M. L. Synthesis of Difficult Cyclic Peptides by Inclusion of a Novel Photolabile Auxiliary in a Ring Contraction Strategy. *J. Am. Chem. Soc.* **1999**, *121*, 9790.

(17) Haack, T.; Mutter, M. Serine Derived Oxazolidines as Secondary Structure Disrupting, Solubilizing Building Blocks in Peptide Synthesis. *Tetrahedron Lett.* **1992**, *33* (12), 1589.

(18) Mutter, M.; Oppliger, H.; Zier, A. Solubilizing Protecting Groups in Peptide Synthesis. Effect of Side-chain-attached Poly(ethylene Glycol) Derivatives upon B-sheet Formation of Model Peptides. *Die Makromol. Chemie, Rapid Commun.* **1992**, *13* (3), 151.

(19) Mutter, M.; Wohr, T.; Gioria, S.; Keller, M. Pseudo-Prolines: Induction of Cis/trans-Conformational Interconversion by Decreased Transition State Barriers. *Biopolymers* **1999**, *51* (2), 121.

(20) Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. Pseudo-Prolines as a Solubilizing, Structure-Disrupting Protection Technique in Peptide Synthesis. *J. Am. Chem. Soc.* **1996**, *118* (39), 9218.

(21) Cremer, G.-A.; Tariq, H.; Delmas, A. F. Combining a Polar Resin and a Pseudo-Proline to Optimize the Solid-Phase Synthesis of a 'difficult Sequence'. *J. Pept. Sci.* **2006**, *12* (6), 437.

(22) García-Martín, F.; Quintanar-Audelo, M.; García-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Côté, S.; Tulla-Puche, J.; Albericio, F. ChemMatrix, a Poly(ethylene Glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J. Comb. Chem.* **2006**, *8* (2), 213.

(23) Tiefenbrunn, T. K.; Blanco-Canosa, J.; Dawson, P. E. Alternative Chemistries for the Synthesis of Thrombospondin-1 Type 1 Repeats. *Biopolymers* **2010**, *94* (4), 405.

(24) Bayer, E.; Hemmasi, B.; , Albert, K.; Rapp, W.; Dengler, M. I. Immobilized Polyoxyethylene, A New Support for Peptide Synthesis. *Pept. Struct. Funct. Proc. 8th Am. Pept. Symp. Hruby, V. J., Rich, D. H., Eds.; Pierce Chem. Co. Rockford, IL.,* **1983**, 87.

(25) García-Martín, F.; White, P.; Steinauer, R.; Côté, S.; Tulla-Puche, J.; Albericio, F. The Synergy of ChemMatrix Resin® and Pseudoproline Building Blocks Renders Rantes, a Complex Aggregated Chemokine. *Biopolymers* **2006**, *84* (6), 566.

(26) Camperi, S. A.; Marani, M.; Iannucci Nancy B; Côté, S.; Albericio, F.; Cascone, O. An Efficient Strategy for the Preparation of One-Bead-One-Peptide Libraries on a New

Biocompatible Solid Support. *Tetrahedron Lett.* **2005**, *46* (9), 1561.

(27) Http://spheritech.com/spheritide-Ks/.

(28) Wade, J. D.; Bedford, J.; Sheppard, R. C.; Tregear, G. W. DBU as an N Alpha-Deprotecting Reagent for the Fluorenylmethoxycarbonyl Group in Continuous Flow Solid-Phase Peptide Synthesis. *Pept. Res.* **1991**, *4* (3), 194.

(29) Tickler, A. K.; Barrow, C. J.; Wade, J. D. Improved Preparation of Amyloid-β Peptides Using DBU as Nα-Fmoc Deprotection Reagent. *J. Pept. Sci.* **2001**, *7* (9), 488.

(30) Sheppeck, J. E.; Kar, H.; Hong, H. A Convenient and Scaleable Procedure for Removing the Fmoc Group in Solution. *Tetrahedron Lett.* **2000**, *41* (28), 5329.

(31) King, D. S.; Fields, C. G.; Fields, G. B. A Cleavage Method Which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis. *Int. J. Pept. Protein Res.* **2009**, *36* (3), 255.

(32) Lundt, B. F.; Johansen, N. L.; Markussen, J. Formation and Synthesis of 3-Tert-Butyltyrosine. *Int. J. Pept. Protein Res.* **2009**, *14* (4), 344.

(33) Blanco-Canosa, 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.

(34) Gunasekera, S.; Aboye, T. L.; Madian, W. A.; El-Seedi, H. R.; Göransson, U. Making Ends Meet: Microwave-Accelerated Synthesis of Cyclic and Disulfide Rich Proteins Via In Situ Thioesterification and Native Chemical Ligation. *Int. J. Pept. Res. Ther.* **2013**, *19* (1), 43.

(35) Harris, P. W.; Kowalczyk, R.; Yang, S. H.; Williams, G. M.; Brimble, M. A. An Important Side Reaction Using the Thiol, 3,6-Dioxa-1,8-Octanedithiol (DODT), in 9-Fluorenylmethoxycarbonyl-Based Solid Phase Peptide Synthesis. *J. Pept. Sci.* **2014**, *20* (3), 186.

(36) Teixeira, A.; Benckhuijsen, W.; de Koning, P.; Valentijn, A.; Drijfhout, J. The Use of Dodt as a Non-Malodorous Scavenger in Fmoc-Based Peptide Synthesis. *Protein Pept. Lett.* **2002**, *9* (5), 379.

(37) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014**, *114* (2), 901.

(38) Irbäck, A.; Mohanty, S. Folding Thermodynamics of Peptides. *Biophys. J.* **2005**, *88* (3), 1560.

(39) Postma, T. M.; Albericio, F. Disulfide Formation Strategies in Peptide Synthesis. *European J. Org. Chem.* **2014**, *2014* (17), 3519.

(40) Okumura, M.; Shimamoto, S.; Hidaka, Y. A Chemical Method for Investigating Disulfide-Coupled Peptide and Protein Folding. *FEBS J.* **2012**, *279* (13), 2283.

(41) Tu, B. P.; Weissman, J. S. Oxidative Protein Folding in Eukaryotes: Mechanisms and

Consequences. *J. Cell Biol.* **2004**, *164* (3), 341.

(42) Johnson, E. C. B.; Durek, T.; Kent, S. B. H. Total Chemical Synthesis, Folding, and Assay of a Small Protein on a Water-Compatible Solid Support. *Angew. Chem. Int. Ed.* **2006**, *45* (20), 3283.

(43) Jayalekshmy, P.; Mazur, S. Pseudodilution, the Solid-Phase Immobilization of Benzyne. *J. Am. Chem. Soc.* **1976**, *98* (21), 6710.

Chapter 5

Chapter 5. Towards the application of the *catch*-and-*release* strategy to very long DRPs

1. Selection of a set of possible targets

In order to evaluate the applicability of our method for longer DRPs, we selected four biologically-relevant DRPs as possible targets for *catch*-and *release* purification. If the NCL-based syntheses of all these compounds have been reported, none of them was ever prepared by single-fragment SPPS, due to their length (65 to 77 amino acids).

1.1. Scorpion α-toxin OD1

OD1 was isolated from the venom of the Iranian yellow scorpion *Odonthobuthus doriae*.[1] It belongs to the alpha-like toxins family. Its sequence contains 65 amino acids: GVRDAYIADDKNCVYTCASNGYCNTECTKNGAESGYCQWIGRYGNACWCIKLPDEVPIRIPGKCR. The peptide is stabilized by four disulfide bonds between cysteines C13-C64, C17-C37, C23-C47 and C27-C49. Recently, Durek *et al.* proposed a strategy for its chemical synthesis which involved the solid phase synthesis of three segments (around 20 amino acids-long) and their assembly *via* one-pot NCL. The folded peptide was used to study in detail the crystal structure and the pharmacological properties of OD1 (Figure 1).[2]

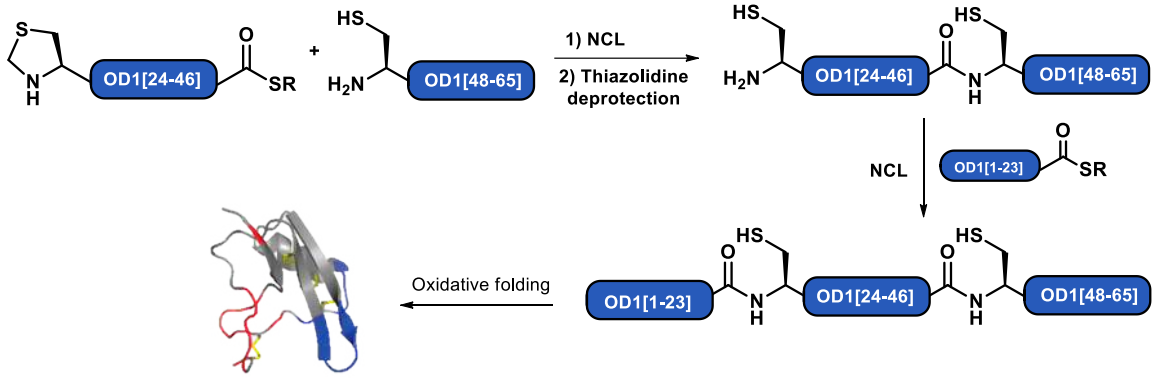


Figure 1. One pot NCL and folding of the OD1 peptide, proposed by Durek et al.[2]

1.2. Spermaurin

Spermaurin has been also isolated from the venom of a scorpion (*Scorpio maurus palmatus)*. It belongs to the La1-like DRP family. Its sequence is constituted by 73 amino acids including 8 cysteines engaged in four disulfide bonds: SGESCKAGKFTVLVGHQQNDKSTCTLYKCLNFKRKY ALQISSCADVDLKSGCRQVPGAASANFPECCPMICK-NH_2. It was identified by Arnoult´s team for its ability to enhance sperm motility in mammals, including human.[3] Our research group realized its synthesis *via* NCL of two fragments, including a C-terminal *N*-Hnb-Cys crypto-thioester which is converted into a thioester under NCL conditions (Figure 2).[4] The CBM NMR group recently solved its 3D structure (unpublished results).

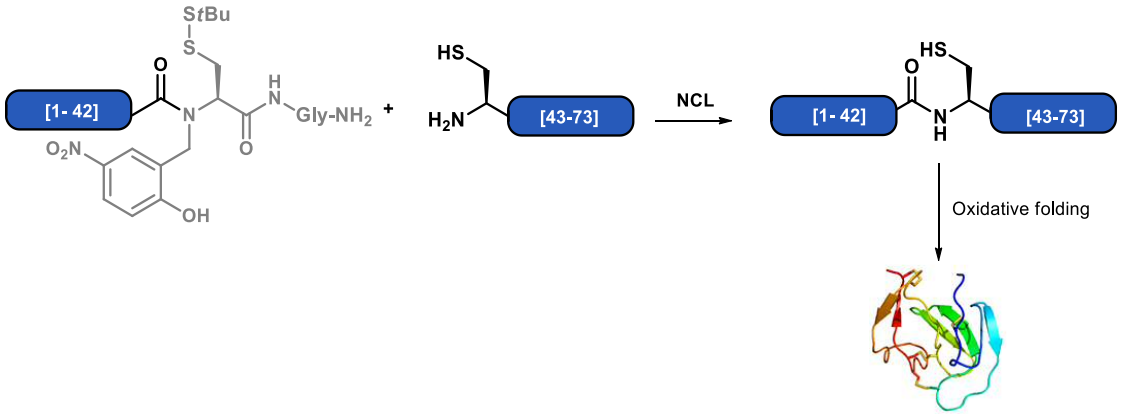


Figure 2. NCL-based synthesis of Spermaurin.

1.3. Bowman-Birk Inhibitors (BBI)

BBI is a dual α-chymotrypsin and trypsin inhibitor isolated from soybeans.[5,6] This peptide contains 71 amino acids (DDESSKPCCDQCACTKSNPPQCRCSDMRLNSCHSACKSCICALSYPAQC FCVDITDFCYEPCKPSEDDKE) including 14 cysteines forming seven disulfide bonds. Tornøe and co-workers recently published its synthesis through the assembly of three segments *via* NCL (Figure 3).[7]

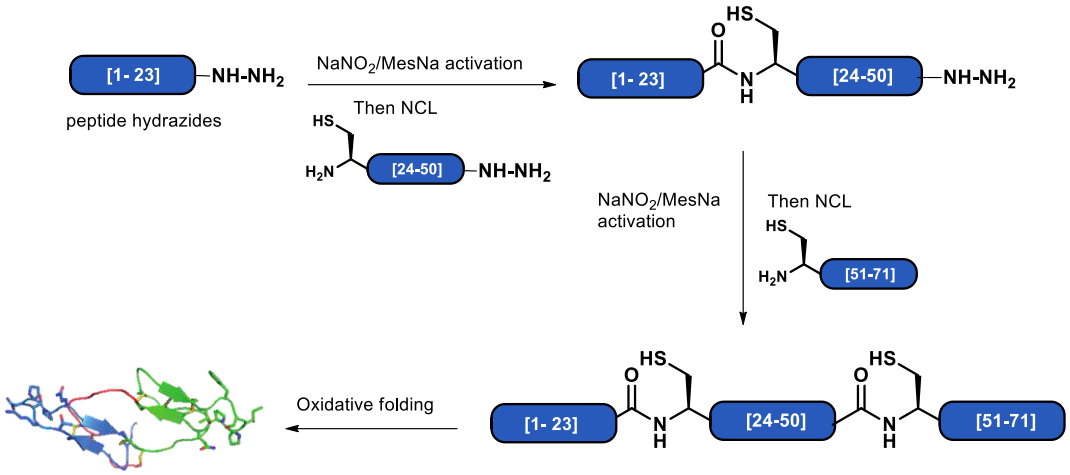


Figure 3. NCL-based synthesis of BBI. [7]

1.4. Bv8 peptide

Bv8 has been isolated from skin secretions of the toad *Bombina variegate* and belongs to the AVIT toxin family. It is constituted by 77 amino acids (AVITGACDKDVQCGSGTCCAASAWSRNIRFCIPLGNSGEDCHPASHKVPYDGKRL SSLCPCKSGLTCSKSGEKFKCS), including ten cysteines which form five disulfide bonds. This DRP is homologous to the mammalian prokineticin (mBv8), and is able to modulate different biological processes such as nociception, neurogenesis or angiogenesis.[8,9] Durek *et al.* reported its synthesis through the NCL-based assembly of three fragments (Figure 4).[8]

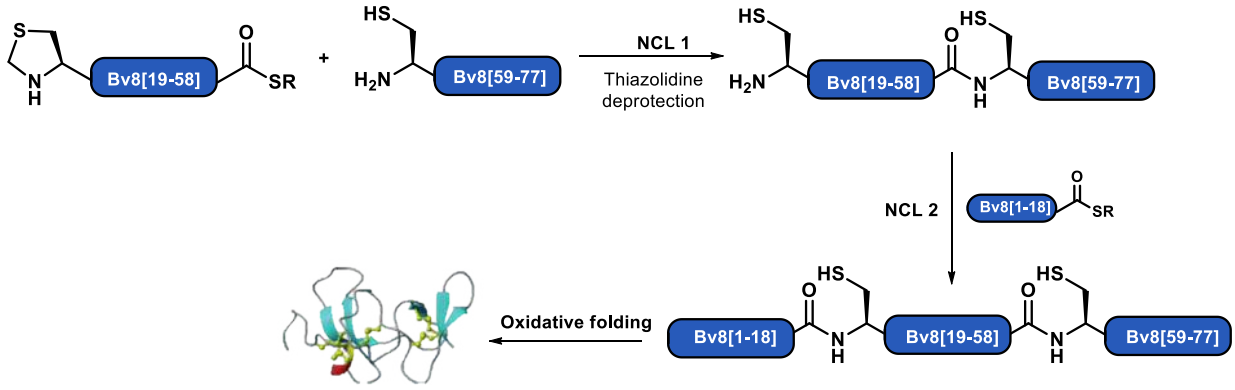


Figure 4. Strategy proposed by Durek et al. for the synthesis of Bv8.[8]

2. Solid phase synthesis of the four DRPs, and selection of a suitable target for *catch*-and-*release* purification

We first evaluated which peptide(s) would constitute the best candidate(s) for the exemplification of our *catch*-and-*release* purification method. Two factors were considered: the difficulty to purify the targets by standard HPLC, and the absence of significant byproducts able to be tagged by our linker (*i.e.* other than acetylated truncated peptides).

Similarly to the previous chapter, we performed the solid phase synthesis of the four sequences lacking their N-terminal amino acid. As we did not have in hand all four linker-functionalized N-terminal amino acids (*ie* Gly, Ser, Asp and Ala) at the time when we performed this preliminary assays, we decided to couple **4b** (Boc-Cys(Trt)-Dpth-Val-OH). Mutant sequences were thus obtained (G1V OD1, S1V spermaurin, D1V BBI and A1V Bv8).

Aliquots of the resulting peptidyl resins were cleaved (Figure 5) and the crude peptides were analyzed by LC-MS and HPLC. The chromatograms obtained for each peptide are analyzed in detail hereafter.

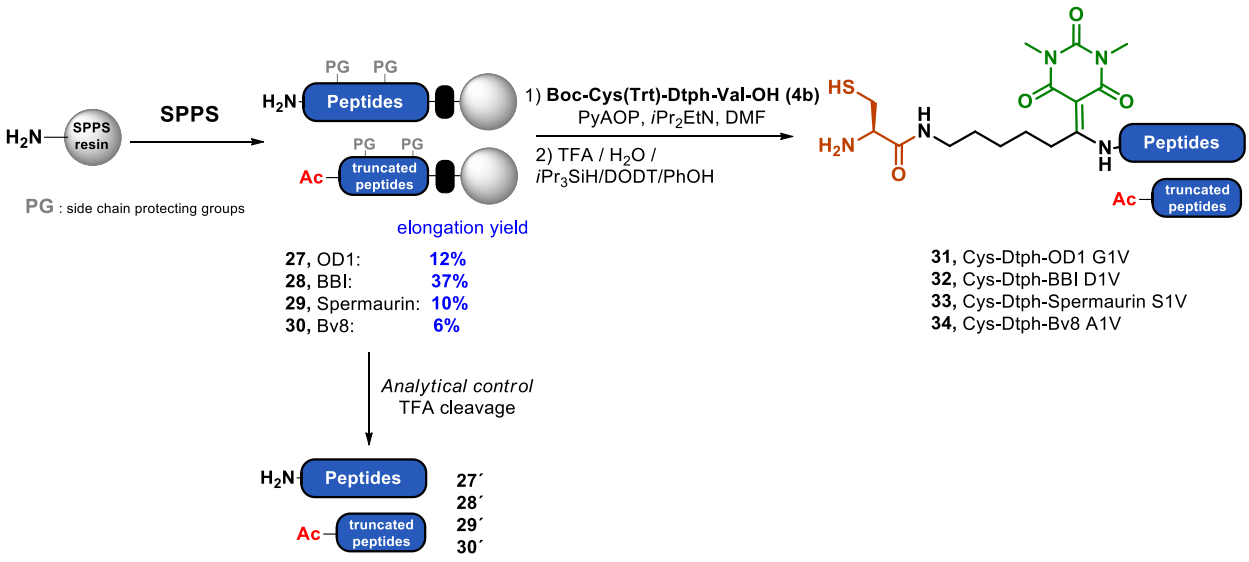


Figure 5. Synthesis of peptidyl resins 27-30 and Cys-Dtph-mutant sequences 31-34.

a) OD1

HPLC analysis of the crude mixture obtained after SPPS showed a major peak which mass was consistent with the target OD1[2-65], **27´** (Figure 6). We believe that its HPLC purification after SPPS should not constitute a big challenge. In addition, the introduction of

our linker did not give a convincing HPLC profile, even if we identified the Cys-Dtph-containing OD1[1-65] Gly1Val mutant. Consequently, this peptide has not been considered as a suitable example for applying our *catch*-and-*release* strategy.

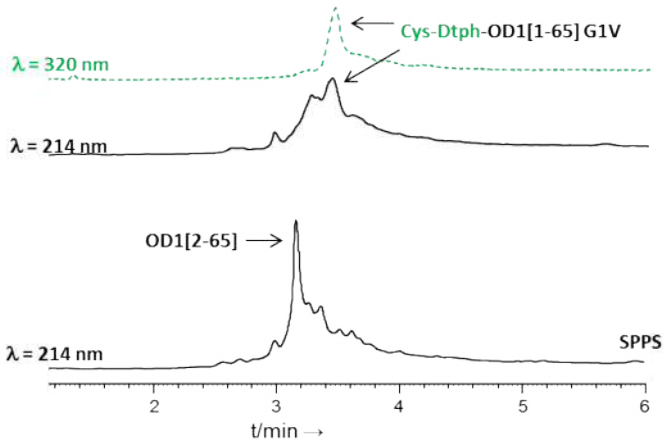


Figure 6. HPLC chromatograms of 27´ and 31.

b) BBI

The solid phase synthesis of BBI resulted in the highest elongation yield: 37 %. Similarly to OD1, the major component of the crude mixture after SPPS was the target peptide BBI[2-71] (Figure 7). Even if the HPLC profile of the Cys-Dtph-containing Asp1Val mutant was more acceptable than the one obtained for OD1, BBI was not considered as a highly challenging target, and we did not selected it for the application of our strategy.

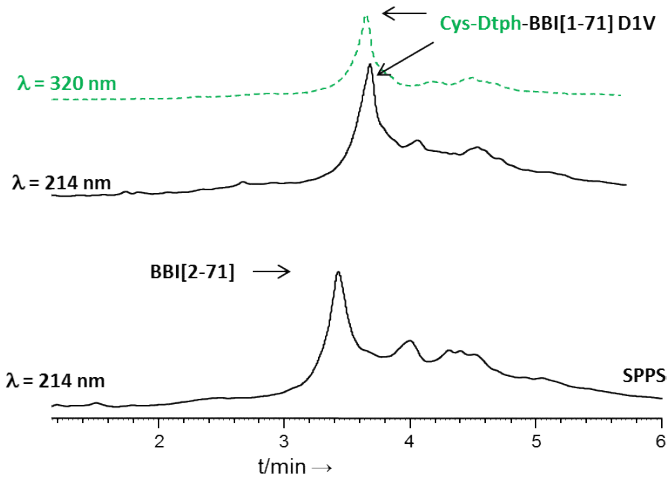


Figure 7. HPLC chromatograms of 28´ and 32.

c) Spermaurin

The crude mixture after SPPS presented a large number of acetylated peptides, some of them co-eluting with the target peptide which was a minor component (Figure 8). After the introduction of our linker, we observed the apparition of one new peak which mass was consistent with the expected Cys-Dtph-containing Ser1Val spermaurin mutant **33**. The chromatogram observed at λ = 320 nm was rather clean, indicating that no significant by-product able to incorporate our linker was present. As we believe that standard HPLC-purification of Spermaurin could not give satisfactory results in terms of purity, it could represent a good candidate for applying our *catch*-and-*release* strategy.

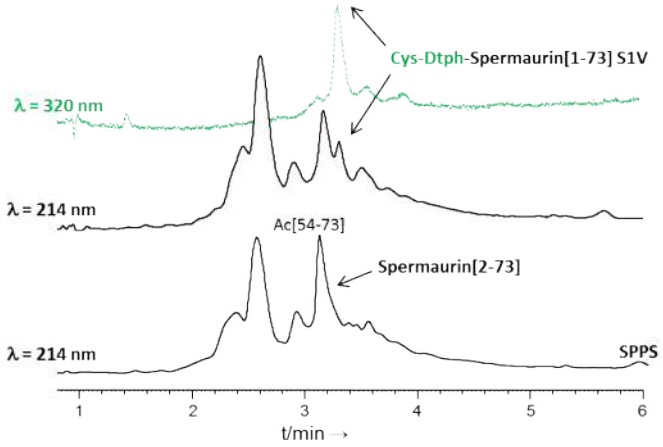


Figure 8. HPLC chromatograms of 29´ and 33.

d) Bv8

Bv8 is the longest peptide of our series (77 amino acids). Its solid phase synthesis resulted in the lowest elongation yield: 6%. Similarly to Spermaurin, the crude peptide after SPPS presented a large number of acetylated peptides and could not be satisfyingly purified by RP-HPLC (Figure 9), and the HPLC profile at 320 nm of the Cys-Dtph-Bv8 [1-77] Ala1Val mutant **34** was rather convincing.

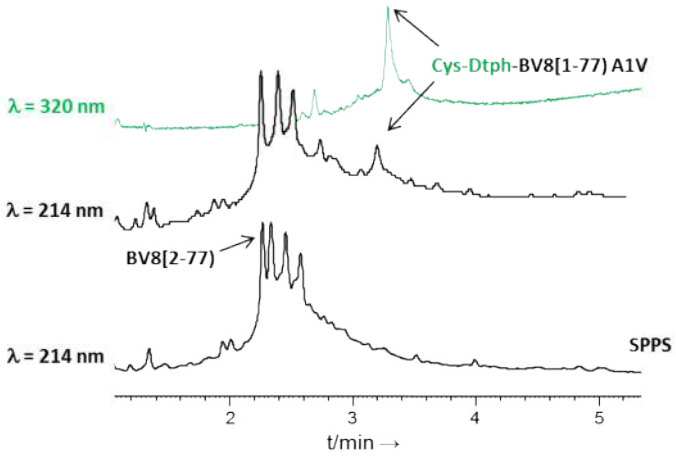


Figure 9. HPLC chromatograms of 30´ and 34.

Both spermaurin and Bv8 peptides seem to constitute good candidates for the application of our strategy. In this manuscript, we will present our preliminary study of the *catch*-and-*release* purification of Bv8.

3. *Catch*-and-*release* purification of Bv8

3.1. Synthesis of Dtph-containing peptides

3.1.1. Synthesis in detail of peptidyl resin **30**

Bv8 [2-77] was synthesized by automated Fmoc-SPPS on Tentagel resin using a methylphenoxypropionic acid linker (MPPA) generating a C-terminal carboxylic acid peptide upon TFA-mediated cleavage of the peptidyl resin **30** (See details in the experimental part).

Bv8 contains an Asp-Gly sequence (Asp50-Gly51) which can be associated with the formation of aspartimide,[10] a well-known SPPS side reaction (Figure 10). During the elongation, the nitrogen atom of the neighboring peptide bond can attack the carbonyl of the ester-protected Asp side chain, the reaction being particularly favored with a Gly residue. A five-membered α-amino succinimide ring is formed, referred as aspartimide.[11] The subsequent Fmoc deprotection treatments can lead to the opening of the ring by piperidine to afford the piperidide peptide.[11,12]

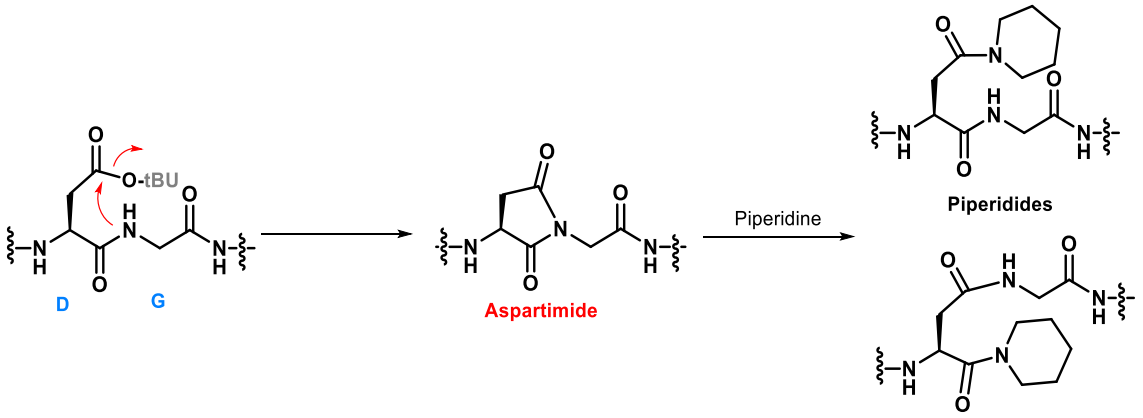


Figure 10. Formation of Aspartimide side products during the Fmoc-removal.

In order to limit this potential side reaction, we decided to protect the Asp side chain with a 3-ethyl-3-pentyl (Epe) ester instead of the standard *tert*-butyl ester (Figure 11). The steric hindrance of Epe has been indeed reported to prevent the formation of aspartimide.[13]

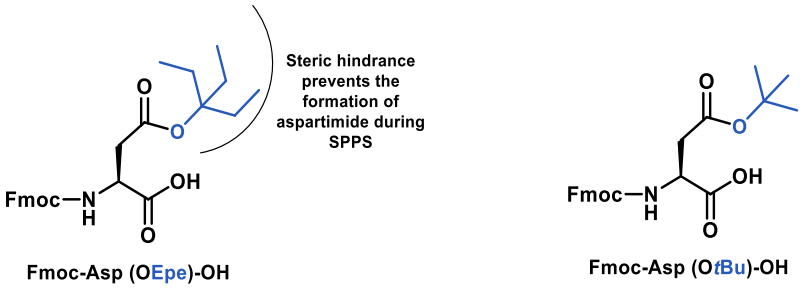


Figure 11. Structure of the Fmoc-Asp (OEpe)-OH AA used in SPPS of BV8 peptide.

Detailed LC-MS analysis of the crude mixture Bv8[2-77] **30´** revealed the presence of minor peaks consistent with aspartimide and piperidides (Figure 12). If these very low amounts of side products seem perfectly acceptable, we also observed a large peak showing a 56 Da mass increase, consistent with a *tert*-butylated cysteine.[14–17]

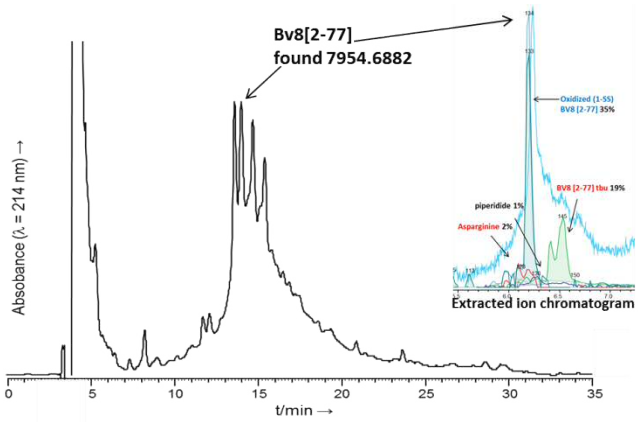


Figure 12. HPLC chromatogram and extracted ion chromatogram (EIC) of 30´.

Note that, around 35 % of the target peptide appeared under an oxidized form (one disulfide bond). TCEP was added to the crude peptide to try to reduce this disulfide bond, however it was not possible to reduce it completely. This illustrates the high tendency of this DRP to undergo spontaneous air oxidation even under acidic conditions, which is quite unusual.

3.2. Synthesis of the Cys-Dtph-containing peptide 35

Following the same protocol used for the other peptides, Boc-Cys(Trt)-Dtph-Ala-OH linker, **4d**, was coupled to the supported peptide **30** to yield the Cys-Dtph-peptide **35** (note that the only difference between **34** and **35** is the N-terminal amino acid : Val in **34** and Ala in **35**).

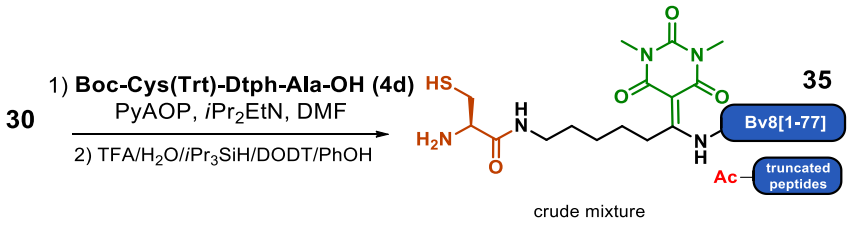


Figure 13. Synthesis of the Cys-Dtph-peptide 35.

LC-MS analysis confirmed a quantitative coupling (Figure 14).

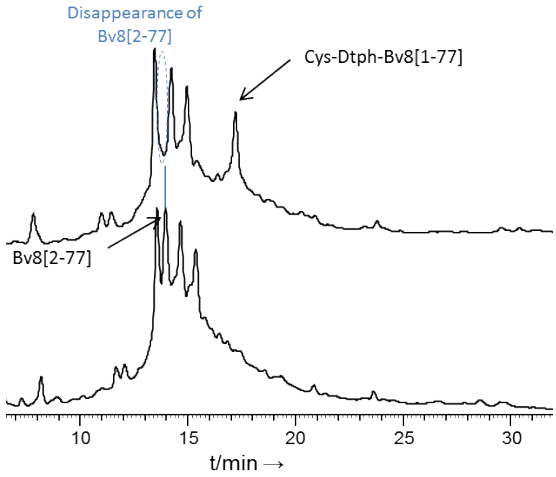


Figure 14. HPLC analyses at 240 nm corresponding to the introduction of the Cys-Dtph linker.

In order to try to limit the amout of the +*t*Bu byproduct we tried to substitute DODT by EDT in the TFA cleavage coktail. However, similarly to what occurred with AvBD2 (Chapter 4), we did not observe any difference in the HPLC profiles and thus decided to continue using DODT.

3.3. *Catch*-and-*release* purification to BV8 peptide

Catch-and-*release* purification was performed using the same conditions as for AvBD2 (Chapter 4) (Figure 15).

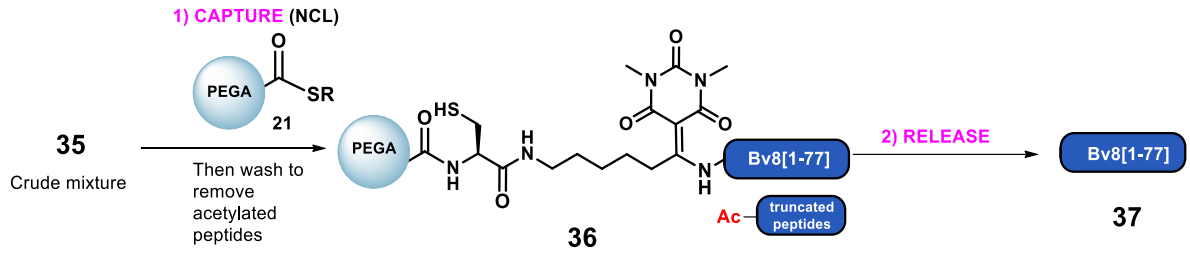


Figure 15. *Catch*-and-*release* purification of Bv8.

NCL-based immobilization proceeded without problem, and the reaction was complete after 14 h as judged from the HPLC analysis of the supernatant (Figure 16). Hydrazine-mediated cleavage of Dtph led to the expected purified Bv8[1-77], albeit in much lower quantities than expected.

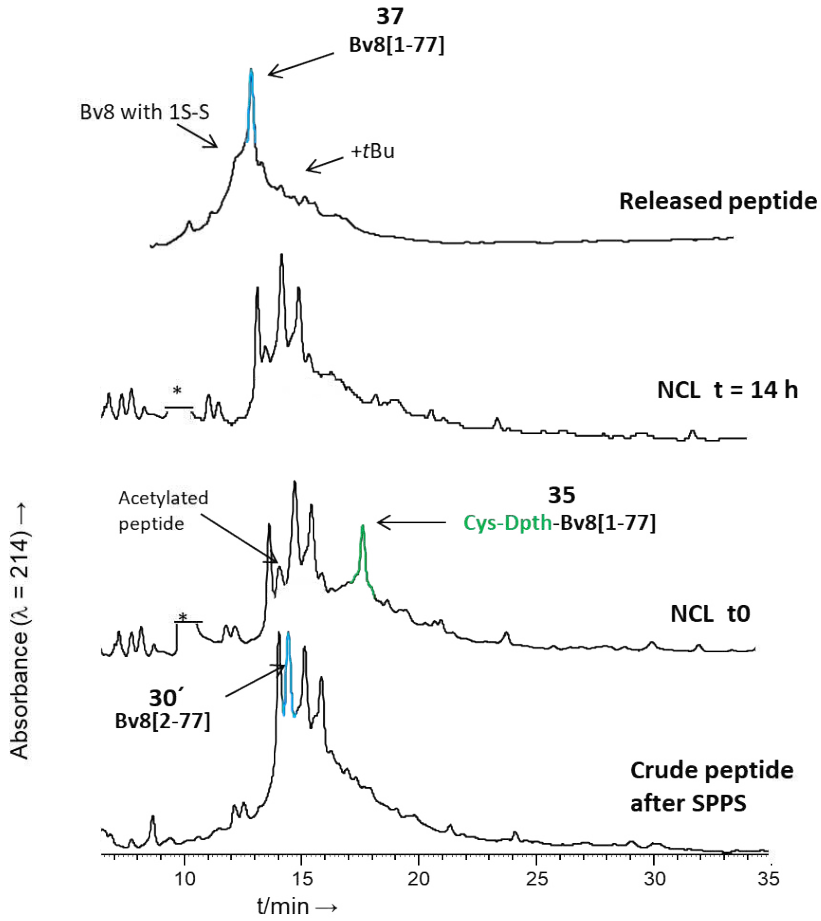


Figure 16. HPLC profiles corresponding to *catch*-and-*release* purification of Bv8. *MPPA traces.

The amount of peptide after the release was determined by UV spectrophotometry (ε^{280} = 6780 M^{-1} cm^{-1}). Comparison with the starting amount of Cys-Dtph-containing peptide (determined at 320 nm) allowed estimating the yield of the *catch*-and-*release* process to only 1.8 %.

Despite all our efforts, we were never able to increase this very deceptive yield that we attributed to the aggregation-prone nature of the reduced form of Bv8.

Purity of the released compound was estimated to 69% by integration of the chromatogram at 280 nm. The corrected overall yield of the entire process: Fmoc SPPS, Boc-Cys(Trt)-Dtph linker coupling, TFA cleavage, catch and release was only 0.05%. This corresponds to 100 µg peptide obtained from a 25 µmol scale synthesis, and thus cannot be seen as a practical option for the preparation of Bv8.

In order to obtain a very pure peptide, we carried out a polishing HPLC which allows increasing the purity of Bv8 in more than 86% (Figure 17).

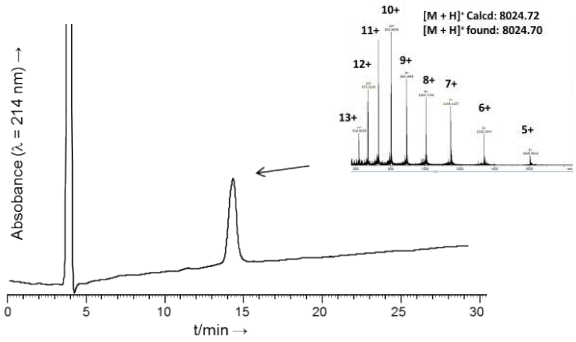


Figure 17. HPLC purification and ESI-HRMS analysis of pure Bv8. Retention time = 14.7 min

4. Conclusions

We have showed that the *catch*-and-*release* approach can be also applied to long DRPs, such as the BV8 peptide (77 amino acids). However we obtained a very low yield compared to the yield obtained for the *catch*-and-*release* purification of AvBD2 (23% overall yield).

The 5 times lowest elongation yield between Bv8 and AvBD2 (6% *vs* 30%) cannot explain the dramatically differences in the yield after the *catch*-and-*release* purification. We suppose that the low yield obtained for Bv8 can be due to the aggregation-prone nature of the reduced form of this peptide.

This work only constitutes a preliminary study, and further optimization will be needed to obtain better yields and purities.

5. References chapter 5

(1) Jalali, A.; Bosmans, F.; Amininasab, M.; Clynen, E.; Cuypers, E.; Zaremirakabadi, A.; Sarbolouki, M.-N.; Schoofs, L.; Vatanpour, H.; Tytgat, J. OD1, the First Toxin Isolated from the Venom of the Scorpion Odonthobuthus Doriae Active on Voltage-Gated Na+ Channels. *FEBS Lett.* **2005**, *579* (19), 4181.

(2) Durek, T.; Vetter, I.; Wang, C.-I. A.; Motin, L.; Knapp, O.; Adams, D. J.; Lewis, R. J.; Alewood, P. F. Chemical Engineering and Structural and Pharmacological Characterization of the α-Scorpion Toxin OD1. *ACS Chem. Biol.* **2013**, *8* (6), 1215.

(3) Martinez, G.; Hograindleur, J.-P.; Voisin, S.; Abi Nahed, R.; Abd El Aziz, T. M.; Escoffier, J.; Bessonnat, J.; Fovet, C.-M.; De Waard, M.; Hennebicq, S.; et al. Spermaurin, an La1-like Peptide from the Venom of the Scorpion Scorpio Maurus Palmatus , Improves Sperm Motility and Fertilization in Different Mammalian Species. *Mol. Hum. Reprod.* **2016**, *23* (2), 116.

(4) 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.

(5) Birk, Y. The Bowman-Birk Inhibitor. Trypsin- and Chymotrypsin-Inhibitor from Soybeans. *Int. J. Pept. Protein Res.* **1985**, *25* (2), 113.

(6) Song, G.; Zhou, M.; Chen, W.; Chen, T.; Walker, B.; Shaw, C. HV-BBI. A Novel Amphibian Skin Bowman.Birk-like Trypsin Inhibitor. *Biochem. Biophys. Res. Commun.* **2008**, *372* (1), 191.

(7) Tornøe, C.; Johansson, E.; Wahlund, P. O. Divergent Protein Synthesis of Bowman-Birk Protease Inhibitors, Their Hydrodynamic Behavior and Co-Crystallization with α-Chymotrypsin. *Synlett.* **2017**, *28* (15), 1901.

(8) Morales, R. A. V.; Daly, N. L.; Vetter, I.; Mobli, M.; Napier, I. A.; Craik, D. J.; Lewis, R. J.; Christie, M. J.; Alewood, P. F.; Durek, T. Chemical Synthesis and Structure of the Prokineticin Bv8. *Chem. Bio. Chem.* **2010**, *11* (13), 1882.

(9) Mollay, C.; Wechselberger, C.; Mignogna, G.; Negri, L.; Melchiorri, P.; Donatell, B.; Kreil, G. Bv8, a Small Protein from Frog Skin and Its Homologue from Snake Venom Induce Hyperalgesia in Rats. *Eur. J. Pharmacol.* **1999**, *374* (2), 189.

(10) Tam, J. P.; Riemen, M. W.; Merrifield, R. B. Mechanisms of Aspartimide Formation: The Effects of Protecting Groups, Acid, Base, Temperature and Time. *Pept. Res.* **1988**, *1* (1), 6.

(11) 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.

(12) 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.

(13) Behrendt, R.; Huber, S.; White, P. Preventing Aspartimide Formation in Fmoc SPPS of Asp-Gly Containing Peptides - Practical Aspects of New Trialkylcarbinol Based Protecting Groups. *J. Pept. Sci.* **2016**, *22* (2), 92.

(14) King, D. S.; Fields, C. G.; Fields, G. B. A Cleavage Method Which Minimizes Side Reactions Following Fmoc Solid Phase Peptide Synthesis. *Int. J. Pept. Protein Res.* **2009**, *36* (3), 255.

(15) Lundt, B. F.; Johansen, N. L.; Markussen, J. Formation and Synthesis of 3-Tert-Butyltyrosine. *Int. J. Pept. Protein Res.* **2009**, *14* (4), 344.

(16) Blanco-Canosa, 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.

(17) Gunasekera, S.; Aboye, T. L.; Madian, W. A.; El-Seedi, H. R.; Göransson, U. Making Ends Meet: Microwave-Accelerated Synthesis of Cyclic and Disulfide Rich Proteins Via In Situ Thioesterification and Native Chemical Ligation. *Int. J. Pept. Res. Ther.* **2013**, *19* (1), 43.

Chapter 6

Chapter 6. Experimental section

Table of contents

1. General information

All reagents and solvents were used without further purification. Protected amino acids, Rink´s linker, Spheritide™ resin, amino-PEGA resin, HBTU and HCTU were purchased from Merck Biosciences (Nottingham, UK). Controlled pore glass beads (TRISOPERL®, pore size: 100 nm, bead diameter: 100-200 µm, lot number: PG L 14/04 AMINO 225) was obtained from VitraBio (Steinach, Germany), Aminomethyl TentaGel R resin was purchased from Rapp polymers (Tuebingen, Germany), H-Rink Amide-ChemMatrix® was obtained from pcas. BioMatrix. Inc (Quebec, Canada). [*N*-(9-fluorenylmethoxycarbonyl)-1,13-diamino-4,7,10-trioxatridecan-succinamic acid (Fmoc-TTDS-OH) and Fmoc-NH-PEG-COOH (3000 Da) were purchased from Iris Biotech GmbH (Marktredwitz, Germany). Fmoc-Ala-methylphenoxypropionic acid (Fmoc-L-Ala-MPAA-OH was obtained from Interchim (Paris, France). PyAOP was obtained from aapptec (France, Europe). Peptide synthesis grade DMF and HATU were obtained from Applied Biosystems (Courtaboeuf, France). Ultrapure water was obtained using a Milli-Q water system from Millipore (Molsheim, France). All other chemicals were from Sigma Aldrich (St-Quentin-Fallavier, France) and solvents from SDS-Carlo Erba (Val de Reuil, France). Polypropylene syringes fitted with polypropylene frits were obtained from Torviq (Niles, MI, USA) and were equipped with PTFE stopcock bought from Chromoptic (Courtaboeuf, France).

High resolution ESI-MS analyses were performed on a MaXis™ ultra-high-resolution QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany), using the positive mode. MALDI TOF analyses were performed on an Ultraflex instrument (Bruker Daltonics, Bremen, Germany) equipped with a 337-nm nitrogen laser and a gridless delayed extraction ion source. The sample was co-crystallized with a solution of α-cyano-4-hydroxy-cinnamic acid (HCCA) as a matrix. The reported *m/z* values correspond to the monoisotopic ions if not specified otherwise. LC/MS analyses were performed on a 6120B single Quadrupole LC/MS system (Agilent Technologies, Les Ulis, France), using positive mode.

HPLC analyses were carried out on a Chromaster system equipped with a Hitachi 5160 pump, a Hitachi 5260 auto sampler and a Hitachi 5430 diode array detector. Chromolith HighResolution RP-18e (130 Å, 10 × 4.6 mm, 3 mL/min flow rate) and Jupiter C4 (300 Å, 5 µm, 250 × 10 mm, 3 mL/min flow rate) column was used for analysis. Semi-preparative

purifications were carried out on a LaChrom Elite system equipped with a Hitachi L-2130 pump, a Hitachi L-2455 diode array detector and a Hitachi L-2200 auto sampler. Nucleosil C18 (300 Å, 5 µm, 250 × 10 mm, 3 mL/min flow rate) or Jupiter C4 (300 Å, 5 µm, 250 × 10 mm, 3 mL/min flow rate) columns were used for purification. Chromatography was conducted at room temperature unless otherwise mentioned. Solvents A and B are 0.1% TFA in H_2O and 0.1% TFA in MeCN, respectively. Each gradient was followed by a washing step (up to 95% B/A) to elute any compound not eluted during the gradient. LC/HRMS and LC/MS analyses were carried out respectively on an Ultimate® 3000 RSLC HPLC system (Dionex, Germering, Germany), coupled with the MaXis™ mass spectrometer and on an Agilent 1260 Infinity HPLC system, coupled with the Agilent 6120 mass spectrometer, and fitted with an Aeris Widepore XB-C18 2 (3.6 µm, 150 × 2.1 mm, 0.5 mL/min flow rate, 40°C) column. Solvents A and B were 0.1% formic acid in H_2O and 0.08% formic acid in MeCN, respectively. Gradient: 3% B/A for 0.6 min, then 3 to 50% B/A over 10.8 min.

Microwave heated reactions were carried out on a Biotage® SP Wave.

All the NCL reactions and oxidative folding studies were performed under a strict argon atmosphere, using solvents that had been freshly deoxygenated through ten successive vacuum (15 mbar)/argon cycles.

Characterization and purification of the building blocks

Flash column chromatography purifications were carried out using Kiesegel C60 (Merck, Germany) as the stationary phase, and thin layer chromatography analyses were performed on precoated silica gel plates (0.25 mm thick, 60 F254, Merck, Germany) and observed under UV light at 254 nm then stained with a standard basic potassium permanganate solution.

Purifications by reverse phase column chromatography were carried out on the Interchim SPOT II Flash, equipped with CHROMABOND® Flash RS 40, 43 g C18 ec (flow rate: 40 mL/min).

^{1}H and ^{13}C NMR spectra were recorded on a Bruker AVANCE III 600 instrument, at a constant temperature of 25°C and on a Bruker AVANCE III 400 instrument at a temperature of 55°C. Chemical shifts are reported in parts per million from low to high field and referenced to tetramethylsilane (TMS). Coupling constants (*J*) are reported in hertz (Hz). Standard

abbreviations indicating multiplicity were used as follows: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, m = multiplet, b= broad signal.

2. General procedures

a. General procedure A for automated solid phase peptide synthesis

Fmoc-based solid phase peptide synthesis (SPPS) was carried out on a Prelude synthesizer from Protein Technologies using Fmoc/tBu chemistry. The side-chain protecting groups used were Arg(Pbf), Asn(Trt), Asp(O*t*Bu), Cys(Acm), Cys(Trt), Glu(O*t*Bu), Gln(Trt), His(Trt), Lys(Boc), Ser(*t*Bu), Thr(*t*Bu), Trp(Boc) and Tyr(*t*Bu). The synthesis, starting from the selected resin (see table) were performed on a 0.025 mmol/reactor scale or 0.15 mmol/reactor scale. Protected amino acids (10 equiv.) were coupled using HCTU (9.5 equiv.) and *i*Pr_2EtN (20 equiv.) in NMP for 30 min. Capping of eventual unreacted amine groups was achieved by treatment with acetic anhydride (60 equiv.), *i*Pr_2EtN (15.5 equiv.) and HOBt (1.8 equiv.) in NMP for 7 min. Fmoc group was deprotected by three successive treatments with 20% piperidine in NMP for 3 min. Manual capping and Fmoc deprotection were achieved using the same protocol.
The SPPS elongation yield was quantified using UV titration of the fluorenylmethylpiperidine byproduct at 301 nm (ε= 7800 $L \cdot mol^{-1} \cdot cm^{-1}$) corresponding to the first and last Fmoc group deprotection.

Resin	Loading (mmol/g)
Aminomethyl Tentagel R	0.21
Rink-ChemMatrix	0.5
Spheritide	0.15
Amino PEGA	0.03

Supplementary table 1. Solid supports (resins) used for Fmoc-SPPS.

b. General procedure B for peptide deprotection and resin cleavage

The crude peptidyl resin was extensively washed with CH_2Cl_2. It was then deprotected and cleaved from the resin by treatment with TFA/H_2O/*i*Pr_3SiH/DODT[1]/phenol, 85/5/2/5/5 for 3.5 h for cysteine-containing peptides and TFA/H_2O/*i*Pr_3SiH/phenol, 85/5/2/8 for 2 h for cysteine-free peptides. The peptide was then precipitated by dilution into an ice-cold 1:1

[1] DODT: 3,6-dioxa-1,8-ethanedithiol

diethyl ether/petroleum ether mixture, recovered by centrifugation and washed twice with diethyl ether.

c. General procedure C for Kaiser test.[2]

The resin was washed with CH_2Cl_2. A few beads of the resin were placed in a hemolysis tube. Then 40 µL of each of the three following solution were added: solution 1 (5 g ninhydrin in 100 mL EtOH), solution 2 (80 g phenol in 100 mL EtOH), and solution 3 (2 mL of a 0.1 mM aqueous KCN in 98 mL) were added successively to the resin. The suspension was then heated at 100°C for 5 min. Blue color indicates positive test thus presence of free amines into the resin.

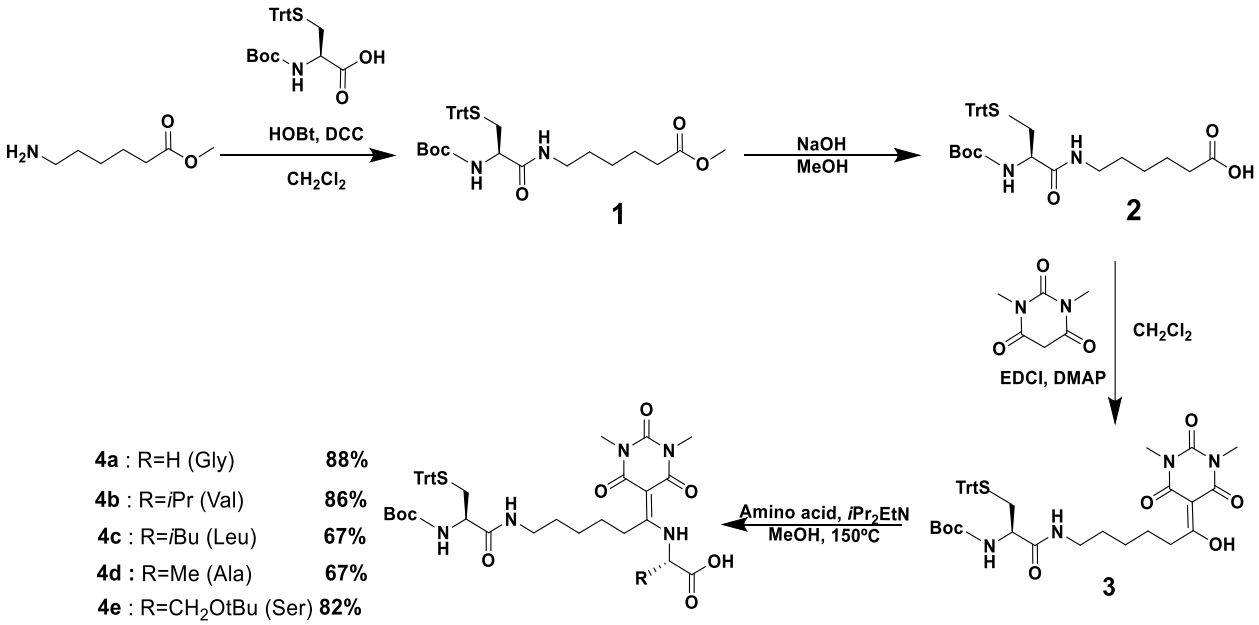
d. General procedure D for coupling of Boc-Cys(Trt)-Dtph-linker into the amine peptidyl resin.

Boc-Cys-(Trt)-Dtph-Xaa-OH linkers (2.5 equiv.) and PyAOP (2.5 equiv.) were dissolved in peptide grade DMF. Then *i*Pr_2EtN (5 equiv.) was added. The resulted mixture was immediately added to the peptidyl resin (1 equiv., swollen in DMF prior to the reaction), and the resulting suspension was stirred for 4 h at room temperature. The resin was then thoroughly washed with DMF. In case of incomplete reaction (positive Kaiser's test), this protocol was repeated once with overnight stirring at room temperature. The resin was then cleaved following the general protocol B.

[2] [2] Kaiser, E. R. L.; Colescott, C. D.; Bossinger, P.; Cook, I. *Anal. Biochem.* **1970**, *34*, 595.

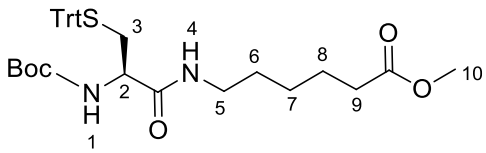
3. Preparation of the model linker-containing peptides

3.1. Synthesis of Boc-Cys(Trt)-Dtph-Xaa-OH



Supplementary figure 1 Synthesis of Boc-Cys(Trt)-Dtph-OH , overall yield 64%

Compound **1**

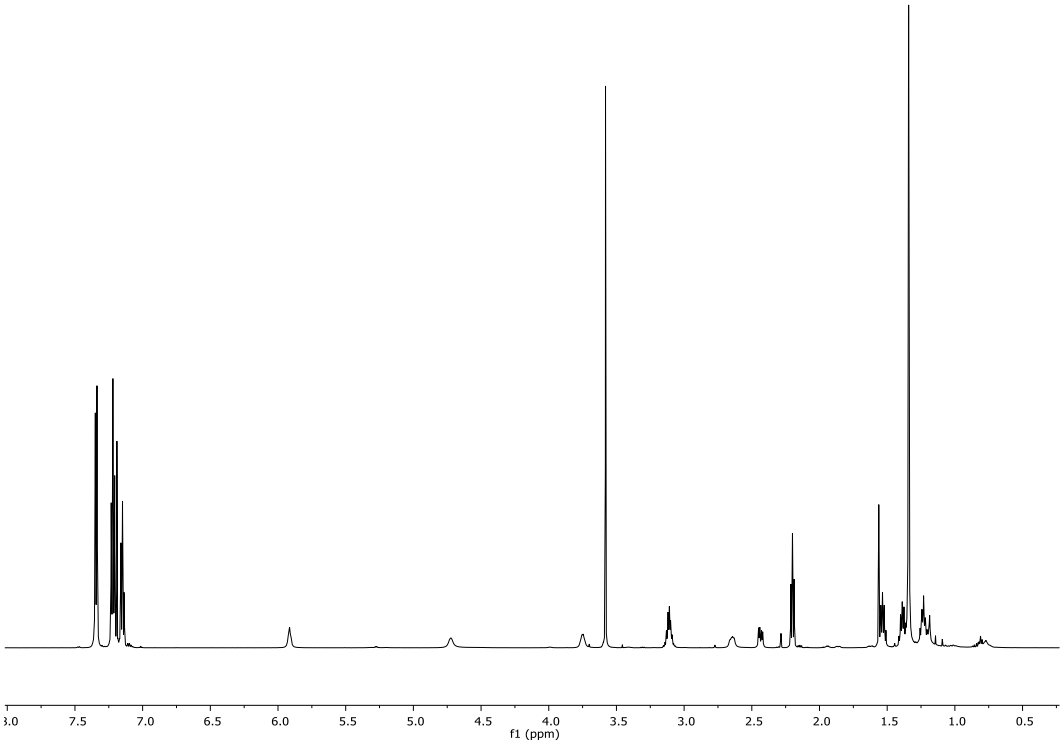


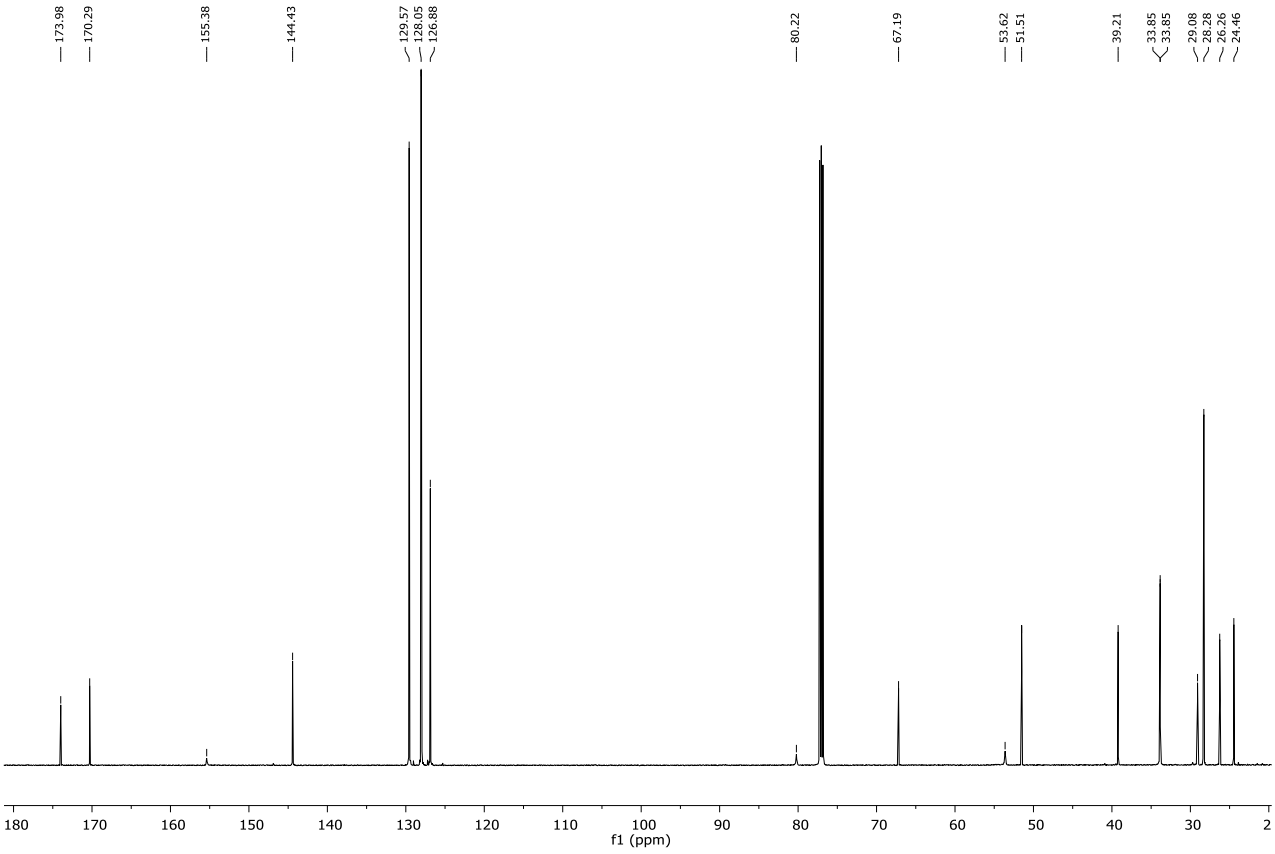
A solution of methyl-6-aminohexanoate hydrochloride (250 mg, 1.6 equiv.) and iPr_2EtN (256.2 µL, 1.7 equiv.) in CH_2Cl_2 (20 mL) was cooled in an ice bath. Then Boc-Cys(Trt)-OH (0.4 g, 0.86 mmol,1 equiv.), HOBt (0.116 g, 1 equiv.) and DCC (0.213 g, 1.2 equiv) were added. The resulted mixture was warmed up to room temperature and was then stirred overnight. An aqueous 1 M HCl solution was added (3 x 10 mL) and the organic phase was then separate washed with a saturated aqueous solution of $NaHCO_3$ (3 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtered then concentrated under reduced pressure. Purification by flash column chromatography (eluent: petroleum ether/ EtOAc, 6:4) afforded compound **1** (442 mg, 87%) as a white amorphous solid.

^{1}H NMR (600 MHz, $CDCl_3$): δ 7.37 - 7.32 (m, 6H, H_{Trt}), 7.22 - 7.18 (m, 6H, H_{Trt}), 7.19 - 7.12 (m, 3H, H_{Trt}), 5.91 (bt, *J* = 5.9 Hz, 1H, H_4), 4.72 (bs, 1H, H_1), 3.75 - 3.71 (m, 1H, H_2), 3.58 (s, 3H, H_{10}), 3.17 - 3.05 (m, 2H, H_7), 2.64 - 2.59 (m, 1H, H_{3a}), 2.43 - 2.39 (m , 1H, H_{3b}), 2.20 (bt, *J* = 7.5 Hz, 2H, H_9), 1.53 - 1.48 (m , 2H, H_8), 1.38 - 1.36 (m , 2H, H_6), 1.34 (s, 9H, H_{Boc}), 1.27 - 1.17 (m, 2H, H_7).

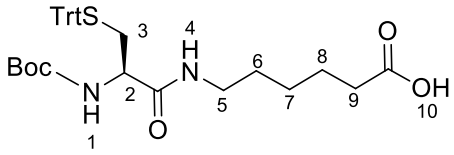
^{13}C NMR (150 MHz, $CDCl_3$): δ 174, 170.3, 155.4, 144.4, 129.6, 128, 126.9, 80.2, 67.2, 53.6, 51.5, 39.2, 33.9, 33.8, 29.0, 28.3, 26.3, 24.5

ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{34}H_{43}N_2O_5S$: 591.2884, found: 591.2887.





Compound **2**

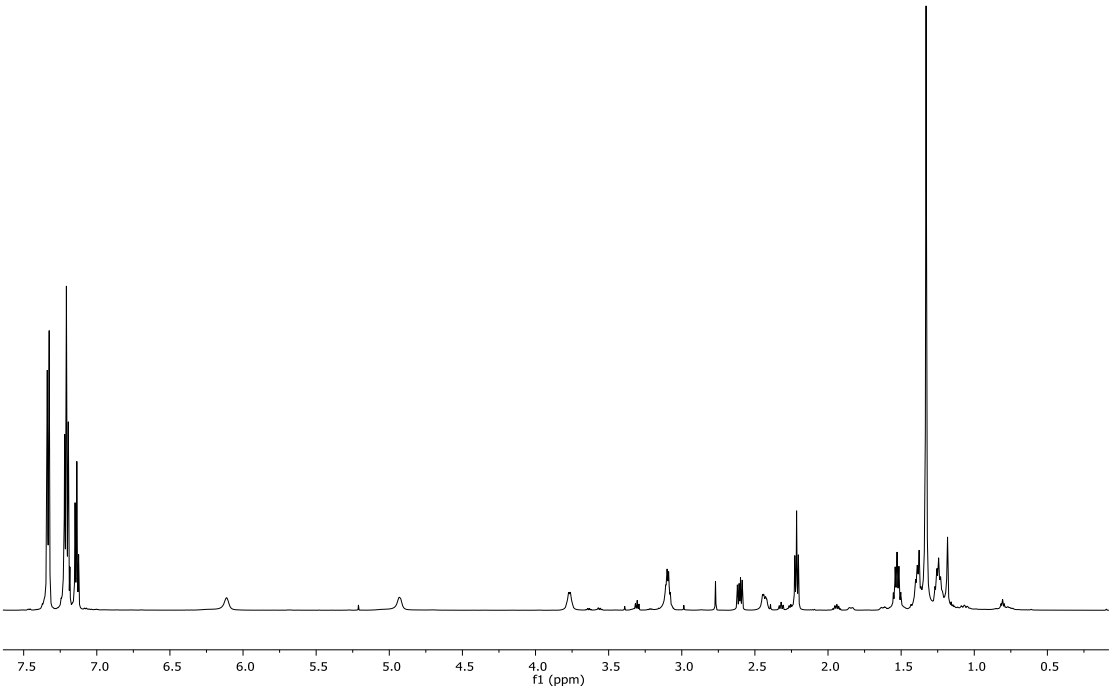


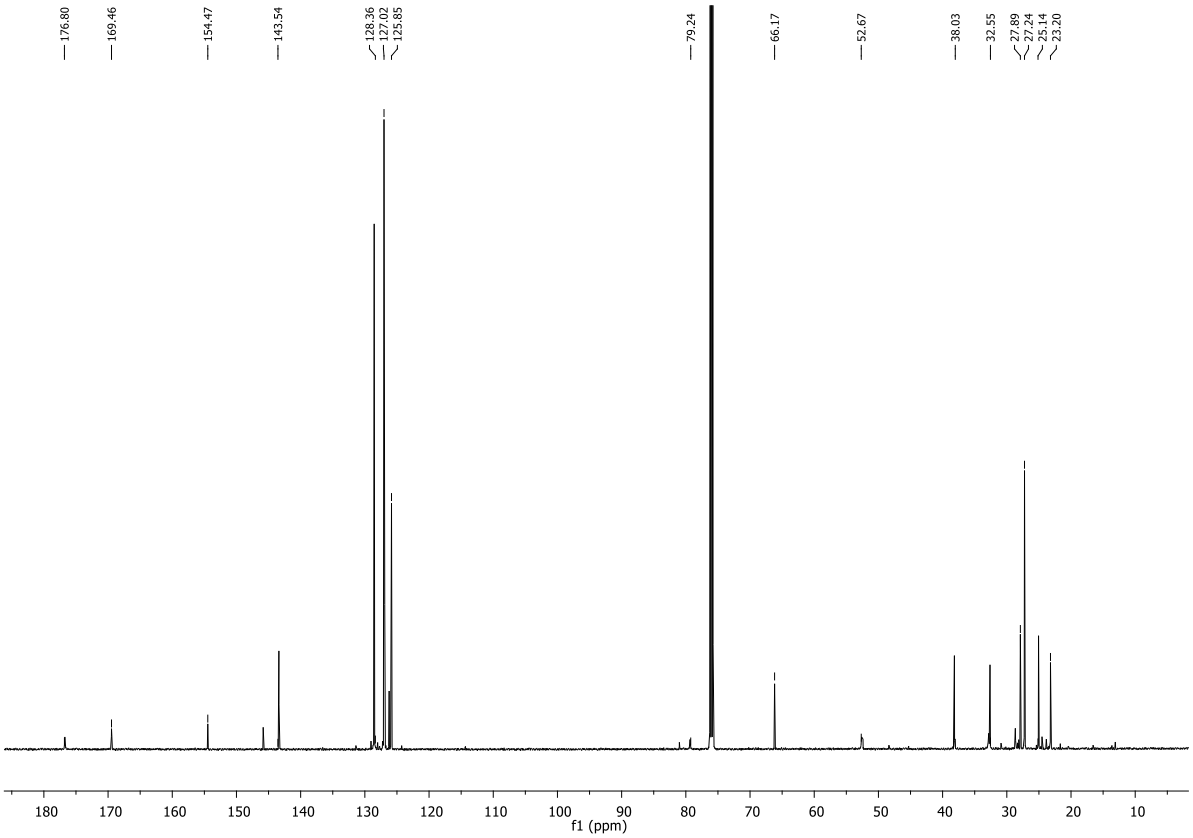
To a solution of **1** (442 mg, 0.71 mmol, 1 equiv.) in MeOH (16 mL) was added a solution of NaOH (120 mg, 4 equiv.) in water (5 mL) and the resulting mixture was heating at 50 °C for 5 h. Then, a solution of 10% citric acid was added dropwise to pH = 2, leading to the precipitation of a yellow solid. Volatiles were then evaporated under reduced pressure. The resulting suspension was diluted with CH_2Cl_2 (30 mL) and washed with aqueous 1 M HCl solution (3 x 10 mL), dried over $MgSO_4$, filtered then concentrated under reduced pressure to afford compound **2** (355 mg, 82%) as a white amorphous solid. Compound **2** was used in the next step without further purification.

^{1}H NMR (600 MHz, $CDCl_3$): δ 7.37 - 7.31 (m, 6H, H_{Trt}), 7.25 - 7.18 (m, 6H, H_{Trt}), 7.17 - 7.11 (m, 3H, H_{Trt}), 6.03 - 5.93 (m, 1H, H_4), 4.84 (bs, 1H, H_1), 3.76 (dt, *J* = 7.4, 5.5 Hz, 1H, H_2), 3.16 - 3.07 (m, 2H, H_5), 2.62 (dd, *J* = 12.9, 7.4 Hz, 1H, H_{3a}), 2.44 (dd, *J* = 12.3, 7.2 Hz, 1H, H_{3b}), 2.23 (t, *J* = 7.4 Hz, 2H, H_9), 1.54 - 1.49 (m, 2H, H_8), 1.43 - 1.37 (m, 2H, H_6), 1.34 (s, 9H, H_{Boc}), 1.29 - 1.20 (m, 2H, H_7).

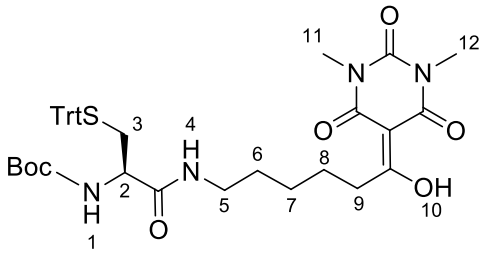
^{13}C NMR (150 MHz, $CDCl_3$): δ 176.8, 169.5, 154.5, 143.5, 128.4, 127.0, 125.9, 79.2, 66.2, 52.7, 38.0, 32.6, 27.9, 27.2, 25.1, 23.2.

ESI-HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{34}H_{41}N_2O_5S$: 577.2730, found: 577.2728.





Compound **3**

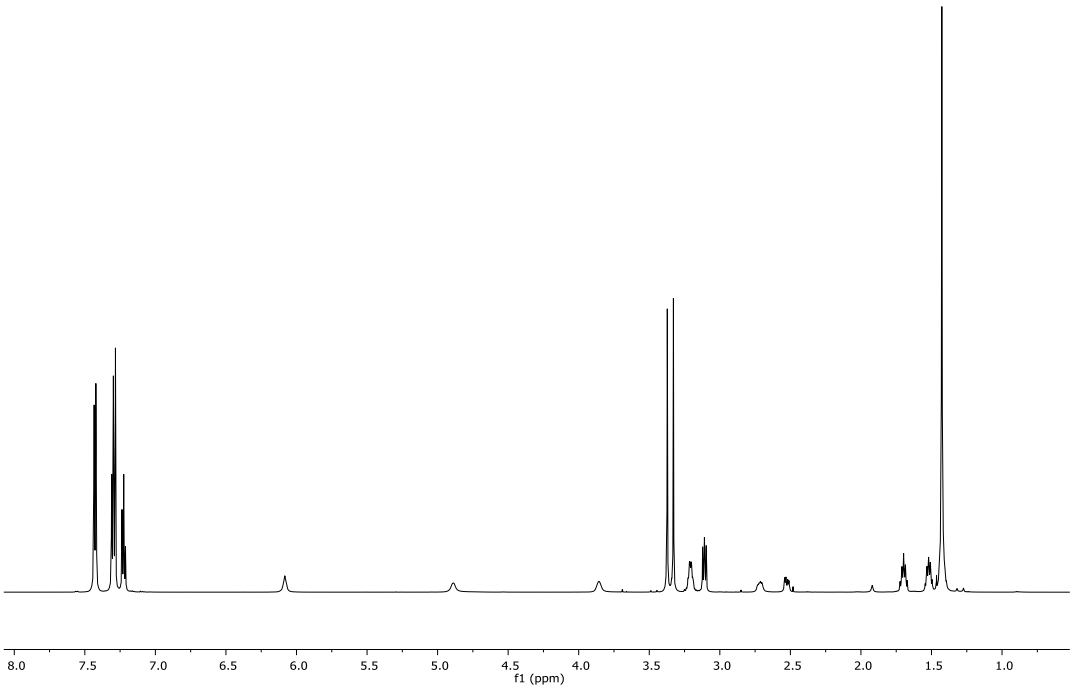


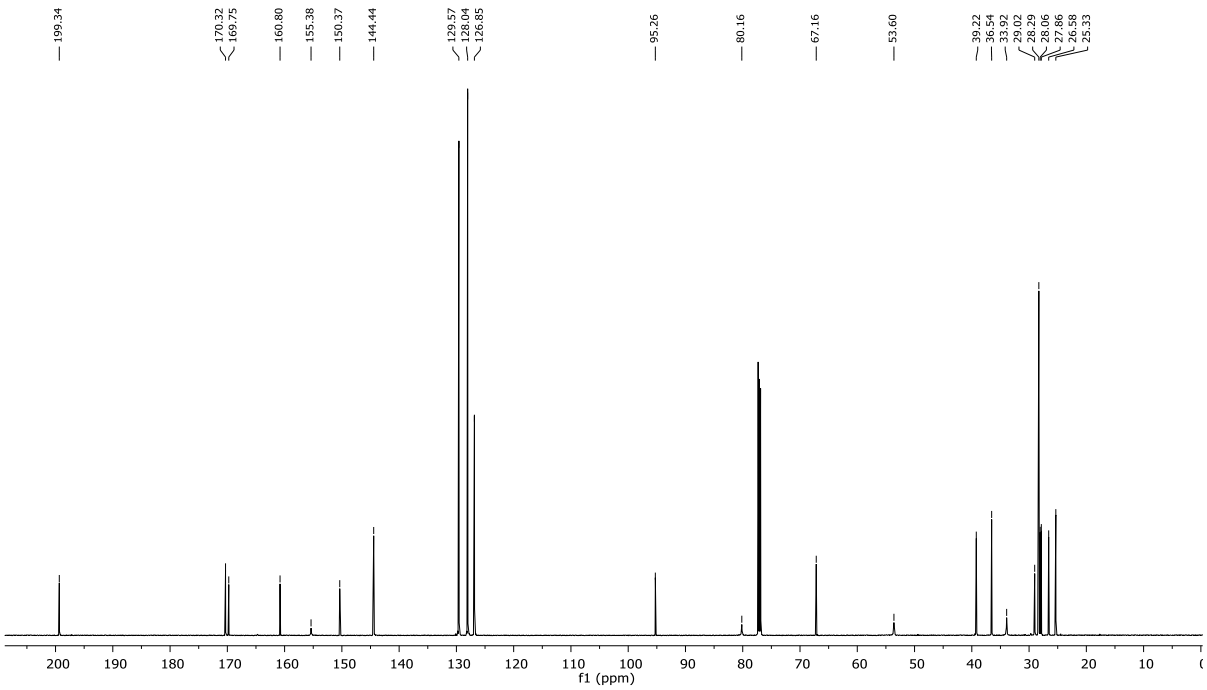
A solution of 1,3-dimethylbarbituric acid (115 mg, 1.2 equiv.), compound **2** (355 mg, 0.61 mmol, 1 equiv.) and DMAP (90 mg, 1.2 equiv.) in CH_2Cl_2 (30 mL) were cooled in an ice bath. EDCI (142 mg, 1.2 equiv.) was then added to the resulted mixture. The reaction mixture was stirred at room temperature for 4 h, washed then with an aqueous 1 M HCl solution (3 x 10 mL). The organic phase was separate washed with an aqueous saturated solution of $NaHCO_3$ (3 x 10 mL). The organic layer was dried over $MgSO_4$, filtered then concentrated under reduced pressure. Purification by flash column chromatography (eluent: CH_2Cl_2/MeOH. 99:1) afforded compound **3** (1.08 g, 44%) as a yellow amorphous solid.

^{1}H NMR (600 MHz, $CDCl_3$): δ 7.48 - 7.36 (m, 6H, H_{Trt}), 7.35 - 7.28 (m, 6H, H_{Trt}), 7.26 - 7.12 (m, 3H, H_{Trt}), 6.13 - 6.01 (m, H , H_4), 4.79 (s, 1H, H_1), 3.81 - 3.74 (m, 1H, H_2), 3.37 (s, 3H, $H_{11\ or\ 12}$), 3.33 (s, 3H, $H_{12\ or\ 11}$), 3.21 - 3.18 (m, 2H, H_5), 3.09 (t, *J* = 7.4 Hz, 2H, H_9), 2.69 (dd, *J* = 12.9, 7.4 Hz, 1H, H_{3a}), 2.50 (dd, *J* = 12.3, 7.2 Hz, 1H, H_{3b}), 1.74 - 1.68 (m, 2H, H_8), 1.52 - 1.48 (m, 2H, H_6), 1.43 - 1.39 (m, 11H, H_{Boc} H_7).

^{13}C NMR (150 MHz, $CDCl_3$): δ 199.3, 170.3, 169.8, 160.8, 150.4, 150.37, 144.4, 129.6, 128.04, 126.9, 95.2, 80.2, 67.2, 53.5, 39.2, 36.5, 33.9, 29.0, 28.3, 28.1, 27.89, 26.4, 25.3.

ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{39}H_{47}N_4O_7S$: 715.3156, found: 715.3159.

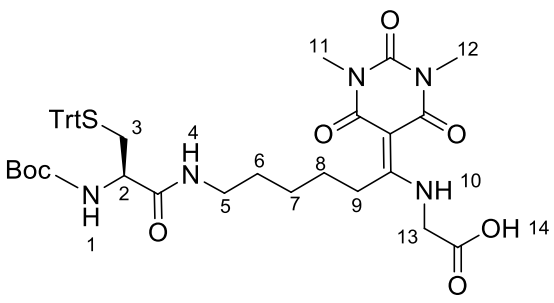




General protocol for the preparation of Boc-Cys(Trt)-Dtph-Xaa-OH: compounds 4a-e

A solution of compound **3** (1 equiv.), amino acid (Gly, Val, Leu, Ala or Ser(O*t*Bu)) (4 equiv), *i*PR_2EtN (4 equiv) in 5 mL MeOH were heated at 150°C for 10 min in a sealed tube using microwave irradiation. The resulting suspension was filtered and evaporated under reduced pressure and diluted with CH_2Cl_2 (30 mL). The organic phase was separate washed with an aqueous 1 M HCl solution. The organic layer was dried over $MgSO_4$, filtered concentrated under reduced pressure then purified by flash column chromatography (eluent: $MeOH/CH_3COOH/CH_2Cl_2$ (2:2:9) to afford compound **4a**-**c**

Compound **4a**: Boc-Cys(Trt)-Dtph-Gly-OH

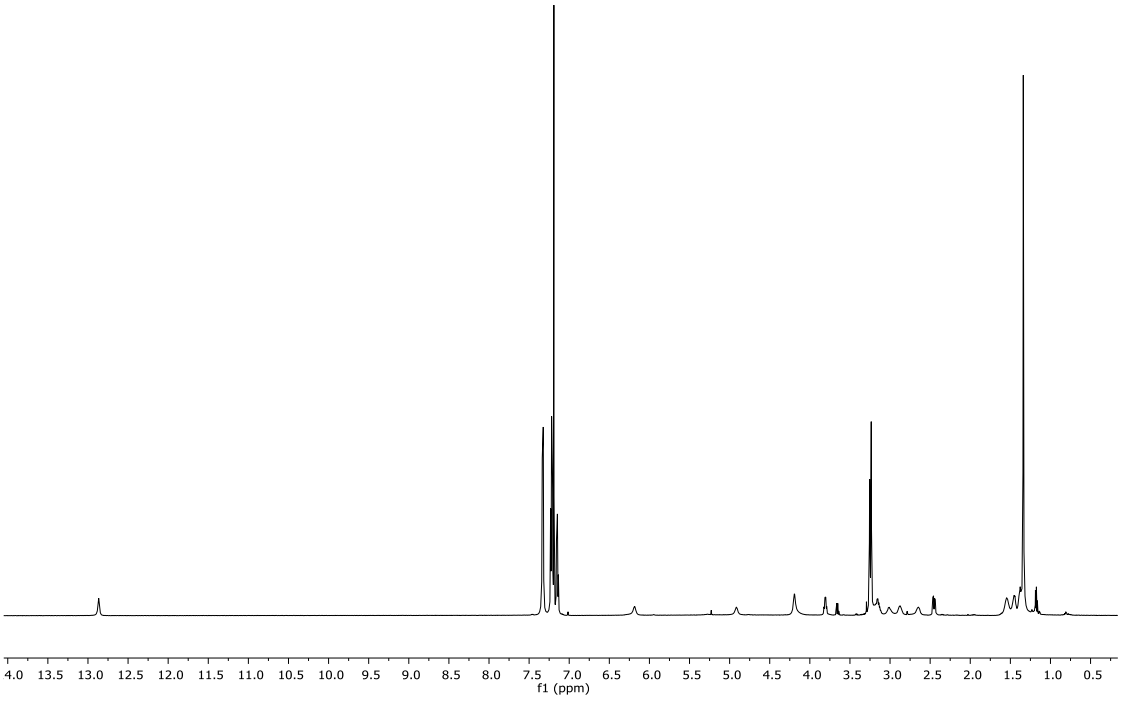


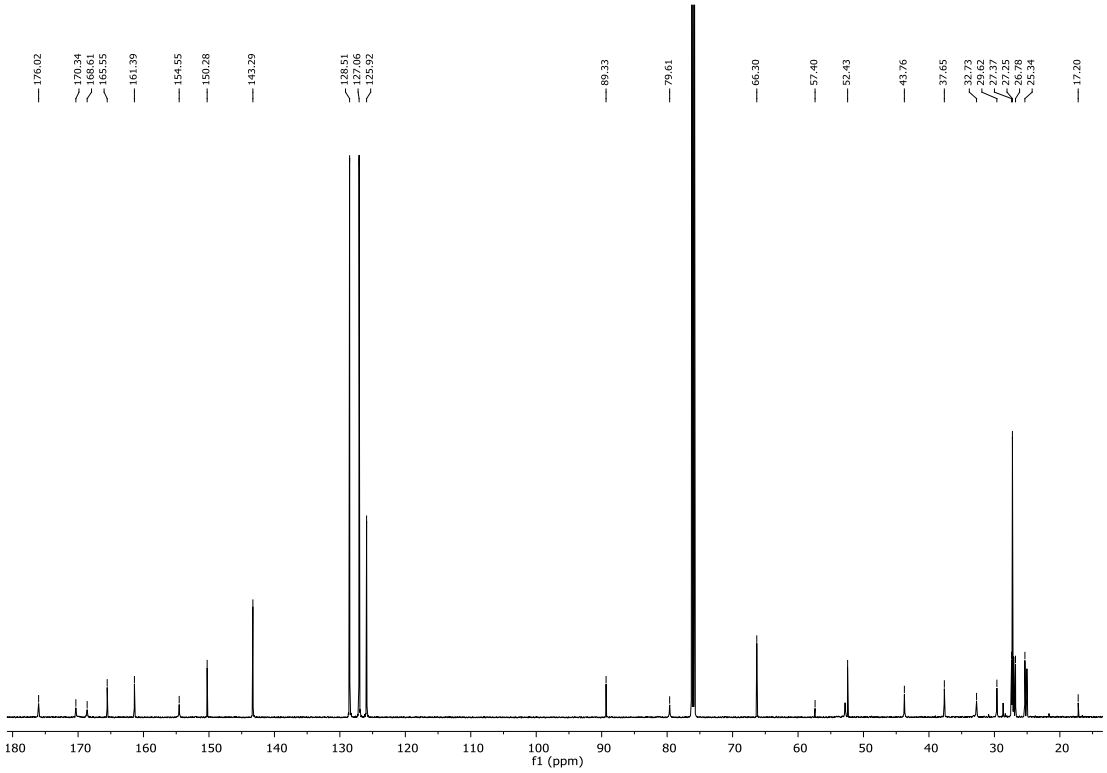
^{1}H NMR (400 MHz, $CDCl_3$, 55°C): δ 12.89 (s, 1H, H_{10}), 7.36 - 7.31 (m, 6H, H_{Trt}), 7.22 - 7.18 (m, 6H, H_{Trt}), 7.15 - 7.12 (m, 3H, H_{Trt}), 6.19 - 5.92 (m, 1H, H_4), 4.92 (s, 1H ,H_1), 4.22 - 4.15 (m, 2H, H_{13}), 3.81 - 3.72 (m, 1H, H_2), 3.21 - 3.14 (m, 2H, H_9), 3.02 - 2.90 (m, 1H, H_{5a}), 2.89 - 2.81 (m, 1H, H_{5b}), 2.69 (dd, *J* = 12.6, 7.4 Hz, 1H, H_{3a}), 2.49 (dd, *J* = 12.6, 7.2 Hz, 1H, H_{3b}), 1.57 - 1.53 (m, 2H, H_8), 1.49 - 1.42 (m, 2H, H_6), 1.36 - 1.31 (m, 2H, H_7), 1.35 - 1.32 (s, 9H, H_{Boc}).

^{13}C NMR (150 MHz, $CDCl_3$): δ 176.0, 170.3, 168.6, 165.6, 161.4, 154.6, 150.2, 143.3, 128.5, 127.1, 125.9, 89.3, 79.6, 66.3, 57.5, 52.4, 43.7, 37.6, 32.7, 29.6, 27.4, 27.2, 26.8, 25.3, 24.9, 17.20.

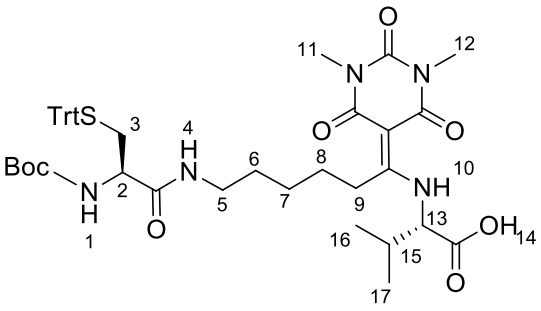
ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{41}H_{50}N_5O_8S$: 772.3371, found: 772.3374.

Yield: 88 % (363 mg). White amorphous solid.





Compound **4b**: Boc-Cys(Trt)-Dtph-Val-OH

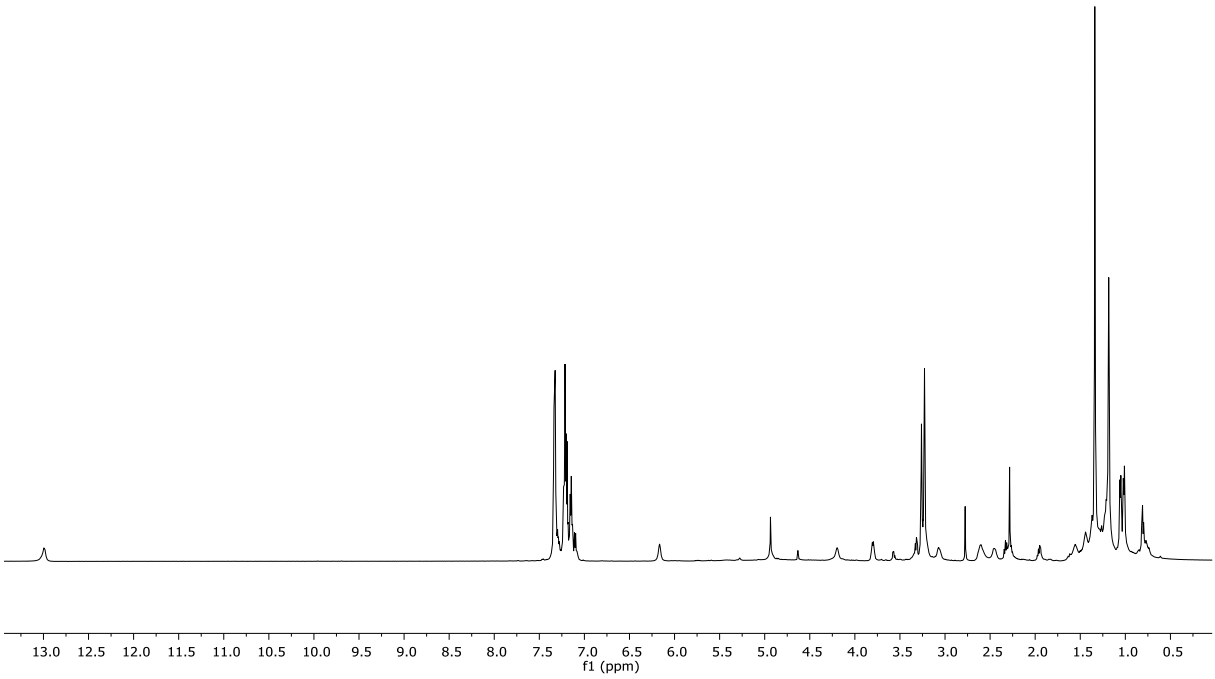


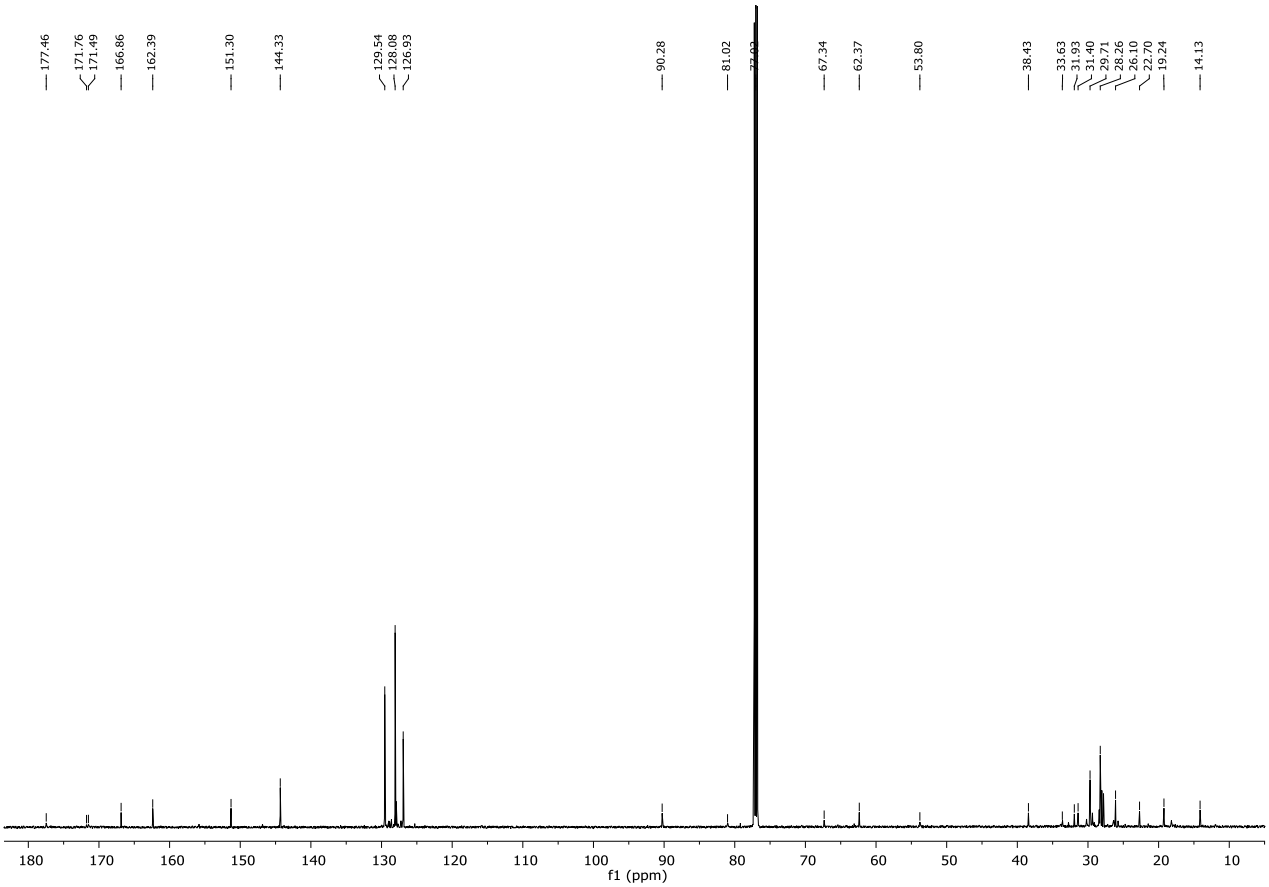
^{1}H NMR (400 MHz, $CDCl_3$, 55°C): δ 13.0 (bs, 1H, H_{10}) 7.33 - 7.31 (m, 6H, H_{Trt}), 7.24 - 7.19 (m, 6H, H_{Trt}), 7.14 - 7.11 (m, 3H, H_{Trt}), 6.21 - 6.16 (m, 1H, H_4), 4.97 - 4.92 (m, 1H, H_1), 4.21 - 4.17 (m, 1H, H_{13}), 3.80 - 3.72 (m, 1H, H_2), 3.17 - 3.28 (m, 2H, H_9), 3.26 (s, 3H, $H_{11\ or\ 12}$), 3.24 (s, 3H, $H_{11\ or\ 12}$), 3.11 - 3.01 (m, 1H, H_{5a}), 2.68 - 2.52 (m, 2H, H_{3a} H_{5b}), 2.54 - 2.40 (m, 1H, H_{3b}), 1.6 - 1.2 (m, 8H, H_6 H_7 H_8 H_{15}), 1.36 - 1.31 (s, 9H, H_{Boc}), 1.06 (d, 3H, J = 6.8 Hz, $H_{16\ or}$ H_{17}), 1.01 (d, 3H, J = 6.8 Hz, $H_{16\ or}$ H_{17}).

^{13}C NMR (150MHz, $CDCl_3$): δ 177.5, 171.8, 166.8, 162.3, 155.8, 151.3, 144.3, 129.5, 128.1, 126.9, 90.2, 80.1, 67.3, 62.3, 38.4, 33.6, 31.9, 31.4, 29.7, 28.4, 28.2, 27.8, 26.1, 22.7, 19.2, 14.1.

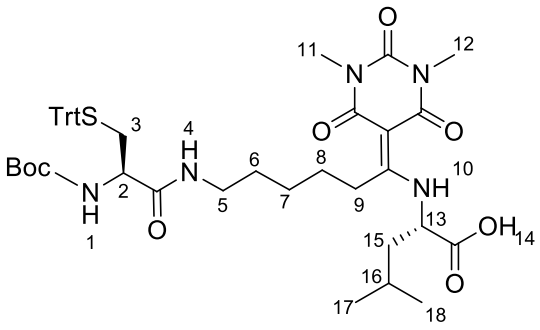
ESI-HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{44}H_56N_5O_8S$: 814.3829, found: 814.3844.

Yield: 82 % (93.3 mg). White amorphous solid.





Compound **4c**: Boc-Cys(Trt)-Dtph-Leu-OH

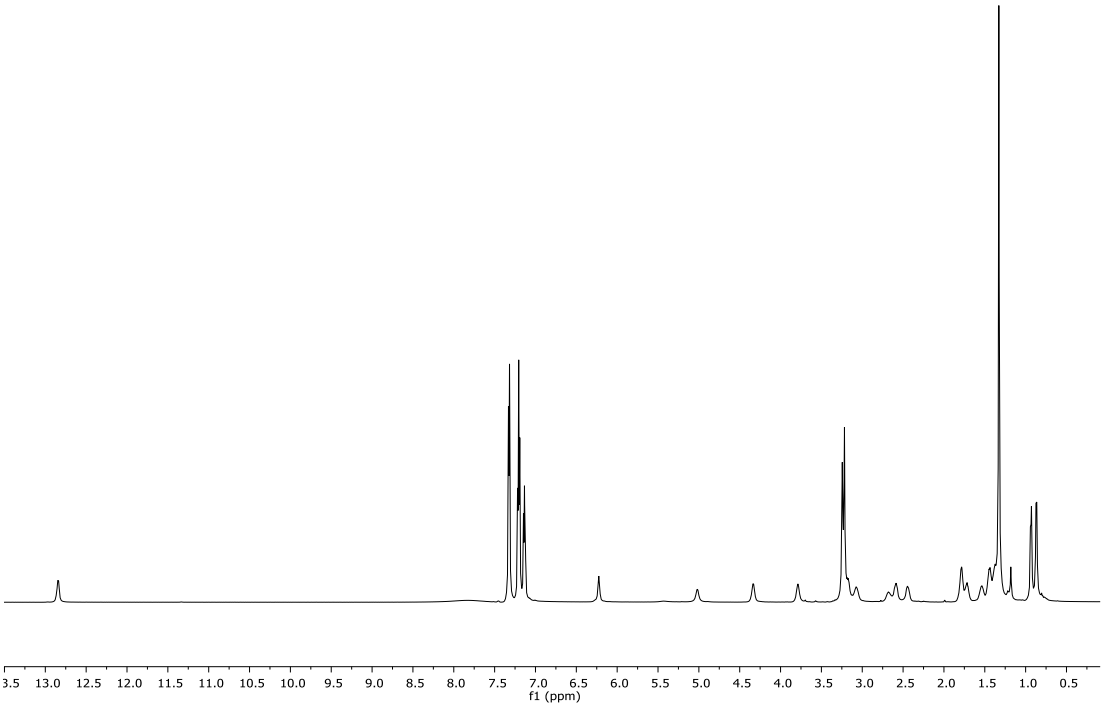


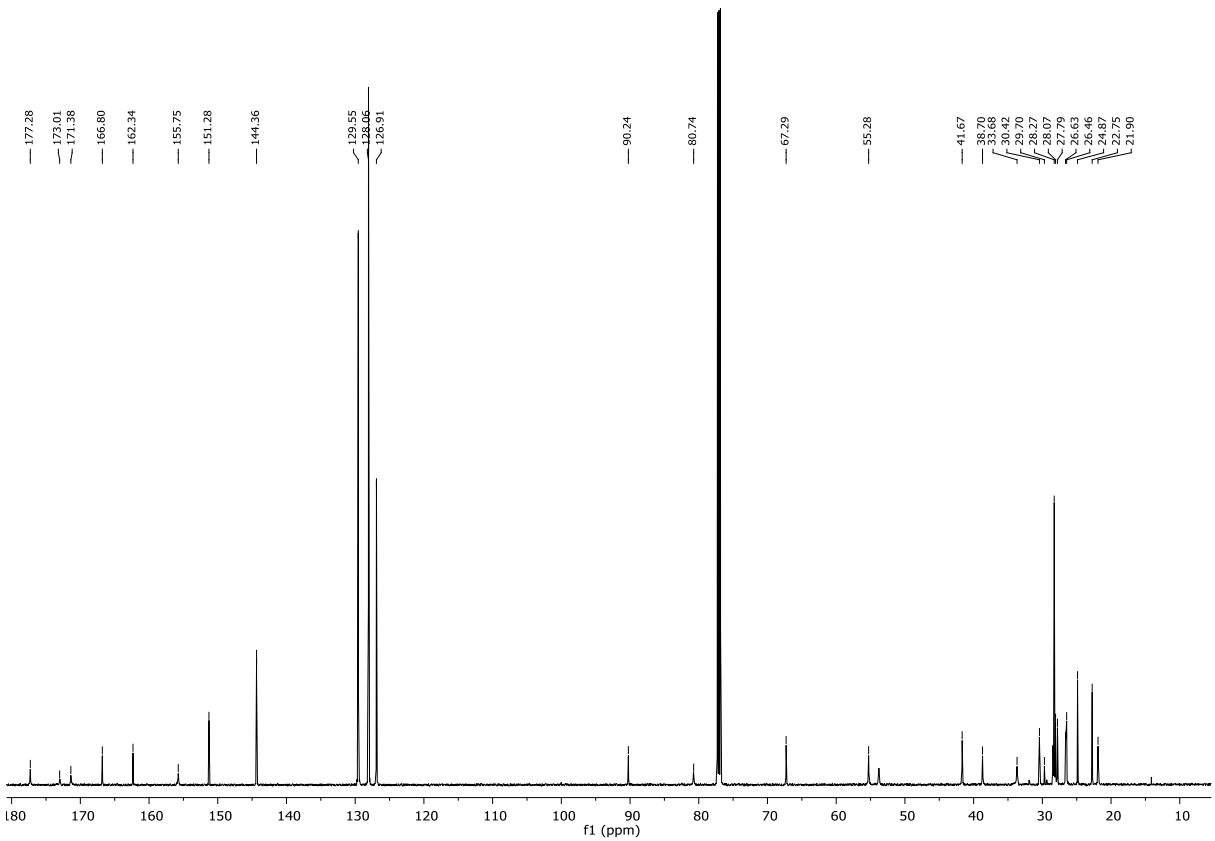
^{1}H NMR (400 MHz, $CDCl_3$, 55°C): δ 12.91 (m, 1H, H_{10}), 7.44 - 7.26 (m, 6H, H_{Trt}), 7.24 - 7.16 (m, 6H, H_{Trt}), 7.16 - 7.05 (m, 3H, H_{Trt}), 6.33 - 6.09 (m, 1H, H_4), 5.04 - 5.00 (m, 1H, H_1), 4.38 - 4.25 (m, 1H, H_{13}), 3.84 - 3.72 (m, 1H, H_2), 3.30 - 3.00 (m, 3H, H_{5a} H_9), 3.24 (s, 3H, $H_{11\ or\ 12}$), 3.22 (s, 3H, $H_{11\ or\ 12}$), 2.74 - 2.62 (m, 2H, H_{5b}), 2.64 - 2.53 (m, 1H, H_{3a}), 2.49 - 2.36 (m, 1H, H_{3b}), 1.83 - 1.39 (m, 8H, H_6 H_7 H_8 H_{15}), 1.39 - 1.34 (m, 1H, H_{16}), 1.34 - 1.30 (s, 9H, H_{Boc}), 0.94 (d, 3H, *J* = 6.4 Hz, H_{17}), 0.87 (d, 3H, *J* = 6.5 Hz, H_{18}).

^{13}C NMR (150 MHz, $CDCl_3$): δ 177.2, 171.4, 166.8, 162.3, 155.6, 151.2, 144.3, 129.5, 128.06, 126.9, 90.2, 80.76, 67.2, 55.2, 53.7, 41.6, 38.7, 33.6, 30.4, 29.7, 28.4, 28.2, 28.1, 27.7, 26.6, 26.46, 24.8.7, 21.9.

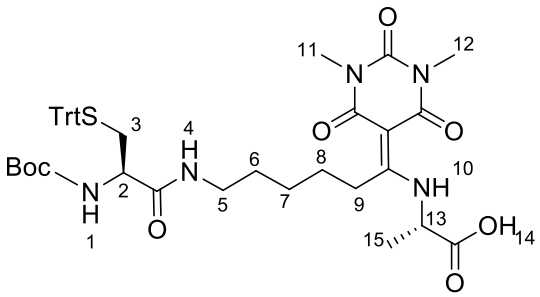
ESI-HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{45}H_58N_5O_8S$: 828.4001, found: 828.3998.

Yield: 67 % (273.8 mg). White amorphous solid.





Compound **4d**: Boc-Cys(Trt)-Dtph-Ala-OH

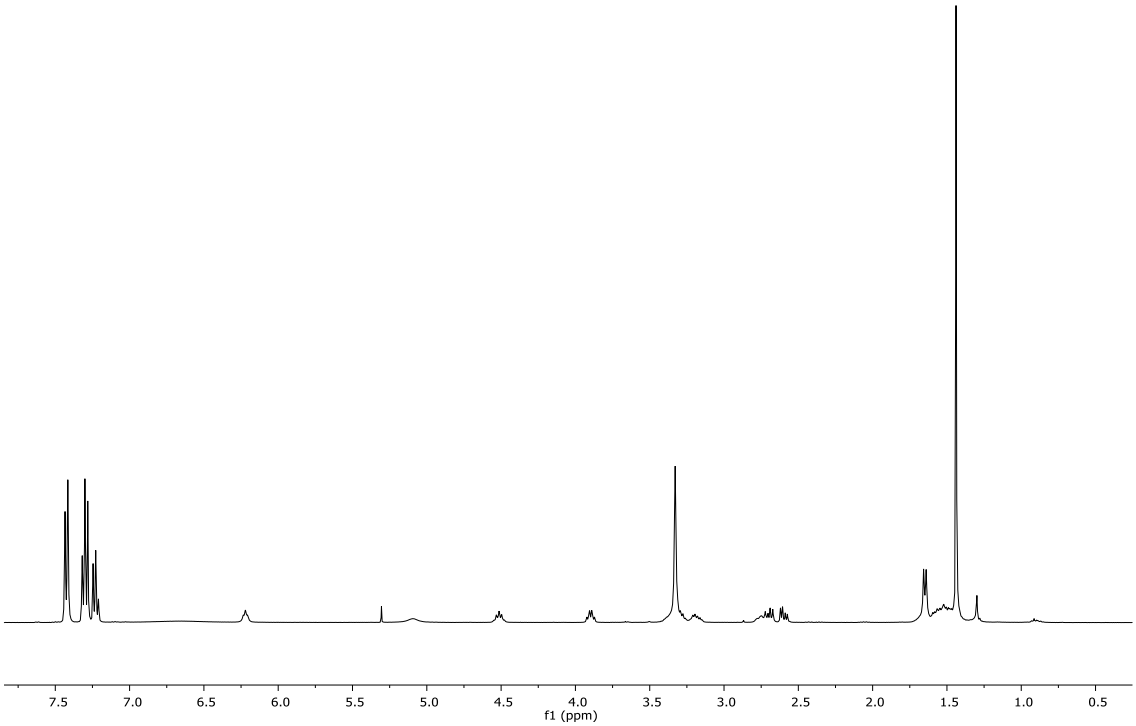


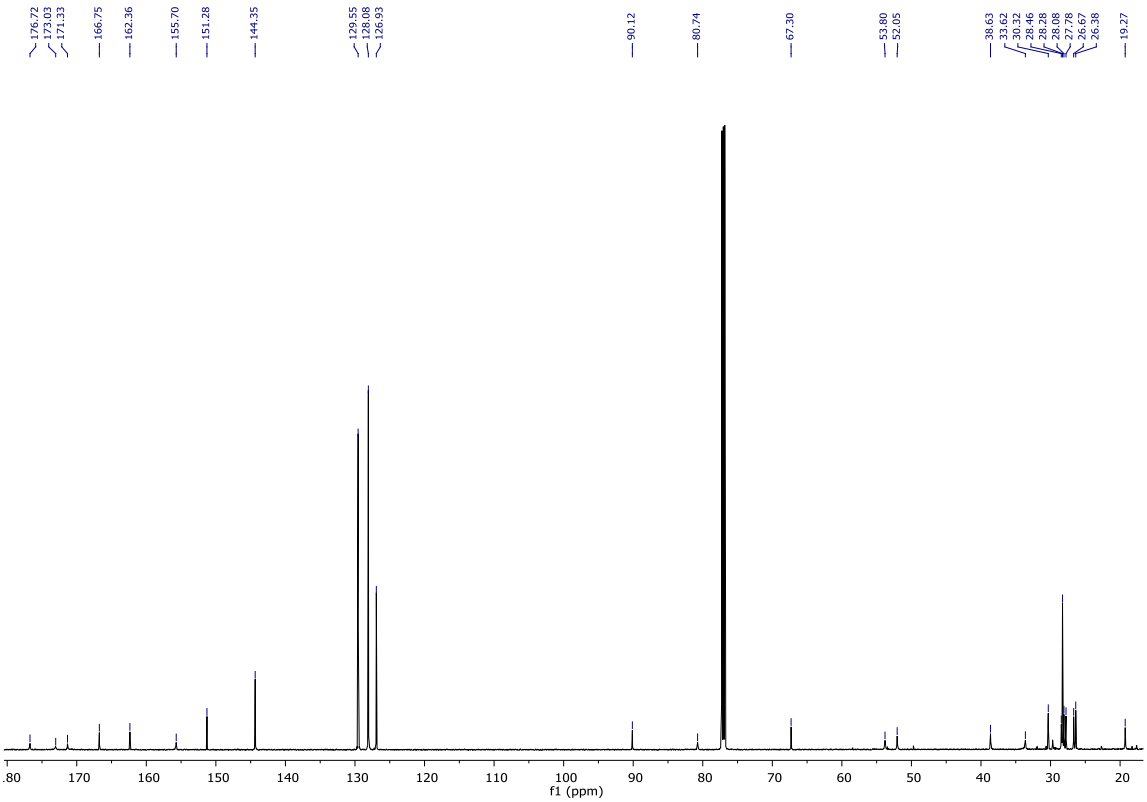
^{1}H NMR (400 MHz, $CDCl_3$, 55°C): δ 7.48 - 7.34 (m, 6H, H_{Trt}), 7.33 - 7.26 (m, 6H, H_{Trt}), 7.26 - 7.17 (m, 3H, H_{Trt}), 6.22 - 6.11 (m, 1H, H_4), 5.10 (s, 1H, H_1), 4.59 - 4.41 (m, 1H, H_{13}), 3.98 - 3.79 (m, 1H, H_2), 3.43 - 3.13 (m, 3H, H_{5a} H_9), 3.33 (bs, 6H, H_{11} H_{12}), 2.80 - 2.70 (m, 1H, H_{5b}), 2.70 (dd, *J* = 12.9, 7.2 Hz, 1H, H_{3a}), 2.60 (dd, *J* = 12.9, 5.7 Hz, 1H, H_{3b}), 167 - 1.62 (d, 3H, *J* = 6.9 Hz H_{15}), 1.60 - 1.41 (m, 6H, H_6 H_7 H_8), 1.46 - 1.40 (s, 9H, H_{Boc}).

^{13}C NMR (150 MHz, $CDCl_3$): δ 176.7, 173.0, 171.3, 166.8, 162.4, 155.7, 151.3, 144.3, 129.5, 128.1, 126.9, 90.1, 80.7, 67.3, 53.8, 52.1, 38.6, 33.6, 30.3, 28.5, 28.3, 28.1, 27.8, 26.7, 26.4, 19.3.

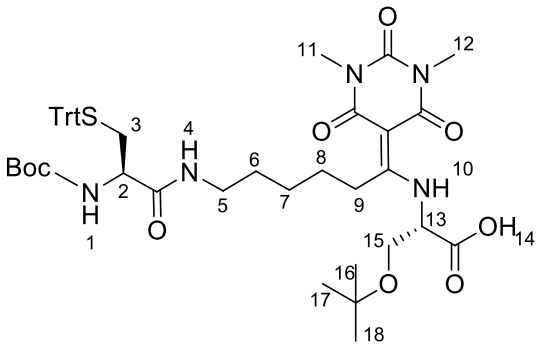
ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{42}H_{52}N_5O_8S$: 786.3458, found: 786.3531.

Yield: 67 % (73.8 mg). White amorphous solid.





Compound **4e**: Boc-Cys(Trt)-Dtph-Ser(O*t*Bu)-OH

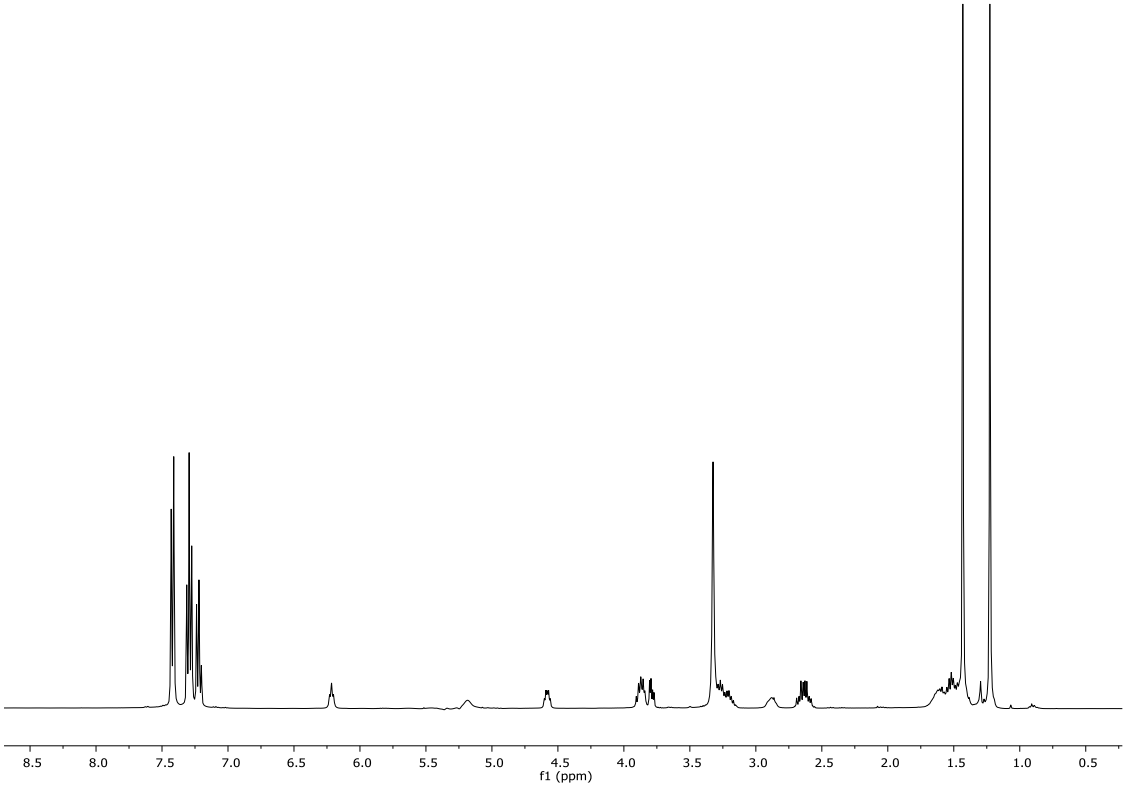


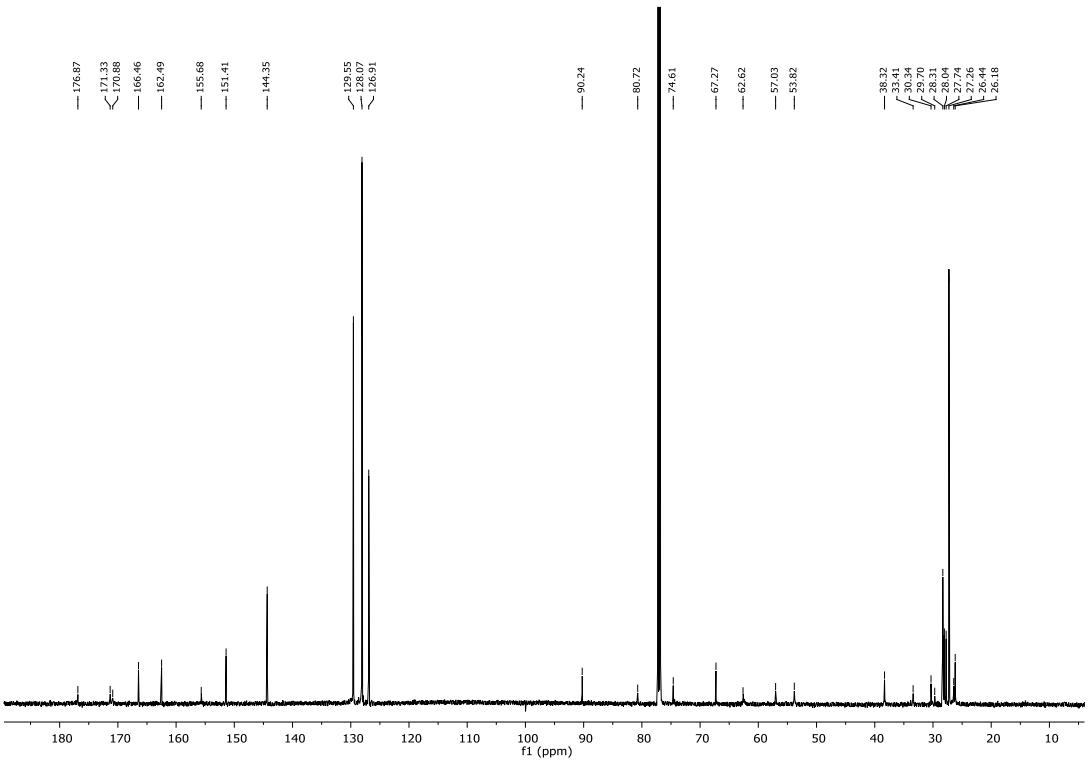
^{1}H NMR (400 MHz, $CDCl_3$, 55°C): δ 7.47 - 7.36 (m, 6H, H_{Trt}), 7.34 - 7.26 (m, 6H, H_{Trt}), 7.26 - 7.16 (m, 3H, H_{Trt}), 6.22 (bt, *J* = 5.7 Hz, 1H, H_4), 5.23 - 513 (m, 1H, H_1), 4.63 - 4.53 (m, 1H, H_{13}), 3.91 - 3.82 (m, 2H, H_2 H_{15a}), 3.79 (dd, *J* = 9.2, 4.2 Hz, 1H, H_{15b}), 3.32 (s, 6H, H_{11} H_{12}), 3.30 - 3.15 (m, 3H, H_9 H_{5a}), 2.95 - 2.82 (m, 1H, H_{5b}), 2.67 (dd, *J* = 12.9, 7.1 Hz, 1H, H_{3a}), 2.62 (dd, *J* = 12.9, 5.2 Hz, 1H, H_{3b}), 1.67 - 1.44 (m, 6H, H_6 H_7 H_8), 1.43 (s, 9H, H_{Boc}), 1.23 (s, 9H, H_{16}).

^{13}C NMR (150 MHz, $CDCl_3$): δ 176.9, 171.3, 170.9, 166.5, 162.5, 155.7, 151.4, 144.4, 129.5, 128.1, 126.9, 90.2, 80.7, 74.6, 67.3, 62.6, 57.0, 53.8, 38.3, 33.4, 30.3, 29.7, 28.3, 28.0, 27.7, 27.3, 26.4, 26.2.

ESI-HRMS (*m/z*): $[M+H]^+$ calcd. for $C_{46}H_{60}N_5O_9S$: 858.4033, found: 858.4106.

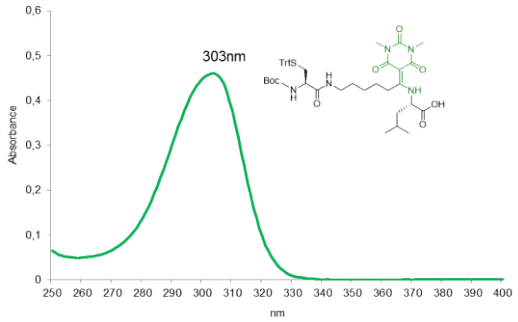
Yield: 82 % (84.4 mg). White amorphous solid.





UV-VIS characterization of Boc-Cys(Trt)-Dtph-Leu-OH (compound **4b**)

Determination of the molar extinction coefficient (ε) of **4b** was obtained by measuring the absorbance of the linker by UV spectroscopy at 280, 303 and 320 nm respectively. The molar extinction coefficient of compound **4b** was determined in a H_2O/MeCN/TFA (1:1:0.001) solution with the Beer-Lambert Law (table 1). The S-Trt group do not absorb at these wavelengths.



Supplementary figure 2: UV spectrum of 4b

Wavelength (nm)	**ε (M^{-1} cm^{-1})**
280	1780
303 (λ max)	**5980**
320	1230

Suplementary table 2: Molar extinction coefficients of **4b** (10 mg in 100 mL) at selected wavelength.

3.2. Solid phase synthesis of model peptide 5

Sequence of 20-mer MUC1-derived model peptide: H-WTSAPDTRPAPGSTAPPAHG-NH2 Synthesis starting from aminomethyl Tentagel R resin was performed on the Prelude synthesizer, on a 0.125 mmol/reactor scale following protocol A.

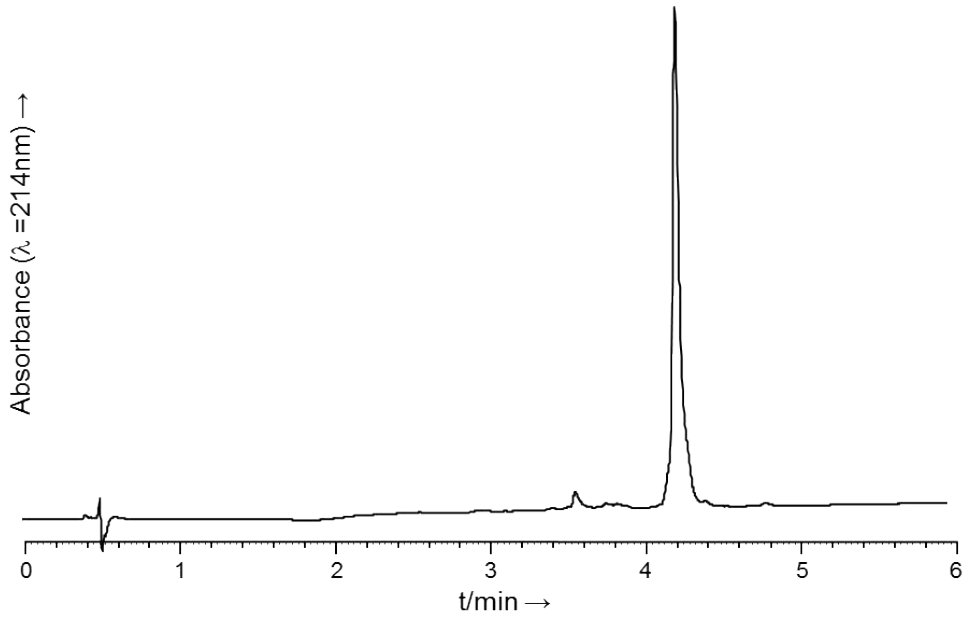
An aliquot of the result peptidyl resin was cleaved for characterization purposes (see protocol B).

The SPPS elongation yields was quantified through UV titration at 301 nm of the fluorenylmethylpiperidine byproduct (ε= 7800 $L \cdot mol^{-1} \cdot cm^{-1}$) corresponding to the first and last Fmoc group deprotection

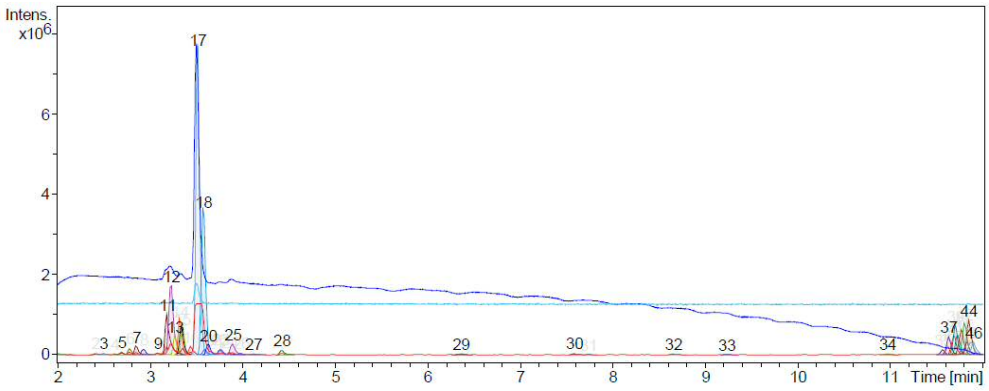
ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{86}H_{130}N_{27}O_{27}$: 1972.9623, found: 1972.9643.

HPLC analysis: retention time: 4.39 min (Chromolith, gradient: 0-30 % B/A over 5 min).

Elongation yields: 79%.



Supplementary figure 3: HPLC trace of crude of **5** at λ = 214 nm.

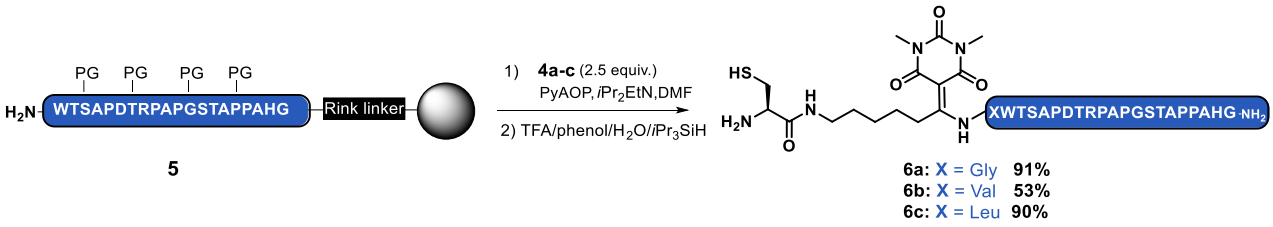


Supplementary figure 4: LC/MS of crude **5**. Blue trace UV 214 nm. Red trace: base peak ion chromatogram.

Peak	$[M+H]^+$ (*m/z*) Calcd.	$[M+H]^+$ (*m/z*) found	Attributed to
12	1256.6422	1256.6495	Ac-[7-20]
17	1357.6899	1357.6969	Ac-[6-20]
	1972.9623	1972.9643	**5**
18	1100.5411	1100,5485	Ac-[8-20]
	1727.8387	1727,8468	Ac-[2-20]
	1568.7696	1569,7769	Ac-[4-20]

Supplementary table 3: Attribution of major peaks observed in LC/MS analysis of crude **5**.

3.3. Preparation of Cys-Dtph-containing peptides

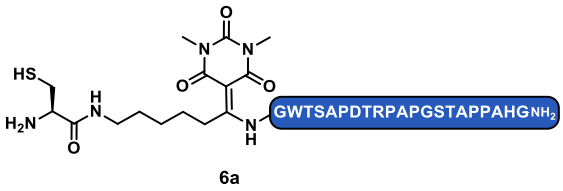


Supplementary figure 5: Syntheses of compounds **6a**-**c**.

Boc-Cys-(Trt)-Dtph-OH linkers (**4a**-**c**) (2.5 equiv.) and PyAOP (2.5 equiv.) were dissolved in peptide grade DMF. Then *i*Pr$_2$EtN (5 equiv.) was added. The resulted mixture was immediately added to the peptidyl resin (1 equiv., swollen in DMF prior to the reaction), and the resulting suspension was stirred for 4 h at room temperature. The resin was then thoroughly washed with DMF. In case of incomplete reaction (positive Kaiser's test), this protocol was repeated once with overnight stirring at room temperature. The resin was then cleaved following the general protocol B. The resulted crude peptide was analysed by analytical RP-HPLC and then purified by semi-preparative HPLC to yield the pure Cys-Dtph-peptides which were analyzed by RP-HPLC and HRMS.

The yield which correspond to the introduction of the linker, TFA cleavage and purification were determined by UV titration at 303 nm (ε_{303}(Dtph) = 5980 M^{-1} cm^{-1}) in a H_2O/MeCN/TFA (1:1:0.001) solution.

Compound **6a**



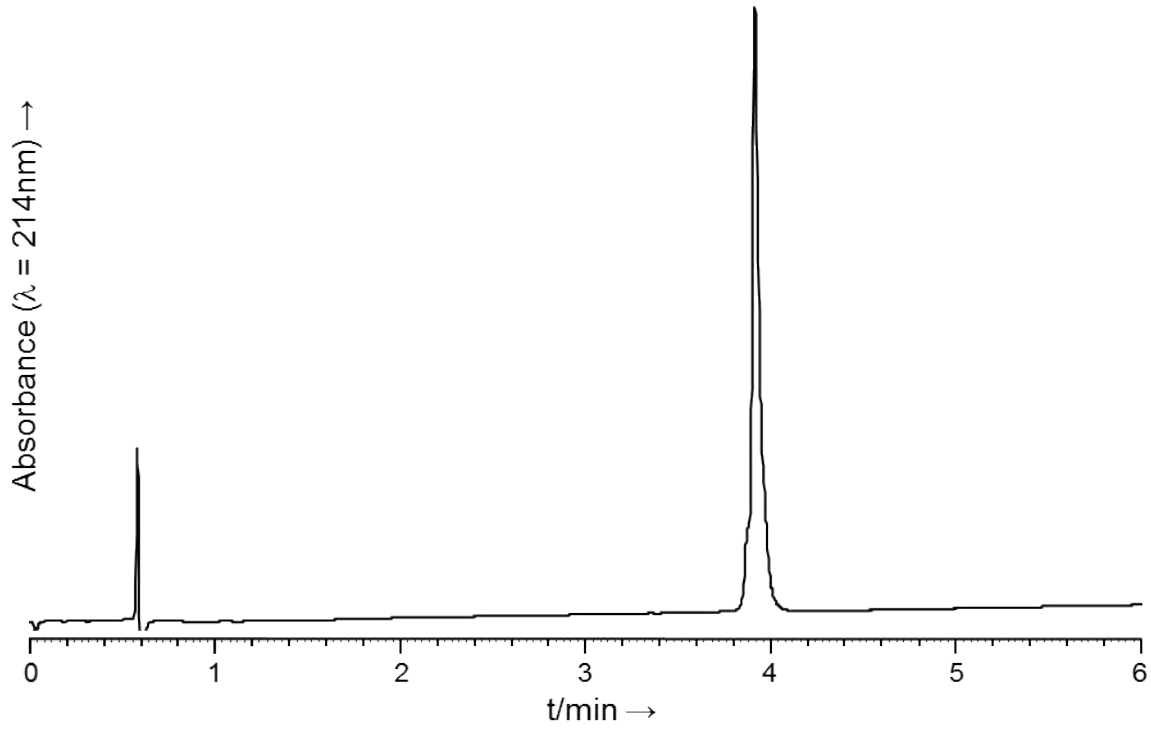
6a

ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{103}H_{154}N_{32}O_{32}S$: 2384.1128, found: 2384.1173.

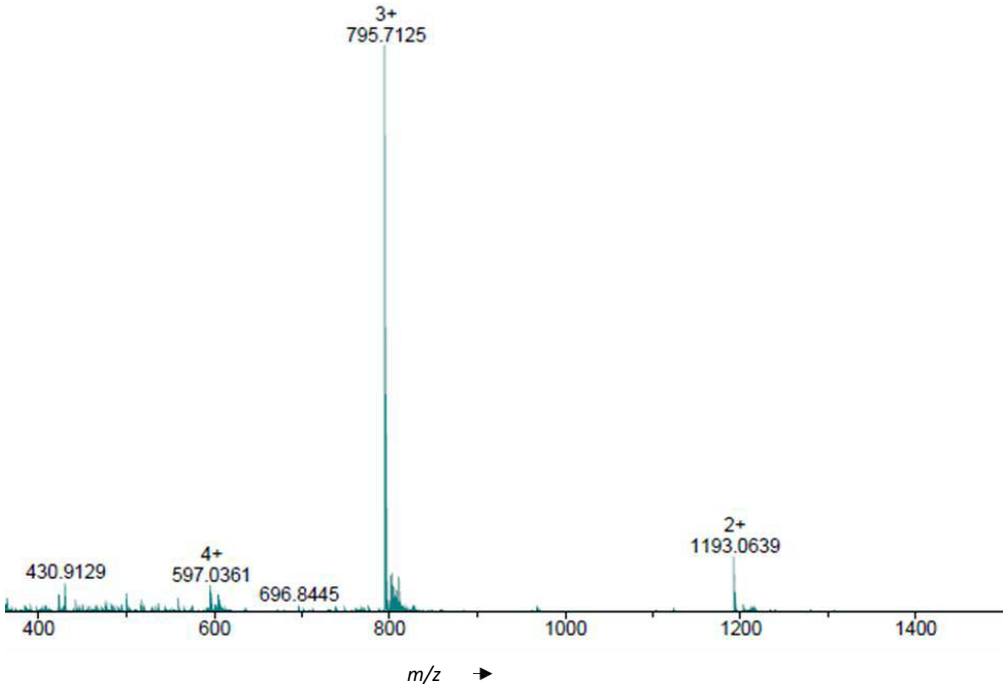
HPLC analysis: retention time: 2.12 min (Chromolith, gradient: 10-40 % B/A over 5 min).

HPLC purification: retention time: 8.05 min (Nucleosil, gradient: 24-31% B/A over 11 min, 55°C).

Yield: 91 % (8.3 µmol).

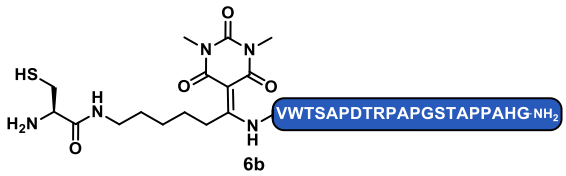


Supplementary figure 6: HPLC trace of pure **6a** at λ = 214 nm**.**



Supplementary figure 7: ESI-HRMS mass spectrum of **6a**.

Compound **6b**



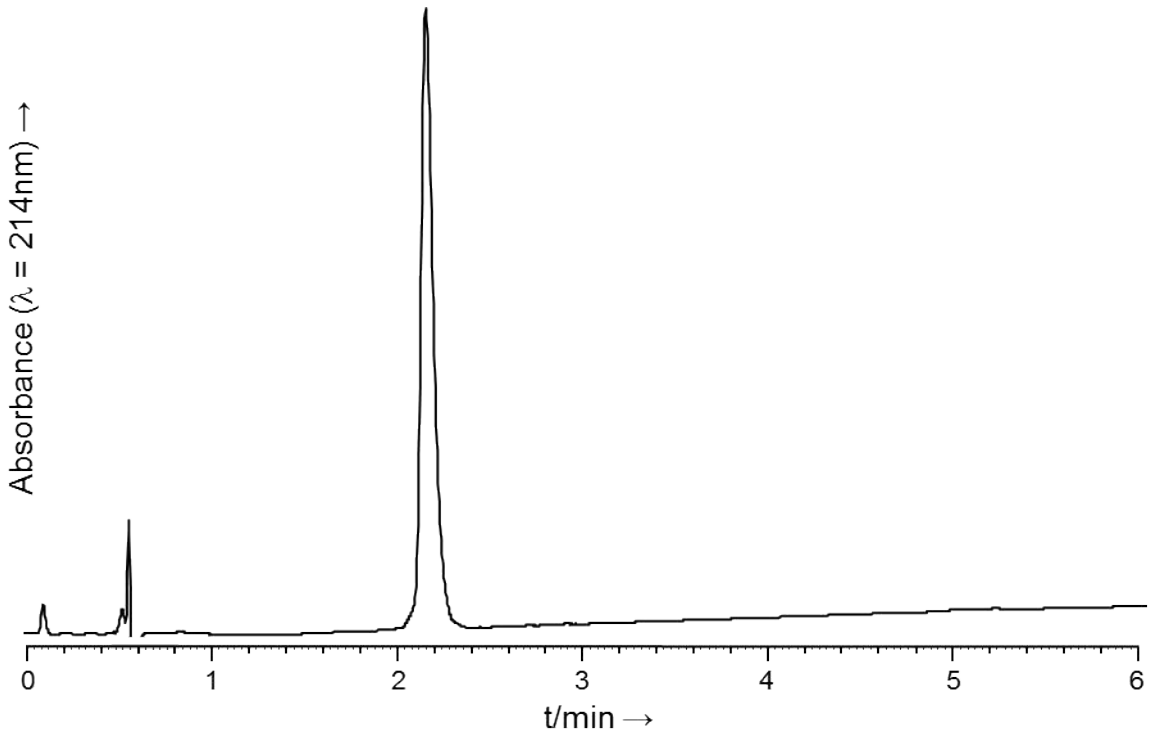
6b

ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{106}H_{160}N_{32}O_{32}S$: 2426.1597, found: 2426.1666.

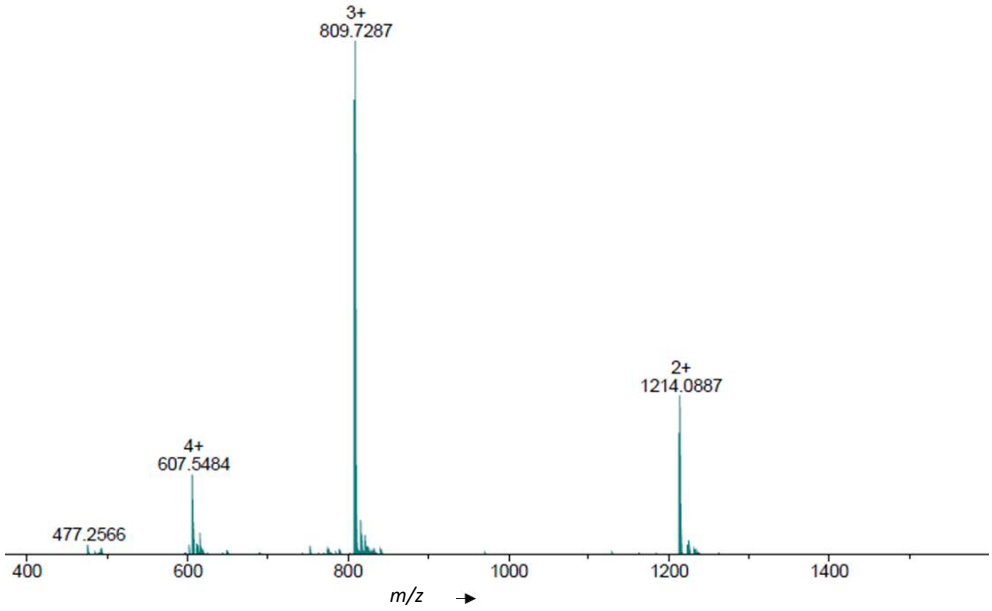
HPLC analysis: retention time: 2.88 min (Chromolith, gradient: 20-50 % B/A over 5 min).

HPLC purification: retention time: 2.1min (Nucleosil, gradient: 24-36 % B/A over 20 min, 55°C).

Yield: 53 % (4.4 µmol).

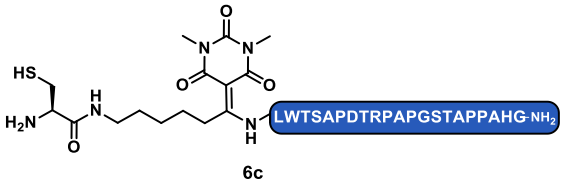


Supplementary figure 8: HPLC trace of pure **6b** at λ = 214 nm.



Supplementary figure 9: ESI-HRMS mass spectrum of **6b**.

Compound **6c**



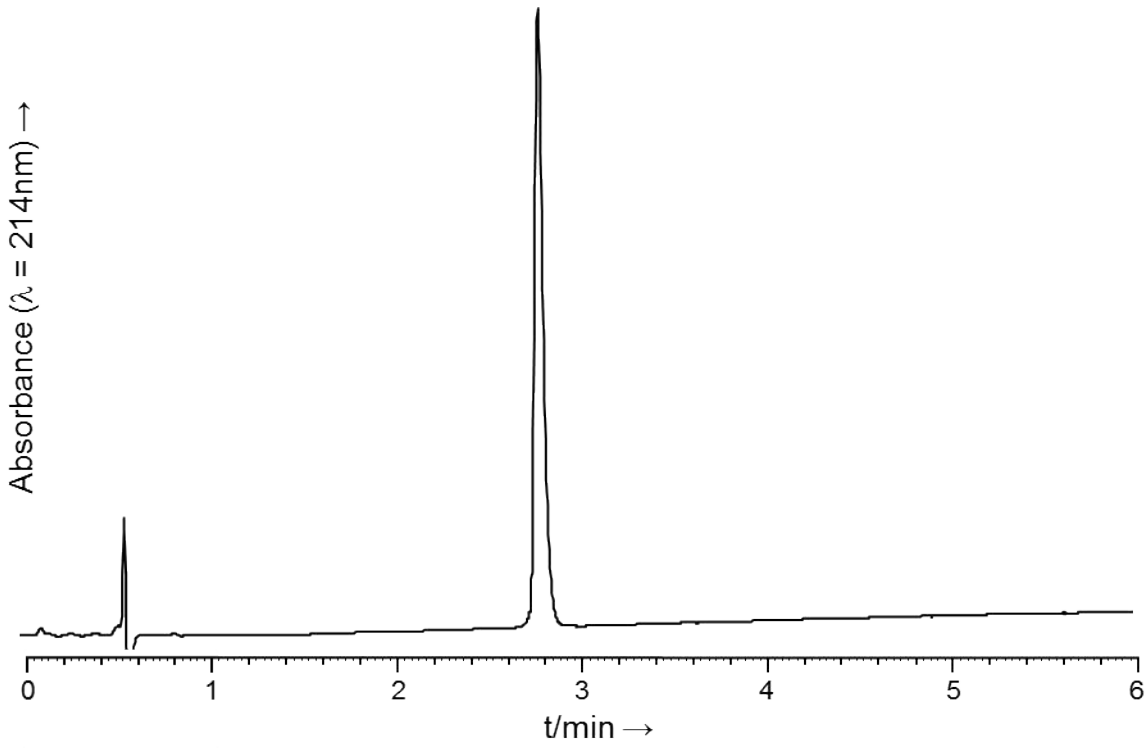
6c

ESI-HRMS (*m/z*): $[M+H]^+$calcd. for $C_{107}H_{162}N_{32}O_{32}S$: 2440.1754, found: 2440.1830.

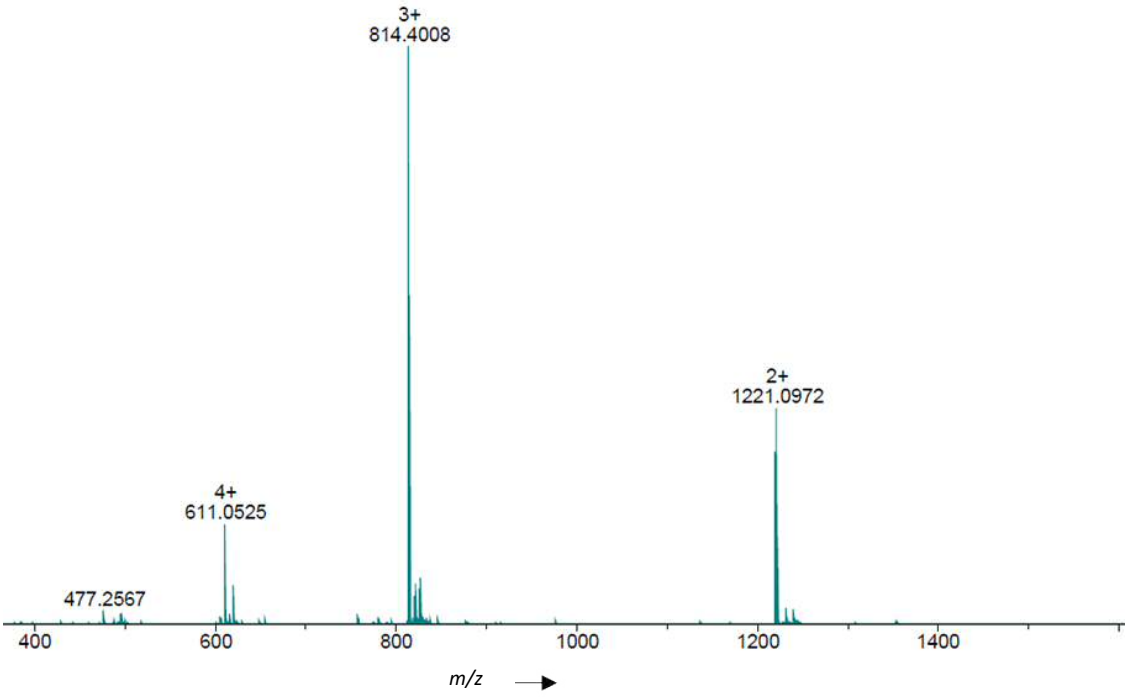
HPLC analysis: retention time: 2.88 min (Chromolith, gradient: 20-50 % B/A over 5 min).

HPLC purification: retention time: 14.29 min (Nucleosil, gradient: 24-36 % B/A over 20 min, 55°C).

Yield: 90 % (4.2 µmol).



Supplementary figure 10: HPLC trace of pure **6c** at λ = 214 nm.

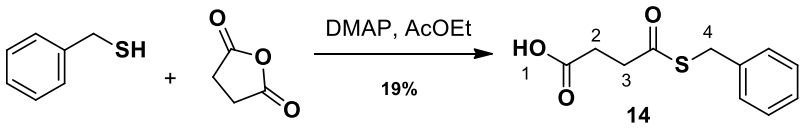


Supplementary figure 11: ESI-HRMS mass spectrum of **6c**.

3.4. Synthesis of acid/thioester heterobifunctional linkers

3.4.1. *S*-benzyl-thiosuccinic acid 14

Compound **14** was prepared following a protocol adapted from the literature.[3]



Supplementary figure 12: Synthesis of **14**.

Benzyl mercaptan (2.6 mL, 22 mmol, 1.1 equiv.) was added under argon to a stirred solution of succinic anhydride (2.0 g, 20 mmol, 1 equiv.) and 4-(dimethylamino)-pyridine (122 mg, 1 mmol, 0.05 equiv.) in 25 mL AcOEt. The resulting solution was stirred at room temperature for 3 h. The solution was then concentrated under reduced pressure. The product was then dissolved in 30 mL of a saturated aqueous sodium bicarbonate solution and the aqueous layer was washed with diethyl ether (2 x 10 mL). The aqueous phase was then cooled in an ice bath and acidified with 5 M hydrochloric acid to pH 2. The precipitate was filtered off,

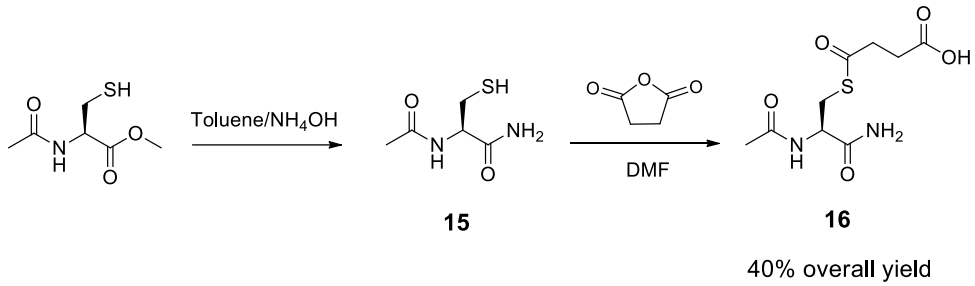
[3] Stetsenko, D. A.; Gait, M. J. *J Org Chem.* **2000**, *65*, 4900

washed with an ice-cold 0.1 M HCl solution, followed by ice-cold water, and then dissolved in 30 mL of CH_2Cl_2. The combined organic layers were dried over $MgSO_4$, filtered then concentrated under reduced pressure to yield **14** as a white powder (854 mg, 19%).

^{1}H NMR (600 MHz, DMSO-d_6): δ 12.28 (s, 1H, H_1), 7.37 - 7.17 (m, 5H, H_{Ar}), 4.13 (s, 2H, H_4), 2.84 (t, J = 7.3, 5.76 Hz, 2H, H_3), 2.54 (t, J = 7.3, 6.4 Hz, 2H, H_2).

^{1}H NMR spectrum was consisted with the literature data.[3]

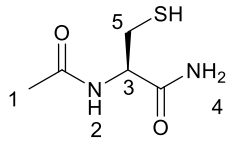
3.4.2. *S*-succinidyl-*N*-acetyl-L-cysteinamide 16



Supplementary figure 13: synthesis of **16**, overall yield 40%.

Preparation of compound *N*-acetylcysteinamide 15

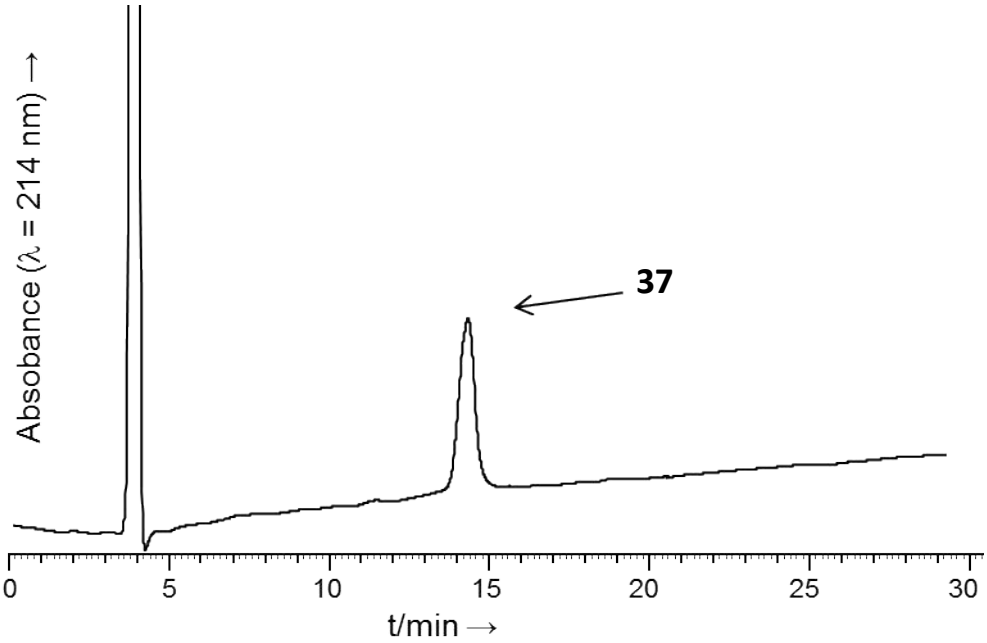
Compound **15** was prepared following a protocol adapted from the literature.[4]



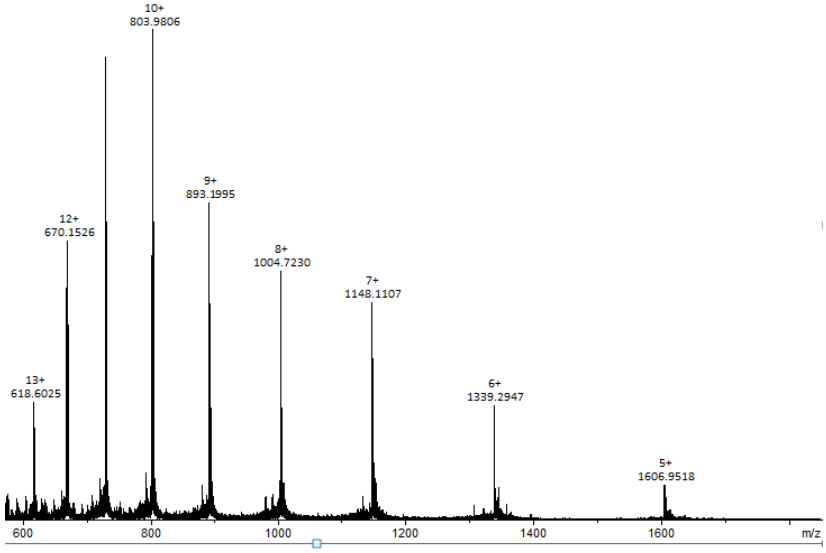
15

A solution of methyl *N*-acetyl-L-cysteinate (1 g, 1 equiv.) was added in a 1:1 mixture of toluene and aqueous ammonium hydroxide (56 mL) stirring for 48 hours under argon atmosphere. The reaction mixture was concentrated under reduced pressure to afford

[4] Bernardes, G. J.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D.; Davis, B. G. *Angew. Chem. Int. Ed. Engl.* **2008**, *47* (12), 2244.



Supplementary figure 93 : Analytical HPLC profile at λ = 214 nm of purified peptide **37**.



Supplementary figure 94 : ESI-HRMS mass spectrum of purified peptide **37**.

Conclusions and Perspectives

Conclusions and Perspectives

Disulfide rich peptides (DRPs) are highly constrained mini-proteins that show interesting pharmacological properties. Since 2000, three DRPs have reached the market, respectively for the treatment of chronic pain and constipation and many of them are currently in clinical and preclinical trials for new pharmacological applications such as treatment of inflammatory or cardiovascular diseases.

Most DRPs currently in clinical and preclinical trials have been produced through recombinant approaches. However, their preparation *via* chemical synthesis offers some advantages, in particular for the introduction of non-natural amino acids that can greatly improve their pharmacological properties and their selectivity towards a receptor subtype. For this reason, new methods for their chemical synthesis have been investigated in the last two decades.

DRPs up to 50 amino acids are, in general, accessible by standard SPPS. However, like other peptides, their HPLC purification can sometimes be highly challenging due to the presence of high amounts of acetylated truncated peptides. This problem is even exacerbated with DRPs, due to the large number of cysteine residues that are classically protected with a very bulky trityl group that slows down kinetics of coupling, leading to an increase in acetylated truncated peptides.

A very promising method that avoids these complicated purifications is referred to as the "*catch*-and-*release*" purification method. It is based on the use of an N-terminal linker which is able to selectively tag the target peptide and allow its immobilization onto a compatible solid support. This method benefits from a non-chromatographic purification step i.e., the unwanted truncated peptides formed during the synthesis are readily eliminated by washing the support while the expected peptide is still anchored on that support. A last step allows releasing the purified peptide

However, after a very detailed study of the existing N-terminal linkers described in the literature (gathered in a review article presented at the beginning of this thesis manuscript), we have realized that nowadays none of these linkers is compatible with the synthesis of DRPs. For this reason, we have decided to synthetize a new N-terminal linker, Boc-Cys(Trt)-Dtph-OH

ought to be completely compatible with the synthesis and *catch*-and-*release* purification of small- to medium-sized DRPs.

This linker has also been conceived to be adapted for the synthesis of longer DRPs, *via* successive solid phase NCLs in the N-to-C direction, which would thus also benefit from a "self-purification effect" and could avoid intermediate purification steps.

We have chosen the NCL reaction for the capture of the linker-containing peptides onto a water compatible solid support. For that purpose, the new N-terminal linker was synthesized with a cysteine moiety. Inspired by our previous work (N_3-Dtph-linker) and in order to allow the release of the peptide from the solid support, we introduced in the new linker a cleavable moiety (the DTPM) susceptible to nucleophilic attack by hydrazine or hydroxylamine.

To evaluate our linker stability and cleavage kinetics, that linker was selectively introduced in three model peptides (20 amino acids in length) presenting different N-terminal amino acids. We demonstrated that the Cys-Dtph linker is stable under a large range of conditions, including those required for the ligation steps. Besides, suitable conditions for the quantitative cleavage of our linker could be readily established.

During this work, a water-compatible thioester-functionalized solid support, which has been used to immobilize the peptide through the N-terminal linker, was also developed.

To validate our methodology, we decided to apply it to the very complex synthesis and purification of chicken Avian Beta Defensin 2 (AvBD2), a DRP which is 36 amino acids in length. The crude peptide obtained after its SPPS is highly contaminated with a large number of truncated peptides and is difficult to purify by standard HPLC due to the co-elution of the truncated peptides with the peptide of interest. In addition, AvBD2 presents a difficult sequence which results in the formation, during the SSPS, of several deletions. We have first optimized the synthesis of AvBD2 by testing different Fmoc deprotection conditions. Then, application of optimized NCL-mediated immobilization and Dtph cleavage conditions to AvBD2 furnished a peptide of good purity and yield, in comparison with this obtained after the Fmoc-SPPS of the entire peptide.

The non-chromatographic purification of AvBD2 constitutes the first example of the application of the *catch*-and-*release* purification method to unprotected cysteine-rich

peptides. In the previous reported methods, cysteines needed to be fully protected to avoid side reactions during both capture and release steps.

The method was also applied to the purification of a longer DRP, Bv8, which is 77 amino acids in length. Although we obtained that peptide with lower yields than AvBD2, our preliminary results have allowed confirming the validity of our strategy for the synthesis of medium length DRPs. Work is currently in progress in the team to apply the *catch*-and-*release* strategy for obtaining the Bv8 peptide in higher yield.

Having demonstrated the validity of our linker and strategy, further work in the team will be dedicated to the expansion of the methodology to successive solid phase NCLs in the N-to-C direction, in order to synthesize very long DRPs. For this purpose, we propose to utilize peptide segments containing a C-terminal thioester surrogate recently introduced by the team, *N*-Hnb-Cys.

Finally, the method proposed in this thesis and its future application to the synthesis of longer targets could open a new pathway to facilitate the high-throughput preparation of biologically-relevant peptides with important pharmacological properties.

Alba CASAS MORA

Développement d'une méthode de purification non-chromatographique de peptides riches en cystéine par immobilisation temporaire

Bien que la synthèse peptidique en phase solide (SPPS) soit maintenant une technique mature et très largement popularisée pour des peptides simples, certaines séquences restent encore compliquées à produire, comme les peptides riches en ponts disulfure (DRPs). Ces produits naturels, ligands sélectifs d'un grand nombre de cibles thérapeutiques, sont des outils pharmacologiques de premier ordre et sont considérés comme de bons candidats médicaments. La proportion importante de cystéines dans leur séquence (plus de 10%) leur confère des propriétés remarquables, mais limite aussi leur synthèse, conduisant à des purifications HPLC délicates, associées à des rendements faibles et une pureté médiocre.

Dans l'optique de simplifier la production de DRPs par voie chimique, notre but est de proposer une méthode de purification non-chromatographique. Pour cela, nous avons développé un bras N-terminal conçu pour être complètement compatible avec les peptides riches en cystéines: Boc-Cys(Trt)-1-{6-[1,3-diméthyl-2,4,6(1*H*,3*H*,5*H*)trioxopyrimidine-5-ylidène]hexyle}. A la fin de l'élongation sur support solide, ce bras est introduit sélectivement à l'extrémité N-terminale du peptide cible, sans réagir avec les peptides tronqués acétylés qui constituent les principales impuretés de la SPPS. Après coupure de la résine SPPS, le peptide cible est immobilisé sur un second support solide par le biais d'une réaction de ligation chimique native (NCL). Les peptides tronqués sont alors éliminés par simple filtration, puis le peptide cible est relargué en solution par coupure du bras N-terminal. Ayant comme objectif d'élargir par la suite notre stratégie à la synthèse de très longs DRPs *via* l'assemblage de multiples segments peptidiques par NCLs successives en phase solide, nous avons étudié en détail la stabilité et les conditions de coupure du bras.

La méthode a été appliquée à la purification de deux séquences naturelles riches en cystéines et biologiquement pertinentes : AvBD2 (36 AA), un peptide antimicrobien appartenant à la famille des β défensines aviaires et Bv8 (77 AA) un peptide d'amphibien de la famille des prokinéticines.

Mots-clés: Synthèse peptidique, ligation chimique native, peptides riches en cystéines, ligation chimique en phase solide.

A *catch*-and-*release* purification method to simplify the synthesis of cysteine-rich peptides

Although solid phase peptide synthesis (SPPS) is a mature and widely popularized technique for simple peptides, some sequences are still complicated to produce, such as disulfide-rich peptides (DRPs). These natural products are able to selectively bind a wide number of therapeutically relevant targets, hence they are considered as promising drug candidates and pharmacological tools. The large proportion of cysteines in their sequence (more than 10%) confers them remarkable properties, but also limits their synthesis, leading to delicate HPLC purifications associated with low yields and poor purity.

With the aim to simplify the chemical production of DRPs, we have developed an N-terminal linker: Boc-Cys(Trt)-1-{6-[1,3-dimethyl-2,4,6(1*H*,3*H*,5*H*)trioxopyrimidine-5-ylidene]hexyl}, which can be used for non-chromatographic *catch*-and-*release* purifications. It has been designed to be completely compatible with unprotected cysteine-containing peptides. Following solid phase elongation, this linker is selectively introduced at the N-terminus of the target peptide, leaving unreacted truncated acetylated peptides which are the main co-products of SPPS. After cleavage from the SPPS resin, the target peptide is immobilized on a second solid support through native chemical ligation (NCL). The truncated peptides are then removed by simple filtration. Cleavage of the linker finally releases the purified peptide into solution.

Having in mind the future extension our strategy to the synthesis of very long DRPs through successive solid-supported NCLs of multiple peptide segments, we have studied in detail the stability and cleavage conditions of the N-terminal linker.

To explore the scope and limitations of the method, it has been applied to the purification of two biologically-relevant cysteine-rich peptide sequences: chicken AvBD2 (36 AA), a β defensin antimicrobial peptide, and Bv8 (77 AA), a prokineticin-like peptide from yellow-bellied toad.

Keywords: Synthetic methods, native chemical ligation, cysteine-rich peptides, solid supported ligation.

Centre de Biophysique Moléculaire
CNRS UPR 4301
Rue Charles Sadron
45071 Orléans Cedex 2