



**HAL**  
open science

# Synthèse biomimétique et automatisée de peptides crypto-thioester pour la ligation chimique native : application à des peptides naturels riches en cystéine

Victor Terrier

► **To cite this version:**

Victor Terrier. Synthèse biomimétique et automatisée de peptides crypto-thioester pour la ligation chimique native : application à des peptides naturels riches en cystéine. Sciences agricoles. Université d'Orléans, 2016. Français. NNT : 2016ORLE2016 . tel-01431163

**HAL Id: tel-01431163**

**<https://theses.hal.science/tel-01431163v1>**

Submitted on 10 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT  
(SSBCV)**

Centre de Biophysique Moléculaire, Orléans

**THÈSE** présentée par :

**Victor TERRIER**

soutenue le : **29 janvier 2016**

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

Discipline : Chimie

**Synthèse biomimétique et automatisée de  
peptides crypto-thioester pour la ligation  
chimique native. Application à des peptides  
naturels riches en cystéine**

**THÈSE dirigée par :**

**Agnès DELMAS**

Directeur de recherche, CNRS Orléans

**RAPPORTEURS :**

**Didier BOTURYN**

Directeur de recherche, Université Joseph Fourier, Grenoble

**Oleg MELNYK**

Directeur de recherche, Institut Pasteur de Lille

---

**JURY :**

**Vincent AUCAGNE**

Chargé de recherche, CNRS Orléans

**Didier BOTURYN**

Directeur de recherche, Université Joseph Fourier, Grenoble

**Agnès DELMAS**

Directeur de recherche, CNRS Orléans

**Oleg MELNYK**

Directeur de recherche, Institut Pasteur de Lille

**Sylvain ROUTIER**

Professeur, Université d'Orléans

**Denis SERVENT**

Chercheur, CEA Saclay

**Gilles SUBRA**

Professeur, Université de Montpellier

*« Wish you were here » - Pink Floyd*

*A Frédérique,*





### Remerciements

Ce travail de thèse a été réalisé au centre de biophysique moléculaire (CBM), unité propre du Centre National de la Recherche Scientifique (CNRS). Naturellement, je voudrais remercier les financeurs des travaux présentés dans ce manuscrit : à commencer par la fondation Pierre Potier pour le développement de la chimie des substances naturelles, liée à l'académie des sciences française qui a financé ma bourse de thèse; je souhaiterais également remercier : le CNRS, l'agence nationale pour la recherche (ANR) ou encore l'association « vaincre la mucoviscidose », organismes financeurs des différents projets abordés lors de ma thèse.

Je voudrais remercier le Dr. Eva Jakab-Toth, directrice du CBM pour m'avoir permis d'effectuer ma thèse au sein de ce laboratoire.

J'adresse mes remerciements aux Drs. Didier Boturyn et Oleg Melnyk d'avoir accepté d'être les rapporteurs de ce manuscrit. Je remercie également le Pr. Sylvain Routier, le Dr. Denis Servent et le Pr. Gilles Subra d'avoir accepté de faire partie du jury de la soutenance de ma thèse.

Je souhaiterais remercier chaleureusement le Dr. Agnès Delmas, ma directrice de thèse, pour la formation riche qu'elle m'a dispensée et sa disponibilité. Je remercie également le Dr. Vincent Aucagne pour tout le temps qu'il m'a consacré ainsi que pour la rigueur et les nombreuses connaissances qu'il m'a transmises. Mes discussions avec Agnès et Vincent ont été toujours très enrichissantes et j'ai le sentiment d'avoir appris de nouvelles choses chaque jour passé au laboratoire.

J'adresse un grand merci à Dominique Lelièvre et à Phillipe Marceau, pour leurs conseils et aides techniques avisés ainsi que pour leur sympathie. Je ne peux pas oublier de remercier le Dr. Mathieu Galibert, post-doc dans l'équipe durant la majorité de ma thèse, je me rappellerai longtemps de ses conseils importants, de son soutien infaillible et de son humour. Je souhaite également remercier le Dr. Nathan Stanley ainsi qu'Alba Casas, Aline Mongis, Guillaume Jousset et les Drs. Véronique et Friedrich Piller, les membres de l'équipe que j'ai côtoyé pour nos échanges intéressants et leur amitié. Je souhaite également remercier Sandra Petinon, stagiaire de troisième année de licence que j'ai encadrée.

Je remercie les Drs. Cyril Colas et Guillaume Gabant pour les nombreuses analyses de spectrométrie de masse qu'ils ont réalisés pour mes composés.

Je remercie également tous mes amis Orléanais (collègues du labo ou non) qui m'ont fait apprécier cette ville, mes collègues des deux mandats que j'ai effectués au sein du bureau de l'association des doctorants du CBM ainsi que tous les membres du CBM que j'ai eu plaisir à côtoyer pendant ma thèse.

Enfin, je souhaiterais adresser un remerciement tout particulier à ma famille et à mes amis jurassiens (mais pas que...) pour leur soutien inconditionnel tout au long de ces 3 ans, ils connaissent l'importance qu'ils ont pour moi.

## Avant-propos

Ce manuscrit est construit à l'aide des articles originaux associés à mon travail de thèse. Pour en faciliter la lecture, les références bibliographiques et les figures sont propres à chaque chapitre, ainsi leur numérotation est réinitialisée au début des chapitres. Le référencement des composés apparaissant dans le corps des chapitres est énoncé par ordre alphabétique pour éviter la confusion avec la numérotation des composés apparaissant dans les articles, sauf dans le cas du chapitre 4 où l'ordre des composés cités est identique à celui de l'article; le référencement des composés est également réinitialisé au début des chapitres. Les références bibliographiques, citées hors des articles, sont rappelées sous forme d'une liste numérotée suivant leur ordre d'apparition dans le corps du chapitre concerné sur les pages de couleur ivoire (pour la version imprimée). Enfin, les parties expérimentales (*supporting information*) de chaque article sont insérées à la fin du manuscrit dans leur forme originale.

## Abréviations

Aa	Acide aminé
Ac	Groupement acétyle
Acm	Groupement acétamidométhyle
ADN	Acide désoxyribonucléique
AIBN	Azobisisobutyronitrile
Alk	Groupement alkyle
All	Groupement allyle
Ar	Groupement aryle
ARN	Acide ribonucléique
BAL	<i>Backbone amide linker</i>
Bn ou Bzl	Groupement benzyle
Boc	Groupement <i>tert</i> -butyloxycarbonyle
Boc-SPPS	Synthèse de peptides sur phase solide en stratégie <i>t</i> Boc/Bzl
CCK	<i>Cyclic cystine knot</i>
cDRP	<i>Cyclic disulfide-rich peptide</i>
CK	<i>Cystine knot</i>
CPC	Cystéinyl prolyl cystéine
CPE	Ester de cystéinyl proline
cO2	Cycloviolacine O2
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ène
Dbz	Bras 3,4-diaminobenzoyle
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
Dde	Groupement 1-(4,4-diméthyl-2,6-dioxacyclohexylidène)éthyle
DEAD	Azodicarboxylate de diéthyle
DIC	<i>N,N'</i> -Diisopropylcarbodiimide
DIEA	Diisopropyléthylamine
Dmb	Groupement 2,4-diméthoxybenzyle
DMF	<i>N,N</i> -Diméthylformamide
Dmmb	Groupement <i>N</i> -4,5-diméthoxy-2-mercaptobenzyle
DMSO	Diméthylsulfoxyde
DODT	2,2'-(Ethylendioxy)diéthanethiol
DRP	<i>Disulfide-rich peptide</i>
EPL	<i>Expressed protein ligation</i>
Equiv.	Equivalent
ESI	<i>Electrospray ionisation</i>
Et	Groupement éthyle
Fmoc	Groupement 9-fluorénylméthoxycarbonyle
Fmoc-SPPS	Synthèse de peptides sur phase solide en stratégie Fmoc/ <i>t</i> Bu
GFP	<i>Green fluorescent protein</i>
GlcNAc	<i>N</i> -Acétyl-D-glucosamine
GM2AP	<i>Ganglioside monosialic 2 protein activator</i>
GP / PG	Groupement protecteur
GPCR	Récepteurs couplés aux protéines G
Gu.HCl	Chlorhydrate de guanidinium

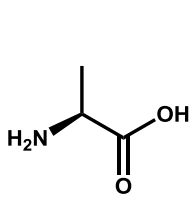
## Abréviations

HATU	Hexafluorophosphate de 2-(1 <i>H</i> -7-azabenzotriazol-1-yl)-1,1,3,3-tetraméthylamminium
HBTU	Hexafluorophosphate de 2-(1 <i>H</i> -benzotriazol-1-yl)-1,1,3,3-tetraméthylamminium
HCTU	Hexafluorophosphate de 2-(6-chloro-1 <i>H</i> -benzotriazole-1-yl)-1,1,3,3-tétraméthylamminium
HEPES	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
HGF	<i>Human growth factor</i>
Hmb	Groupement 2-hydroxy-4-méthoxybenzyle
Hnb	Groupement 2-hydroxy-5-nitrobenzyle
HOAt	1-Hydroxy-7-azabenzotriazole
HOBT	1-Hydroxybenzotriazole
HPLC	Chromatographie en phase liquide à haute performance
HRMS	Spectrométrie de masse haute à résolution
IL-8	Interleukine 8
<i>i</i> Pr	Groupement <i>iso</i> -propyle
KAHA	<i>Ketoacide-hydroxylamine ligation</i>
kB1	Kalata B1
KCL	Ligation sous contrôle cinétique ( <i>kinetically controled ligation</i> )
LC/HRMS	Chromatographie en phase liquide couplée à la spectrométrie de masse haute résolution
LC/MS	Chromatographie en phase liquide couplée à la spectrométrie de masse
MALDI-TOF	Ionisation/désorption assistée par une matrice / Temps de vol
Me	Groupement méthyle
MeNbz	<i>N</i> -Acyl- <i>N'</i> -méthylbenzimidazolone
MESNa	2-Mercaptoéthanesulfonate de sodium
MPA	Acide 3-mercaptopropionique
MPAA	Acide 4-mercaptophénylacétique
MS	Spectrométrie de masse
MT7	<i>Muscarinic toxin 7</i>
NBS	<i>N</i> -Bromosuccinimide
NBz	<i>N</i> -Acylbenzimidazolone
NCL	Ligation chimique native ( <i>native chemical ligation</i> )
NMP	<i>N</i> -Méthyl-2-pyrrolidinone
Pa1b	<i>Pea albumin 1 subunit b</i>
PAM	Phénylacétamidométhyle
Pbf	Groupement 2,2,4,6,7-pentaméthyl-dihydrobenzofuran-5-sulfonyle
Ph	Groupement phényle
PMT	2-Pyrrolidineméthanethiol
PyAOP	Hexafluorophosphate de 7-azabenzotriazol-1-yl-oxytripyrrolidinophosphonium
PyBOP	Hexafluorophosphate de benzotriazol-1-yl-oxytripyrrolidinophosphonium
RMN	Résonance magnétique nucléaire
RP-HPLC	Chromatographie liquide haute performance en phase inverse
RTD-1	<i>Rhesus <math>\theta</math>-defensin 1</i>
SAL	<i>Sugar assisted ligation</i>
SEA	<i>N,N</i> -Bis(2-sulfanyléthyl)amide
SEAlide	<i>N</i> -Sulfanyléthylanilide
SFTI-1	<i>Sunflower trypsin inhibitor 1</i>

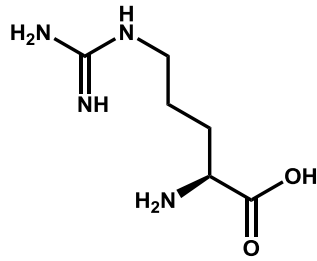
## Abréviations

SPPS	Synthèse de peptides sur phase solide
StBu	Groupement <i>tert</i> -butylsulfanyle
STL	<i>Serine/Threonine ligation</i>
<i>t</i> Bu	Groupement <i>tert</i> -butyle
TCEP.HCl	Chlorhydrate de <i>tris</i> (2-carboxyéthyl)phosphine
TEA	Thioéthylalkyl-amide
TFA	Acide trifluoroacétique
TFFH	Hexafluorophosphate de tetraméthylfluoroformamidinium
TIS	Triisopropylsilane
Trt	Groupeement triphénylméthyle (trityle)
UV	Ultra-violet
VA.044	Dihydrochlorate de 2,2'-azobis[2-(2-imidazol-2-yl)propane]
VIH	Virus de l'immunodéficience humaine
Xaa	Acide aminé quelconque
3FT	<i>Three-finger toxins</i>

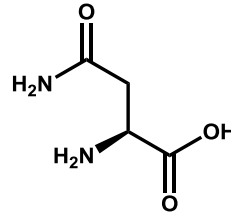
## Acides aminés canoniques



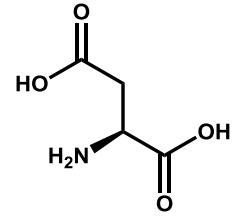
Alanine  
Ala - A



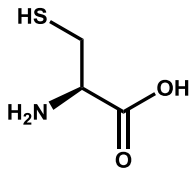
Arginine  
Arg - R



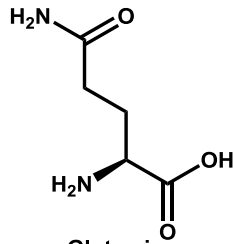
Asparagine  
Asn - N



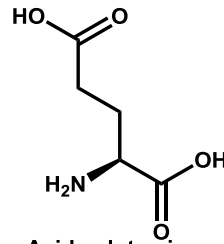
Acide aspartique  
Asp - D



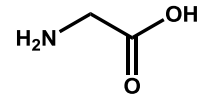
Cystéine  
Cys - C



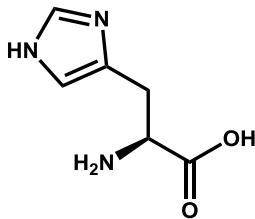
Glutamine  
Gln - Q



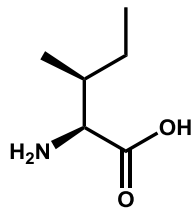
Acide glutamique  
Glu - E



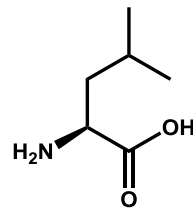
Glycine  
Gly - G



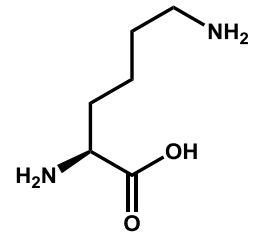
Histidine  
His - H



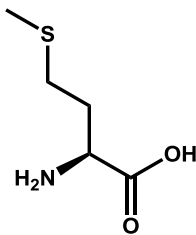
Isoleucine  
Ile - I



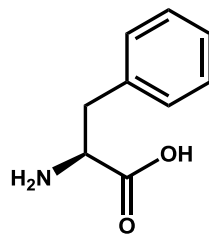
Leucine  
Leu - L



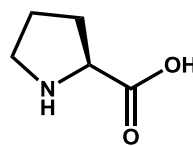
Lysine  
Lys - K



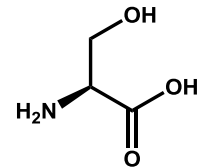
Méthionine  
Met - M



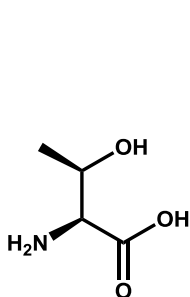
Phénylalanine  
Phe - F



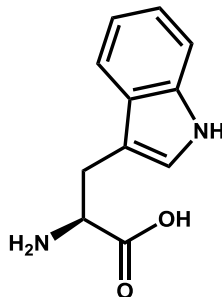
Proline  
Pro - P



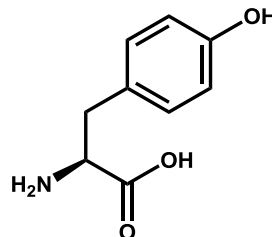
Serine  
Ser - S



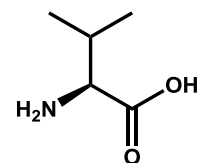
Thréonine  
Thr - T



Tryptophane  
Trp - W



Tyrosine  
Tyr - Y



Valine  
Val - V

## Sommaire

Remerciements.....	1
Avant-propos.....	2
Abréviations .....	3
Acides aminés canoniques.....	6
Introduction générale .....	11
Chapitre 1 : Etude Bibliographique .....	15
Partie 1 : La synthèse totale de mini-protéines par NCL.....	15
1.1 Intérêt de la synthèse de peptides et protéines par voie chimique .....	15
1.2 Développements historique de la synthèse chimique de peptides et protéines.....	16
1.2.1 Synthèse en solution.....	16
1.2.2 Synthèse peptidique sur phase solide .....	19
1.2.3 Approche hybride : couplage en solution de fragments protégés synthétisés par SPPS.....	22
1.3 La ligation chimique native (NCL).....	24
1.3.1 Le concept de ligation chimique.....	24
1.3.2 Travaux précurseurs .....	25
1.3.3 La ligation chimique native.....	27
1.3.4 Ligations chimiques natives successives.....	29
1.3.5 Développements pour la généralisation de la NCL et méthodes apparentées .	34
1.4 Autres stratégies de ligation native .....	44
1.4.1 Ligation de Staudinger native .....	44
1.4.2 Ligation native par capture sous forme d'imide puis transfert d'acyle $N \rightarrow N$ ...	45
1.4.3 Ligation KAHA.....	46
1.4.4 Ligation native par formation d'une oxazolidine .....	48
1.5 Synthèse de peptides thioester par stratégie Fmoc .....	49
1.5.1 Introduction de la fonction thioester sur un peptide protégé en solution .....	50
1.5.2 Greffage par une chaîne latérale .....	51
1.5.3 Greffage par un amide du squelette peptidique .....	52
1.5.4 Utilisation de <i>safety-catch linkers</i> .....	53
1.5.5 Peptides hydrazide comme précurseurs de thioester.....	55
1.5.6 Bras 3,4-diaminobenzoylé (Dbz) .....	56

## Sommaire

1.5.7 Bras pyroglutamylimide.....	59
1.5.8 Transfert d'acyle intramoléculaire $O \rightarrow S$ ou $N \rightarrow S$ .....	60
1.6 Bilan .....	81
Partie 2 : Les peptides et protéines riches en ponts disulfure : des cibles idéales pour la NCL.....	83
Références bibliographiques .....	89
Chapitre 2 : Développement d'une méthode simple et bio-inspirée de synthèse de peptides crypto-thioester par Fmoc-SPPS.....	103
Article 1 .....	115
Références bibliographiques .....	123
Chapitre 3 : Stratégie pour la synthèse de peptides comportant une cystéine C-terminale par ligation chimique native .....	127
Article 2.....	137
Références bibliographiques .....	141
Chapitre 4 : Synthèse de peptides cycliques naturels riches en cystéine.....	145
Article 3.....	151
Références bibliographiques .....	155
Conclusions et perspectives .....	159
Partie expérimentale article 1 : .....	165
Partie expérimentale article 2 : .....	262
Partie expérimentale article 3 : .....	292
Annexe : Note technique .....	325



# INTRODUCTION GENERALE



# Introduction générale

Les peptides et protéines sont impliqués dans la plupart des processus biologiques et montrent une très grande diversité de fonctions et de structures au sein des organismes vivants. L'étude de ces macromolécules est un champ très vaste de la recherche scientifique dans divers domaines comme la biologie et biochimie, les applications biomédicales ou encore la science des matériaux. Le développement de technologies permettant l'obtention de ces biomacromolécules est donc primordial. Faisant l'objet de nombreuses avancées et travaux récents, la synthèse de peptides et protéines par voie chimique apparaît comme complémentaire aux techniques de biologie moléculaire en surmontant certaines limitations des procédés basés sur l'ADN recombinant (limitation aux acides aminés naturels, contrôle difficile des modifications post-traductionnelles, difficulté de production de protéines cytotoxiques). La synthèse peptidique sur phase solide (SPPS) a permis une rupture technologique qui a facilité la production d'une grande variété de séquences peptidiques grâce à la rapidité et à la modularité de sa mise en œuvre. Néanmoins, la taille des peptides pouvant être produits par cette technique est restreinte (environ 50 acides aminés en routine) alors que certaines protéines d'intérêt biologique peuvent dépasser aisément la taille de cent acides aminés. Des méthodes permettant le couplage de segments peptidiques ont donc été développées : ce concept est nommé « ligation chimique ». La stratégie la plus utilisée actuellement est appelée « ligation chimique native » (NCL). Elle consiste en le couplage régio- et chimiosélectif, en milieu aqueux, de deux fragments peptidiques totalement déprotégés, équipés, respectivement, de deux groupements chimiques mutuellement réactifs : un thioester C-terminal et un  $\beta$ -amino thiol (une cystéine N-terminale). Le produit de ligation qui en résulte contient une liaison amide en tant que jonction entre les deux fragments. C'est dans ce contexte que se situe mon travail de thèse.

Après quelques rappels sur la synthèse de peptides et protéines par voie chimique, la première partie du **chapitre 1** de cette thèse décrit en détails la réaction de NCL : son origine ainsi que son mécanisme, ses variantes mais aussi ses limitations et les méthodologies développées récemment pour y répondre. En complément, d'autres techniques récentes de ligation native sont exposées dans cette partie. Enfin, une attention particulière sera portée à l'étude des méthodes existantes pour résoudre le problème de la synthèse délicate des peptides C $\alpha$ -thioester, partenaires clés de la NCL. Cette étude permet de contextualiser le sujet méthodologique de cette thèse.

Les peptides riches en ponts disulfure (DRP) sont des petites protéines (< 100 acides aminés), rencontrées dans la nature, suscitant un fort intérêt en chimie thérapeutique. Ces peptides dont les séquences contiennent de nombreuses cystéines (typiquement > 10 % des acides aminés de la séquence) sont des cibles idéales pour la NCL. La deuxième partie du **chapitre 1** décrit brièvement la biodiversité de ces peptides et donne un aperçu des familles majeures de DRPs naturels.

La synthèse délicate des peptides C $\alpha$ -thioester ralentit la démocratisation de la NCL dans les laboratoires. Le **chapitre 2** du manuscrit présente les travaux de conception et d'optimisation d'une nouvelle méthode bio-inspirée simple, générale et peu coûteuse pour la synthèse de précurseurs de peptides thioester directement utilisables en NCL. Nous avons réalisé une étude préliminaire du mécanisme de la NCL utilisant des peptides portant un nouveau dispositif de thioestérification et commencer à explorer les limites de son applicabilité. Nous avons également appliqué cette stratégie à la synthèse des formes réduites de deux longs DRPs (65 et 93 acides aminés). **[Publication 1]**

L'accessibilité synthétique à certains DRPs représente encore des défis pour la synthèse de peptides par voie chimique. Par exemple, la synthèse de peptides comportant une cystéine C-terminale est un cas difficile à cause de réactions secondaires intrinsèques aux méthodes utilisées historiquement. Le **chapitre 3** présente une solution que nous proposons à ce problème particulier. Profitant de la simplicité de mise œuvre de notre méthode de synthèse de précurseurs de peptides thioester, nous avons appliqué cette stratégie à la synthèse des formes réduites de deux peptides naturels comportant une cystéine acide C-terminale par introduction post-synthétique de la cystéine C-terminale par NCL. **[Publication 2]**

Les peptides cycliques riches en ponts disulfure (cDRP) ont un fort potentiel thérapeutique. Des méthodes robustes pour leur synthèse sont donc nécessaires pour permettre leur étude ou leur utilisation en tant que médicaments. Le **chapitre 4** de cette thèse décrit l'application de notre méthode à la cyclisation par NCL intramoléculaire du squelette peptidique de plusieurs DRPs naturels. **[Publication 3]**

Enfin, ce manuscrit se terminera par les conclusions et les perspectives que nous pouvons tirer de ce travail de thèse.

En annexe, une note technique est fournie, décrivant, point par point, le protocole à suivre pour appliquer notre méthode de synthèse de peptides thioester pour la réaction de NCL. **[Publication 4]**

# CHAPITRE 1



# Chapitre 1 : Etude Bibliographique

## Partie 1 : La synthèse totale de mini-protéines par NCL

### 1.1 Intérêt de la synthèse de peptides et protéines par voie chimique

La plupart des processus biochimiques intervenant au sein des organismes vivants sont régis par des macromolécules appelées « protéines ». Les protéines sont constituées de chaînes polypeptidiques caractérisées par une séquence d'acides aminés qui confère à celles-ci leur structure et leur fonction au sein du vivant. L'étude de ces macromolécules, permettant de comprendre et de moduler leurs activités biologiques, représente un thème important de la recherche scientifique notamment dans le domaine biomédical. Ce travail est rendu très laborieux par l'incroyable diversité des protéines. A titre d'exemple, le séquençage récent du génome humain a permis de dénombrer plus de 20000 gènes codant pour la séquence d'acides aminés de protéines.<sup>1,2</sup> De plus, l'épissage de l'ARN messager ou les modifications post-traductionnelles des acides aminés incorporés dans les chaînes polypeptidiques augmentent considérablement le nombre de protéines pouvant être exprimées par une cellule (formant le protéome). Dans le cas de l'être humain, ce nombre a été estimé à jusqu'à entre 100 000 et 1 000 000 de protéines distinctes.<sup>3</sup>

Des techniques de production de protéines par biologie moléculaire utilisent la technologie de l'ADN recombinant pour programmer des organismes tels que la bactérie *Escherichia coli* à exprimer une protéine de séquence définie. Ces méthodes ont largement été utilisées au cours des trente dernières années, permettant de grandes avancées dans la compréhension des fonctions biologiques des protéines ou dans leur utilisation à des fins thérapeutiques. En particulier, des études de relations structure/activité ont été réalisées en faisant varier la séquence en acides aminés de la protéine étudiée et en observant l'activité biologique de la protéine mutante qui en résulte. La production de protéines par voie recombinante s'est démocratisée au laboratoire mais son utilisation n'est pas encore universelle à cause de certaines limitations majeures. En effet, l'introduction de modifications post-traductionnelles (phosphorylation, glycosylation, ...) est techniquement difficile voire impossible à contrôler précisément, et l'introduction d'acides aminés non-encodés génétiquement est également complexe. De nombreux développements ont permis la possibilité d'ajout de près de 150 acides aminés non-canoniques au code génétique d'organismes, permettant la production

---

<sup>1</sup> International Human Genome Sequencing Consortium, *Nature*, **2001**, 409, 860-921.

<sup>2</sup> International Human Genome Sequencing Consortium, *Nature*, **2004**, 431, 931-945.

<sup>3</sup> Unwin R. D.; Gaskell S. J.; Evans C. A.; Whetton A. D., *Exp. Hematol.*, **2003**, 31, 1147-1159.

de protéines ayant des propriétés modulées ou nouvelles ou encore de permettre d'introduire des sondes pour leur suivi à un niveau cellulaire ou de l'organisme entier.<sup>4</sup> Bien qu'attrayantes, ces méthodes sont encore réservées à des spécialistes et ne sont pas utilisables en routine.

Grâce à des recherches intenses dans les dernières décennies, la synthèse de peptides et protéines par voie chimique se présente aujourd'hui comme une alternative convaincante à la voie biotechnologique. La structure des molécules produites par ce biais peut être contrôlée avec une précision atomique. De plus, des modifications non-naturelles (acides aminés non codés génétiquement, sondes pour la spectroscopie) peuvent être incorporées facilement aux protéines sans restriction de nombres ou de nature, ayant pour limite principale l'imagination des chimistes.<sup>5</sup> Ces possibilités repoussent ainsi les limites de l'étude structurale et fonctionnelle des protéines. La synthèse de peptides et protéines par voie chimique a été le sujet de développements constants depuis le début du vingtième siècle, chacune des avancées méthodologiques ayant permis d'augmenter progressivement la complexité des cibles synthétisables, tout en simplifiant l'accès. Ce domaine connaît encore une grande expansion de nos jours; le développement de nouvelles méthodes générales et facilement applicables pour la production de ces macromolécules reste nécessaire et continue à représenter un véritable défi pour les chimistes contemporains.

## 1.2 Développements historique de la synthèse chimique de peptides et protéines

### 1.2.1 Synthèse en solution

Fischer (prix Nobel de chimie en 1902) est considéré comme le pionnier de la synthèse chimique de peptides. C'est lui le premier, avec Hofmeister, qui a introduit le terme de « peptide » pour décrire un oligomère d'acides  $\alpha$ -aminés. Il a décrit, avec Fourneau, la synthèse d'un dipeptide, la glycyl-glycine, en 1901.<sup>6</sup> Les travaux de Fischer puis la découverte de groupements protecteurs réversibles pour la fonction amine d'un acide aminé,<sup>7</sup> ont pavé la voie vers la synthèse de petits peptides en solution. En 1953, du Vigneaud (prix Nobel de chimie 1955) et ses collègues ont notamment décrit le premier exemple de synthèse d'une hormone peptidique, l'ocytocine (9 acides aminés).<sup>8</sup>

---

<sup>4</sup> Dumas A.; Lercher L.; Spicer C. D.; Davis B. G., *Chem. Sci.*, **2015**, *6*, 50-69.

<sup>5</sup> Nilsson B. L.; Soellner M. B.; Raines R. T., *Annu. Rev. Biophys. Biomol. Struct.*, **2005**, *34*, 91-118.

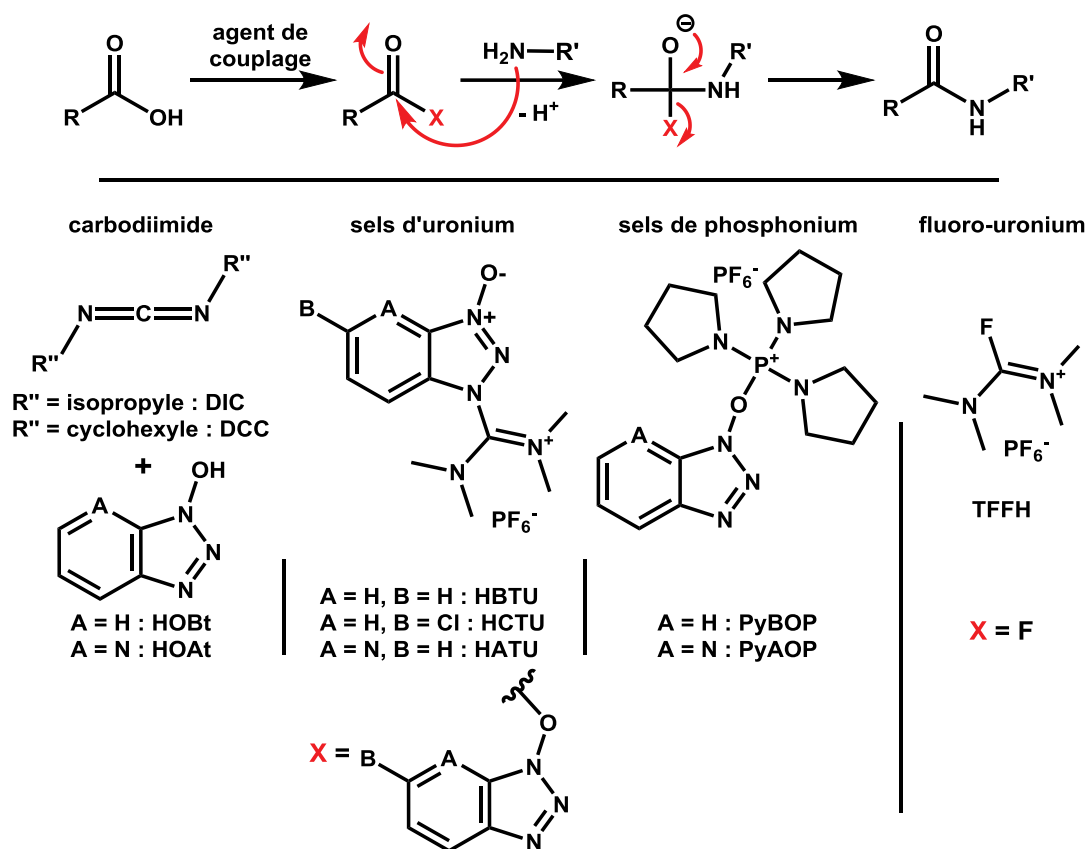
<sup>6</sup> Fischer E.; Fourneau E., *Ber. Deutsch. Chem. Ges.*, **1901**, *34*, 2868-2877.

<sup>7</sup> Bergmann M.; Zervas L., *Ber. Deutsch. Chem. Ges.*, **1932**, *65*, 1192-1201.

<sup>8</sup> Du Vigneaud V.; Ressler C.; Swan J. M.; Roberts C. W.; Katsoyannis P. G.; Gordon S., *J. Am. Chem. Soc.*, **1953**, *75*, 4879-4880.



De telles synthèses font appel à une stratégie de couplages successifs de petits fragments peptidiques (entre deux et dix résidus d'acides aminés) en milieu organique. Ces couplages consistent en la réaction entre un dérivé activé de la fonction carboxylique d'un acide aminé avec la fonction amine d'un autre acide aminé pour obtenir une liaison amide (figure 1). L'activation de l'acide carboxylique repose sur la substitution du groupement hydroxyle par un groupement électroattracteur constituant également un bon groupement partant (groupement X, figure 1). Les réactifs permettant une telle activation sont appelés « agents de couplage ».<sup>9</sup> De très nombreux agents de couplage ont été développés pour la synthèse peptidique; la figure 1 décrit quelques exemples couramment utilisés.

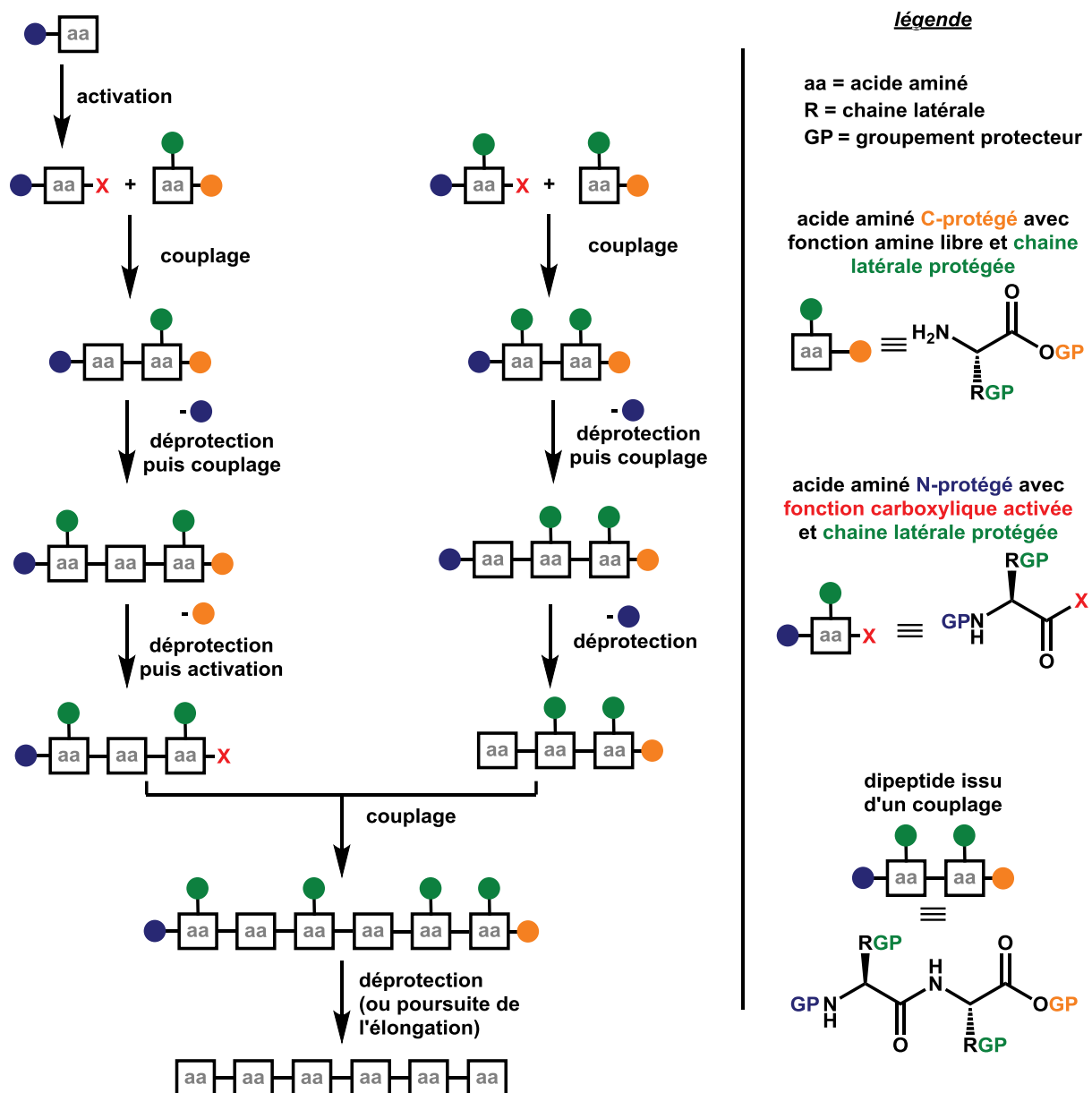


**Figure 1 :** Exemples d'agents de couplage couramment utilisés en synthèse peptidique pour l'activation d'un acide carboxylique <sup>9</sup>

Un point clé de la synthèse peptidique repose sur l'utilisation de plusieurs gammes orthogonales de groupements protecteurs réversibles pour masquer temporairement les fonctions n'intervenant pas dans la réaction de couplage. Lors d'une synthèse en solution, après un premier couplage entre deux acides aminés, la fonction amine du dipeptide formé est ensuite libérée pour permettre sa réaction avec la fonction carboxylique activée d'un troisième acide aminé ou d'un petit peptide produit par la même stratégie (figure 2). Les cycles

<sup>9</sup> El-Faham A.; Albericio F., *Chem. Rev.*, **2011**, *111*, 6557–6602.

déprotection/couplage sont ensuite répétés plusieurs fois et ce processus peut être réalisé de façon itérative et/ou convergente afin d'obtenir la séquence cible, puis tous les groupements protecteurs sont éliminés en fin de synthèse par un traitement chimique approprié.<sup>10</sup>

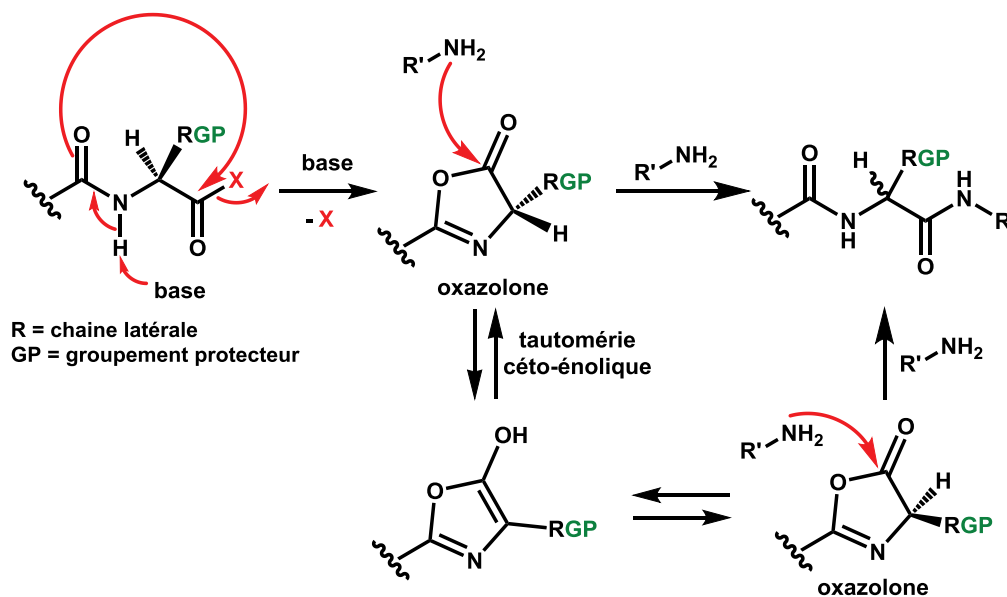


**Figure 2 :** Représentation schématique simplifiée d'une stratégie convergente de synthèse de peptides en solution <sup>10</sup>

Plusieurs exemples de préparation de longs peptides suivant cette stratégie de synthèse en solution ont été décrits.<sup>10</sup> Ces applications sont de véritables tours de force ayant nécessité un travail considérable et un nombre très important d'étapes de purification par chromatographie.

<sup>10</sup> Sakakibara S., *Biopolymers*, **1999**, 51, 279-296.

Cette stratégie de synthèse en solution est donc laborieuse et n'est viable pour l'accès rapide à une grande diversité de produits que pour de toutes petites séquences. De plus, l'approche est limitée par l'épimérisation possible de l'acide aminé C-terminal (hors glycine ou proline) d'un fragment peptidique ayant sa fonction carboxylique activée (figure 3). La voie majoritairement reconnue pour l'épimérisation est le passage par la formation d'une oxazolone suivie de l'inversion de la configuration du centre chiral par une tautomérisation céto-énolique (figure 3).<sup>9</sup>



**Figure 3 :** Mécanisme d'épimérisation d'un acide aminé activé en milieu basique <sup>9</sup>

Par ailleurs, la présence de groupements protecteurs limite grandement la solubilité des différents intermédiaires de synthèse, entraînant des réactions de couplage lentes donc parfois incomplètes, et rendant les étapes de purifications difficiles.<sup>11</sup>

### 1.2.2 Synthèse peptidique sur phase solide

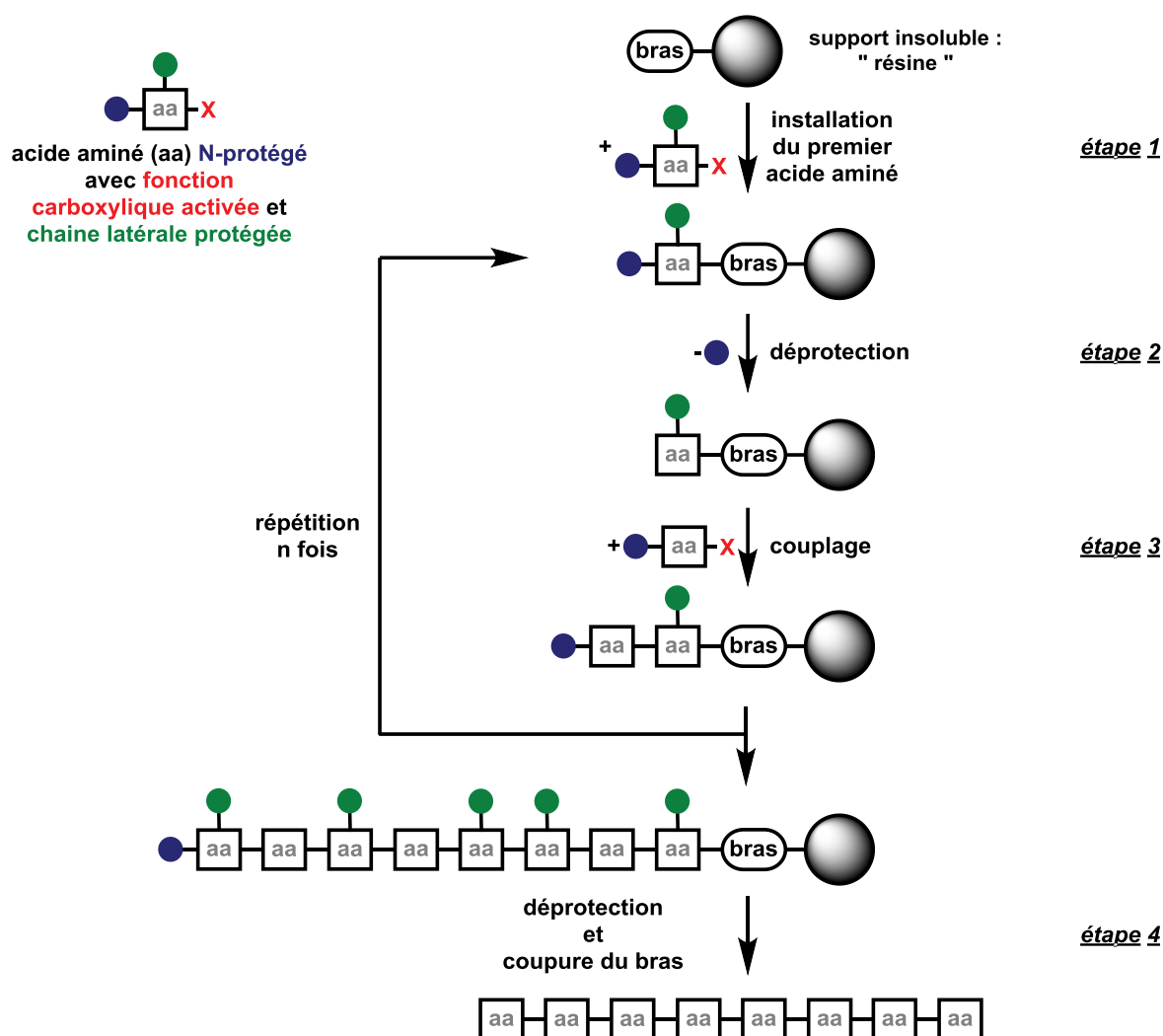
En 1963, Merrifield (prix Nobel de chimie en 1984) a mis au point la synthèse peptidique sur phase solide ou SPPS (figure 4) dans le but de limiter le nombre d'étapes de purification.<sup>12</sup> Le premier acide aminé est greffé sur un support solide polymérique insoluble dans le milieu réactionnel (appelé « résine ») par l'intermédiaire d'un lien covalent ou « bras » (étape 1, figure 4). L'assemblage de la chaîne peptidique est réalisé par répétition de cycle déprotection/couplage (étape 2 et 3, figure 4) dans le sens C-terminus vers N-terminus afin d'éviter l'épimérisation par formation d'oxazolone inhibée par la présence d'un groupement

<sup>11</sup> Kent S. B. H., *Annu. Rev. Biochem.*, **1988**, 57, 957-989.

<sup>12</sup> Merrifield R. B., *J. Am. Chem. Soc.*, **1963**, 85, 2149-2154.

carbamate plutôt qu'amide. En fin d'élongation, un traitement chimique permet d'éliminer les groupements protecteurs et de rompre le lien covalent liant le peptide au support solide (étape 4, figure 4).

Cette technique présente des avantages importants comparée à l'approche en solution : les réactifs et coproduits des réactions peuvent être éliminés facilement par simple filtration. Ceci permet l'emploi de larges excès de réactifs, afin de conduire à des rendements quasi-quantitatifs à chaque étape. Quelques années après l'introduction de la SPPS, Merrifield a décrit un instrument permettant l'automatisation complète de cette stratégie.<sup>13</sup> La SPPS a révolutionné la synthèse peptidique par voie chimique.



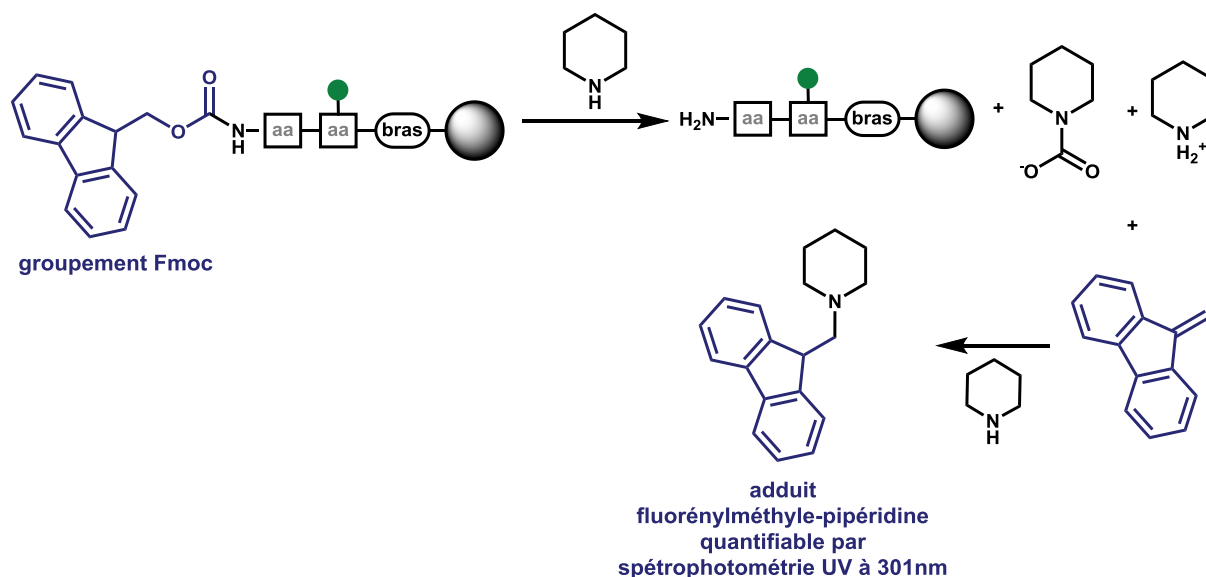
**Figure 4 :** Stratégie de synthèse peptidique sur support solide <sup>12</sup>

La méthode développée par Merrifield, qui s'est rapidement imposée, est appelée stratégie Boc/Bzl; elle consiste en l'utilisation de groupements acido-labiles: *tert*-

<sup>13</sup> Merrifield R. B.; Stewart J. M.; Jernberg N., *Anal. Chem.*, **1966**, 38, 1905-1914.

butoxycarbonyle (Boc) pour protéger les amines N-terminales, et groupements benzyles (Bzl) et dérivés pour la majorité des fonctions des chaînes latérales. L'amine N-terminale est libérée à chaque cycle à l'aide d'acide trifluoroacétique (TFA). La coupure du bras et l'élimination des groupements protecteurs des chaînes latérales en fin de synthèse est réalisée par traitement au fluorure d'hydrogène (HF). L'HF est un composé volatil (point d'ébullition = 20 °C), très toxique et corrosif capable de dissoudre le verre, sa manipulation demande donc des précautions importantes. De plus, les conditions acides utilisées pour cette stratégie (TFA à chaque cycle et HF) sont incompatibles avec la synthèse de peptides portant des modifications post-traductionnelles acido-sensibles (comme les glycosylations ou les phosphorylations).

Introduite par Sheppard en 1978, la stratégie de SPPS Fmoc/tBu est aujourd'hui la stratégie la plus utilisée (en 2003, une étude comptabilisait 71 % d'articles faisant état de l'utilisation de la stratégie Fmoc/tBu parmi les articles à propos de SPPS publiés dans le *Journal of Peptide Science*).<sup>14</sup> Pour cette méthode, le groupement baso-labile 9-fluorénylméthoxycarbonyle est utilisé pour protéger les amines N-terminales (figure 5). Ce groupement peut être éliminé à l'aide d'un traitement à la pipéridine qui est une base faible et nucléophile. Un avantage de la stratégie est la possibilité de quantification de l'efficacité d'un cycle couplage/déprotection; par exemple, par dosage de l'adduit fluorénylméthyle-pipéridine libéré au cours du traitement par spectrophotométrie UV à 301 nm.



**Figure 5 :** Elimination du groupement protecteur Fmoc N-terminal à l'aide de pipéridine

<sup>14</sup> Sheppard R., *J. Pept. Sci.*, **2003**, 9, 545-552.

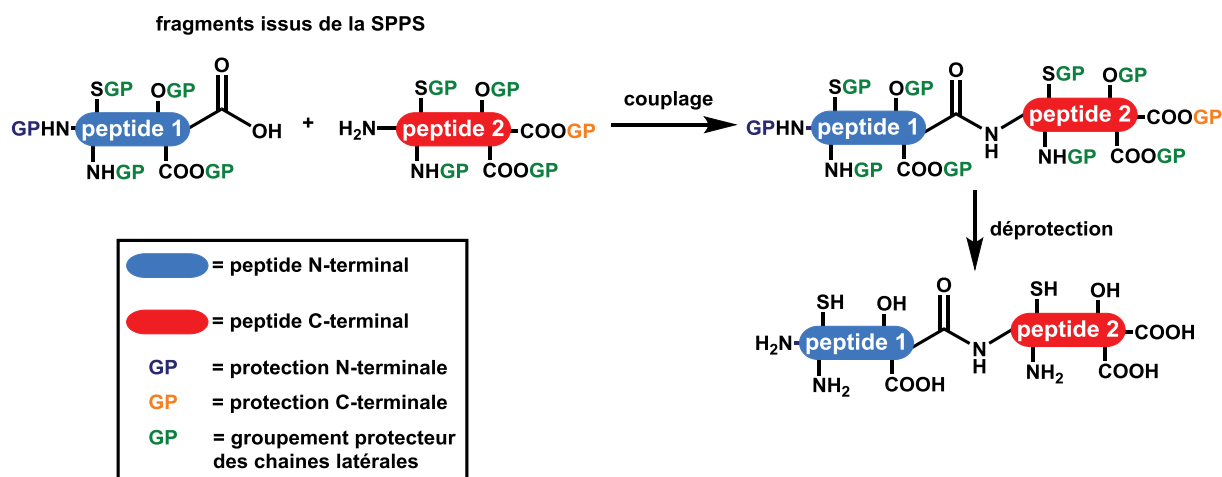
Grâce à de nombreux développements et optimisations (en termes de supports solides,<sup>15</sup> bras,<sup>16</sup> groupements protecteurs,<sup>17</sup> agents de couplage<sup>9</sup> ou synthétiseurs), la SPPS s'est imposée dans les laboratoires et a permis la synthèse de très nombreuses cibles. Néanmoins, malgré quelques exceptions pour lesquelles la stratégie a été poussée à ses limites pour obtenir des protéines de plus de 100 acides aminés, cette méthode est restreinte, en routine, à l'obtention de peptides d'une taille d'une cinquantaine d'acides aminés.<sup>18</sup> Au-delà de cette taille, l'accumulation de co-produits indésirables principalement issus de couplages incomplets conduit à des rendements trop faibles et à une grande difficulté à isoler le produit désiré par chromatographie.

### 1.2.3 Approche hybride : couplage en solution de fragments protégés synthétisés par SPPS

L'évolution logique pour augmenter la taille des protéines synthétisables facilement par voie chimique pourrait être une approche hybride convergente de couplage en solution de fragments peptidiques protégés de taille moyenne synthétisés aisément par SPPS.<sup>18</sup>

#### a) Couplage de fragments peptidiques protégés

Suivant le principe de la synthèse en solution (vu en 1.2.1), deux fragments préparés par SPPS ayant leurs chaînes latérales complètement protégées peuvent être condensés par couplage peptidique en milieu organique pour obtenir une chaîne plus longue; le produit de couplage est ensuite déprotégé (figure 6).



**Figure 6 :** Principe de condensation de fragments peptidiques totalement protégés

<sup>15</sup> Garcia Martin F.; Albericio F., *Chimica Oggi*, **2009**, 26, 29-34.

<sup>16</sup> Góngora-Benítez M.; Tulla-Puche J.; Albericio F., *ACS Comb. Sci.*, **2013**, 15, 217-228.

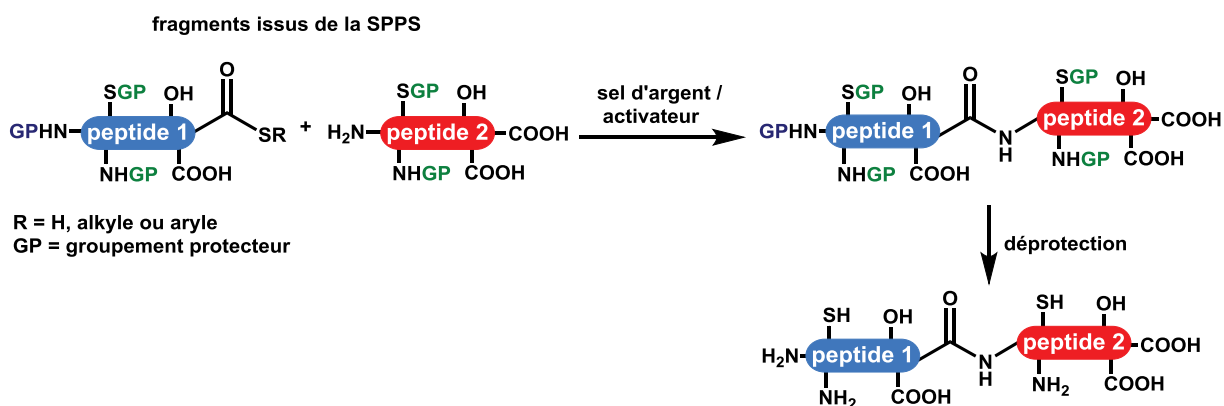
<sup>17</sup> Isidro-Llobet A.; Alvarez M.; Albericio F., *Chem. Rev.*, **2009**, 109, 2455-2504.

<sup>18</sup> Albericio F., *Curr. Opin. Chem. Biol.*, **2004**, 8, 211-221.

Quelques cas remarquables de synthèses *via* cette approche ont été décrits; par exemple, la *green fluorescent protein* (GFP) de la méduse *Aequorea victoria* (238 acides aminés) a été synthétisée en 1998 par couplages successifs de 26 fragments.<sup>19</sup> Cet exemple impressionnant, étant donné la longueur de la chaîne obtenue, a aussi permis d'étudier la formation du chromophore de la GFP. Cependant, cette stratégie possède les inconvénients inhérents à l'utilisation de fragments peptidiques protégés (solubilisation difficile des fragments, réactions lentes et possibilité d'épimérisation de l'acide aminé activé lors du couplage).

### b) Couplage de fragments peptidiques partiellement protégés

Blake et ses collaborateurs, en 1981,<sup>20</sup> puis Hojo et Aimoto,<sup>21</sup> en 1991, ont décrit une méthode de condensation de fragments peptidiques issus de la SPPS qui ne sont que partiellement protégés, appelée « méthode thioester ». Cette approche semi-chimiosélective repose sur l'activation du groupement thiocarboxylate ou thioester C-terminal d'un peptide par des ions argent permettant l'acylation sélective de l'amine N-terminale libre d'un autre peptide. La régiosélectivité de la réaction est garantie par la protection des fonctions amines et thiols des fragments mis en jeu. En revanche, les fonctions des chaînes latérales des résidus Asn, Asp, Arg, Glu, Gln, His, Ser, Thr, Trp et Tyr peuvent être libres (figure 7).<sup>22</sup>



**Figure 7 :** Principe de condensation de fragments peptidiques partiellement protégés<sup>20-22</sup>

L'utilisation de fragments possédant la majorité des fonctions de leurs chaînes latérales libres présente le grand avantage d'améliorer nettement la solubilité de ces fragments dans des solvants polaires (eau, DMF, DMSO, ...) comparée à des fragments totalement protégés. Néanmoins, les groupements protecteurs des fonctions amines et thiols ne confèrent pas à ces

<sup>19</sup> Nishiuchi Y.; Inui T.; Nishio H.; Bodi J.; Kimura T.; Tsuji F. I.; Sakakibara S., *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 13549-13554.

<sup>20</sup> Blake J.; Li C. H., *Proc. Natl. Acad. Sci. USA*, **1981**, *78*, 4055-4058.

<sup>21</sup> Hojo H.; Aimoto S., *Bull. Chem. Soc. Jpn*, **1991**, *64*, 111-117.

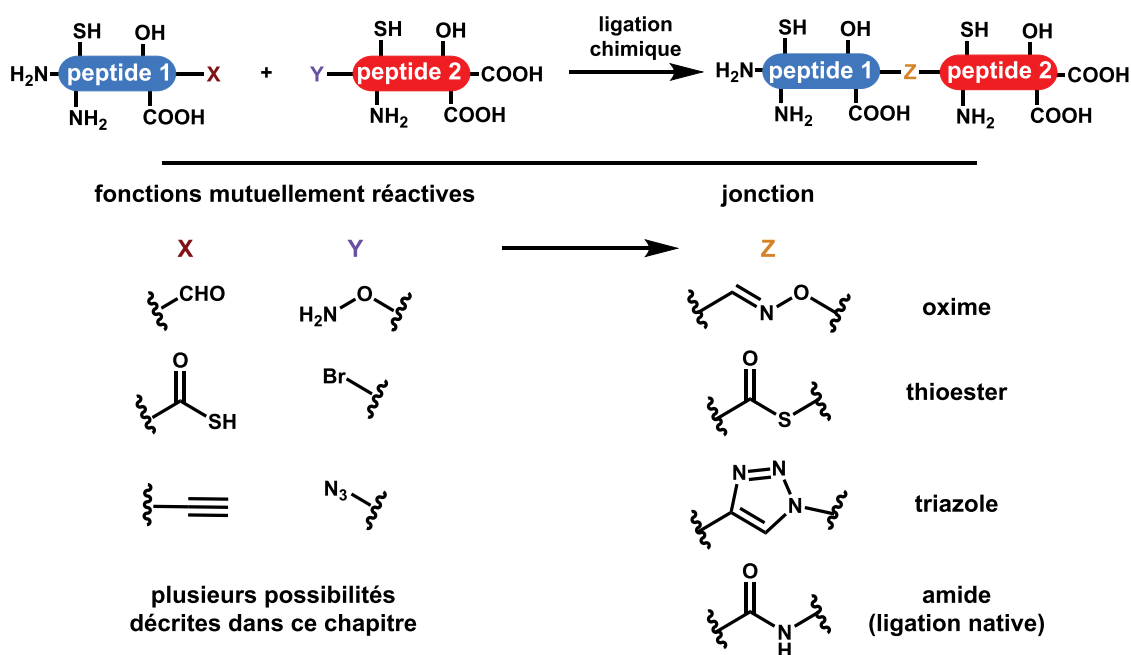
<sup>22</sup> Sato T.; Aimoto S., *Tetrahedron Lett.*, **2003**, *44*, 8085-8087.

fragments une solubilité optimale, demandant des temps de réactions prolongés.<sup>23</sup> De plus, par cette approche, le risque d'épimérisation de l'acide aminé C-terminal du fragment thiocarboxylate ou thioester est diminué contrairement à un couplage de peptides complètement protégés mais ce risque n'est pas complètement évité. Malgré cela, cette méthode et d'autres méthodes apparentées font encore l'objet de recherches riches de nos jours.<sup>22,24</sup>

### 1.3 La ligation chimique native (NCL)

#### 1.3.1 Le concept de ligation chimique

En 1992, Schnolzer et Kent ont introduit un nouveau concept consistant en la réaction complètement régio- et chimiosélective entre deux fragments peptidiques munis de fonctions mutuellement réactives. La chimiosélectivité de la réaction permet d'utiliser des fragments complètement déprotégés, évitant ainsi la nécessité d'une étape de déprotection finale. De plus, les fragments déprotégés montrent une meilleure solubilité que leurs équivalents protégés, ce qui facilite leur manipulation. Cette approche, qui a révolutionné la synthèse de protéine par voie chimique, a été nommée « ligation chimique » (figure 8).<sup>25</sup> Le terme « ligation » s'inspire des propriétés des ligases, des enzymes qui catalysent la formation de liaisons covalentes entre deux biomolécules.



**Figure 8 :** Principe de la ligation chimique de peptides déprotégés N.B. : par convention, sauf exception, les fonctions libres des chaînes latérales ou extrémités N- et C-terminales ne sont pas représentées sur les segments peptidiques dans les figures de la suite de ce manuscrit

<sup>23</sup> Aimoto S., *Biopolymers*, **1999**, 51, 247-265.

<sup>24</sup> Roberts A. G.; Johnston E. V.; Shieh J. H.; Sondey J. P.; Hendrickson R. C.; Moore M. A. S.; Danishefsky S. J., *J. Am. Chem. Soc.*, **2015**, 137, 13167-13175.

<sup>25</sup> Schnolzer M.; Kent S. B. H., *Science*, **1992**, 256, 221-225.



Le développement de ces réactions chimio- et régiosélectives est un véritable challenge pour les chimistes des peptides et de nombreuses stratégies répondant à ce principe ont émergé; l'utilisation de ces stratégies conduit à la formation de différentes fonctions à la jonction entre deux peptides (figure 8). La suite de cette étude bibliographique est volontairement limitée aux ligations dites « natives » formant une liaison peptidique, car les autres stratégies sont en dehors du champ de cette thèse.

### 1.3.2 Travaux précurseurs

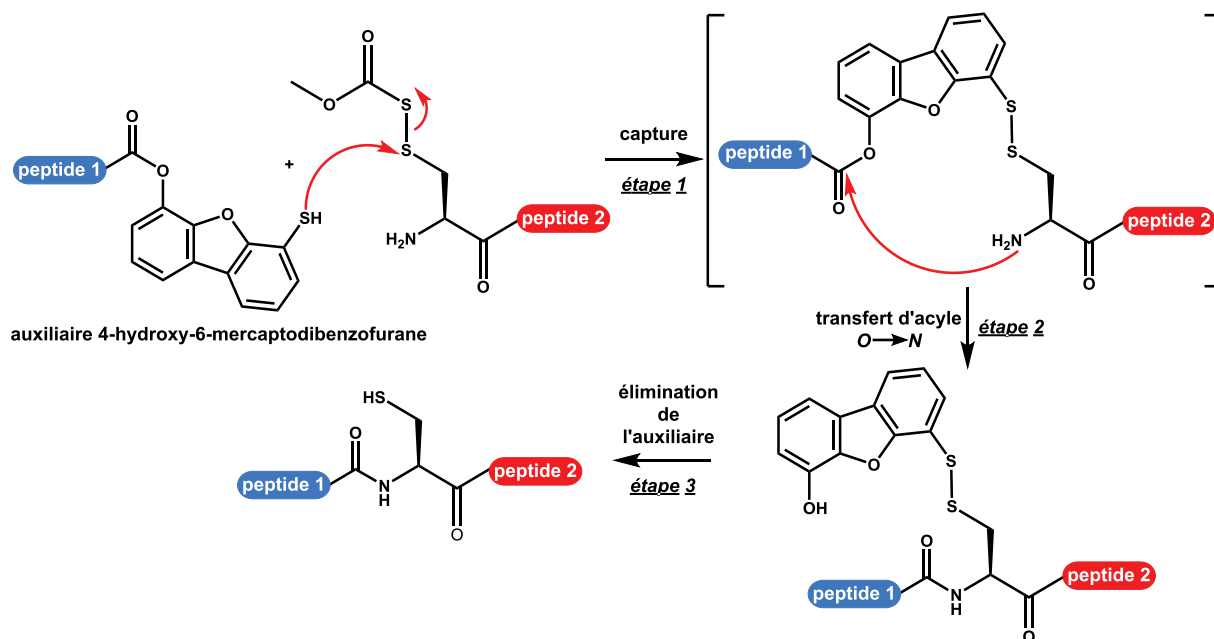
#### a) Stratégie de Kemp

Les premiers exemples de ligation native de peptides complètement déprotégés ont été décrits par Kemp et ses collaborateurs au début des années quatre-vingts. La stratégie repose sur l'utilisation d'un groupement auxiliaire pour rapprocher des fragments peptidiques par capture chimiosélective, les plaçant dans une géométrie favorable à leur réaction intramoléculaire. De nombreux auxiliaires ont été décrits dont par exemple l'utilisation d'un motif 4-hydroxy-6-mercaptodibenzofurane (figure 9).<sup>26,27</sup> Dans ce cas, la capture est réalisée par un échange efficace entre le groupement thiol de l'auxiliaire présent sur le fragment N-terminal et un pont disulfure mixte présent du côté N-terminal du fragment partenaire (figure 9, étape 1). Le rapprochement des fonctions réactives favorise ainsi un réarrangement par transfert d'acyle  $O \rightarrow N$  passant par un état de transition relativement rigide à 12 chaînons pour conduire à la formation de la liaison amide (figure 9, étape 2). La fonction ester de l'auxiliaire étant seulement faiblement activée, le risque d'épimérisation de l'acide aminé au point de couplage est diminué par rapport à une activation par un agent de couplage classique. Après la formation de la liaison amide, l'auxiliaire est éliminé par une simple réduction du pont disulfure, libérant ainsi une cystéine (figure 9, étape 3). Il est à noter que cette approche n'est pas complètement chimiosélective car incompatible avec la présence de cystéines dans les fragments peptidiques mis en jeu.

---

<sup>26</sup> Kemp D. S.; Galakatos N. G., *J. Org. Chem.*, **1986**, *51*, 1821-1829.

<sup>27</sup> Kemp D. S.; Galakatos N. G.; Bowen B.; Tan K., *J. Org. Chem.*, **1986**, *51*, 1829-1838.



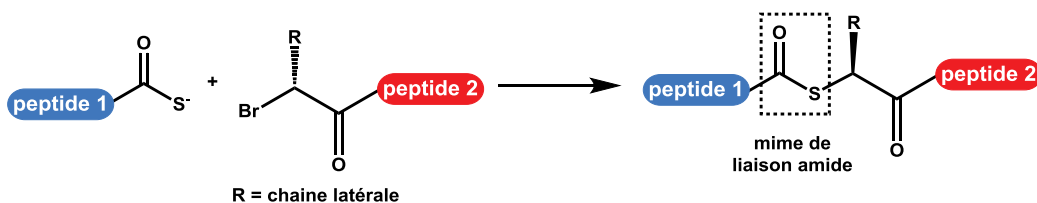
**Figure 9 :** Ligation native de peptides déprotégés assistée par un groupement dibenzofurane <sup>26,27</sup>

La stratégie de Kemp est considérée comme le point de départ conceptuel des méthodes modernes de ligation native. Toutes les stratégies introduites depuis cette approche utilisent les mêmes deux premières étapes :

- une étape de capture opérée par une réaction chimiosélective pour créer un lien covalent entre les deux fragments peptidiques
- un réarrangement qui crée la liaison amide.<sup>5</sup>

### b) Ligation thioester peptidomimétique

Au cours de leurs recherches de nouvelles stratégies convergentes de synthèse peptidique, Schnolzer et Kent ont développé, en 1992, une méthode chimiosélective de couplage de peptides complètement déprotégés par une réaction, en phase aqueuse, de substitution nucléophile entre un groupement thiocarboxylate et un groupement bromoacétyle conduisant à la formation d'une liaison thioester (figure 10, le terme « ligation chimique » a été introduit lors de la description de cette méthode).<sup>25</sup> A noter qu'à pH acide, cette technique de ligation est compatible avec la présence de cystéines libres qui ne sont que très faiblement réactives dans ces conditions.

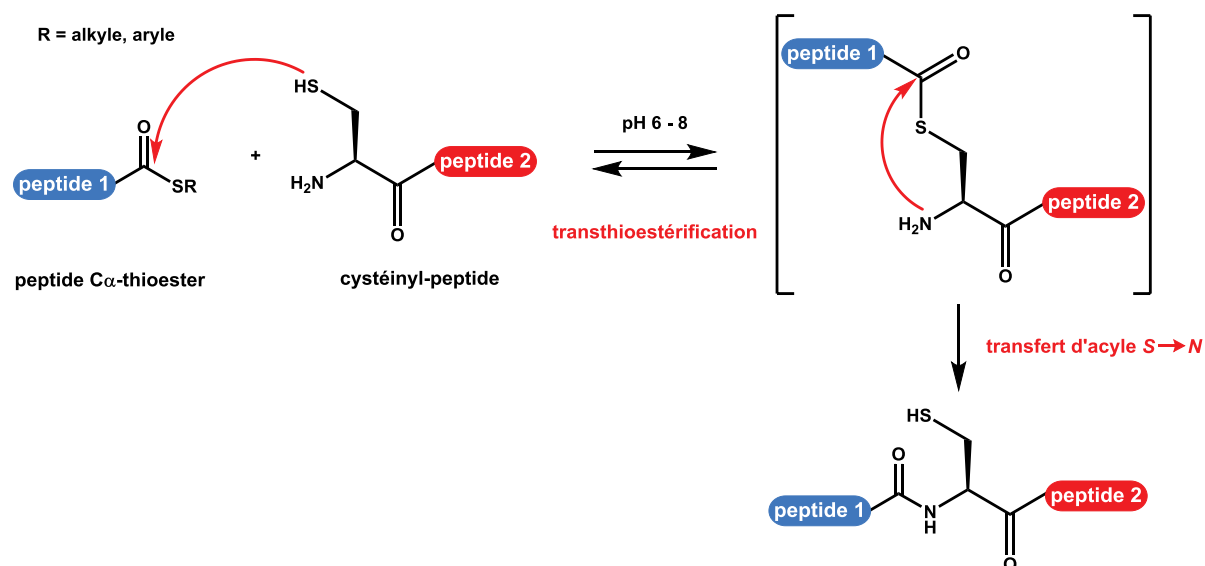


**Figure 10 :** Ligation chimique par formation d'une liaison thioester <sup>25</sup>

Cette approche constitue le premier exemple de ligation complètement chimiosélective pour la synthèse d'analogues de petites protéines et a ouvert la voie au développement d'autres stratégies de ligation. Ce n'est toutefois pas une méthode native, mais peptidomimétique, car créant un mime de liaison amide au point de ligation.

### 1.3.3 La ligation chimique native

En 1994, Kent et ses collaborateurs ont développé une méthode pour lier par une liaison amide deux fragments peptidiques complètement déprotégés. Cette stratégie, appelée « ligation chimique native » (abrégée NCL pour *Native Chemical Ligation*),<sup>28</sup> consiste à faire réagir, dans l'eau à un pH proche de la neutralité, un peptide C $\alpha$ -thioester et un peptide possédant une cystéine N-terminale, ou cystéinyl-peptide (figure 11). Le premier exemple d'application de la NCL par Kent et ses collaborateurs a été la synthèse de l'interleukine 8 humaine (IL-8, 72 acides aminés).<sup>28</sup>

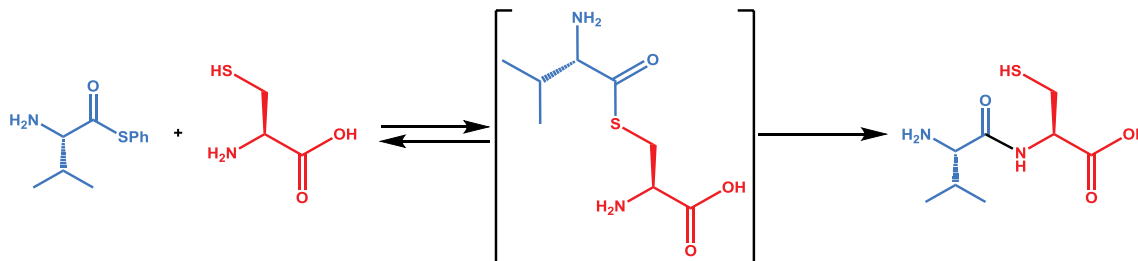


**Figure 11 :** Mécanisme de la ligation chimique native (NCL)<sup>28</sup>

Dans la première étape de transthioestérification, le groupement thiol du résidu cystéine N-terminal attaque le carbone électrophile du groupement carbonyle de la fonction thioester C-terminale. La transthioestérification permet le rapprochement des fonctions amine et thioester dans une géométrie favorable pour conduire à un réarrangement rapide *via* un état de transition à 5 chaînons, par transfert d'acyle intramoléculaire de S vers N (S $\rightarrow$ N), formant la liaison amide. La formation de la liaison amide est irréversible dans les conditions de la réaction. La transthioestérification avec des cystéines internes des fragments peptidiques ne peut pas conduire à la formation d'une liaison amide et est réversible.

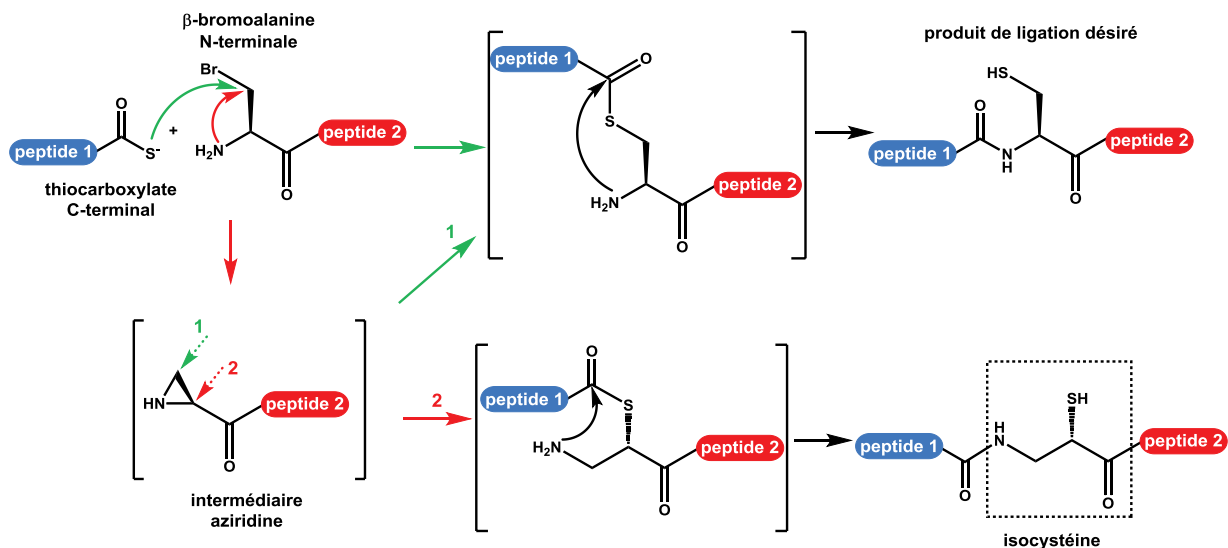
<sup>28</sup> Dawson P. E.; Muir T. W.; Clark-Lewis I.; Kent S. B. H., *Science*, **1994**, 266, 776-779.

La ligation chimique native est inspirée d'une réaction observée par Wieland et ses collaborateurs en 1953. Cette réaction consiste à coupler dans l'eau, de façon chimio- et régiosélective, un thioester de valine avec la cystéine pour conduire à l'obtention du dipeptide Val-Cys (figure 12).<sup>29</sup>



**Figure 12 :** Formation d'une liaison amide entre une valine C $\alpha$ -thioester et la cystéine <sup>29</sup>

Il est à noter que peu de temps après l'introduction de la NCL, Tam et ses collaborateurs ont décrit une approche de ligation native apparentée mettant en œuvre un peptide C $\alpha$ -thioacide et un peptide comportant une  $\beta$ -bromoalanine du côté N-terminal. Le mécanisme mis en jeu fait intervenir le même intermédiaire thioester que lors d'une NCL (figure 13).<sup>30</sup> L'inconvénient majeur de cette stratégie est la possibilité de formation alternative d'un intermédiaire aziridine à partir de la  $\beta$ -bromoalanine, dont l'ouverture par attaque du thioacide peut former un produit comportant une isocystéine comme jonction (voie 2, figure 13). Cette méthode a été, de ce fait, très peu utilisée.



**Figure 13 :** Stratégie de ligation par capture d'acyle et formation d'un intermédiaire aziridine <sup>30</sup>

<sup>29</sup> Wieland T.; Bokelmann E.; Bauer L.; Lang H. U.; Lau H., *Liebigs Ann. Chem.*, **1953**, 583, 129-149.

<sup>30</sup> Tam J. P.; Lu Y. A.; Liu C. F.; Shao J., *Proc. Natl. Acad. Sci. Usa*, **1995**, 92, 12485-12489.

La NCL s'est, au contraire, imposée comme une méthode puissante et très générale pour la synthèse totale de peptides et protéines par voie chimique. Cette stratégie a révolutionné le domaine et a été utilisée de très nombreuses fois pour la synthèse de cibles diverses. L'article original publié dans le journal *Science* en 1994 a fait l'objet de plus de 2300 citations (source : *Google Scholar*, 15/10/2015), ce qui démontre l'impact très important de cette méthodologie au sein de la communauté des sciences des peptides et protéines.

La NCL peut être réalisée à la plupart des sites de jonction acide aminé-cystéine (Xaa-Cys) sans difficultés et avec peu ou pas d'épimérisation. Néanmoins, à cause de réactions secondaires, la NCL pour les sites Asn/Asp-Cys et Gln/Glu-Cys sont complexes et peuvent nécessiter l'utilisation de groupements protecteurs spéciaux pour la chaîne latérale de ces acides aminés.<sup>31</sup> Les aminés encombrés (Ile, Thr ou Val), dont la fonction thioester est peu accessible, ainsi que la proline mènent à des cinétiques très lentes.<sup>32</sup> Dans le cas de la proline, la désactivation de la fonction thioester par interaction  $n \rightarrow \pi^*$  avec le carbonyle de l'acide aminé précédant la proline peut expliquer le manque de réactivité observé.<sup>33</sup>

### 1.3.4 Ligations chimiques natives successives

Comme noté précédemment, la synthèse par SPPS de fragments utilisables pour des réactions de ligation est limitée en routine à des peptides d'une cinquantaine d'acides aminés. Le produit de la ligation de deux fragments ne peut donc logiquement qu'atteindre la taille relativement restreinte d'une centaine d'acides aminés sans optimisation poussée de la synthèse de plus longs fragments. Des stratégies d'assemblage successif par NCLs multiples ont été développées pour atteindre des protéines de plus grande taille (> 100-150 acides aminés) et font l'objet d'un champ de recherche extrêmement dynamique. Il est donc important de les introduire ici mais elles ne seront pas développées car elles sont hors du champ de recherche de cette thèse.

Ces stratégies reposent sur l'utilisation de groupements protecteurs ou de fonctions non réactives ou bien masquées temporairement pour éviter l'oligomérisation ou la cyclisation intramoléculaire des fragments portant les deux fonctionnalités nécessaires à la NCL.

---

<sup>31</sup> Villain M.; Gaertner H.; Botti P., *Eur. J. Org. Chem.*, **2003**, 3267-3272.

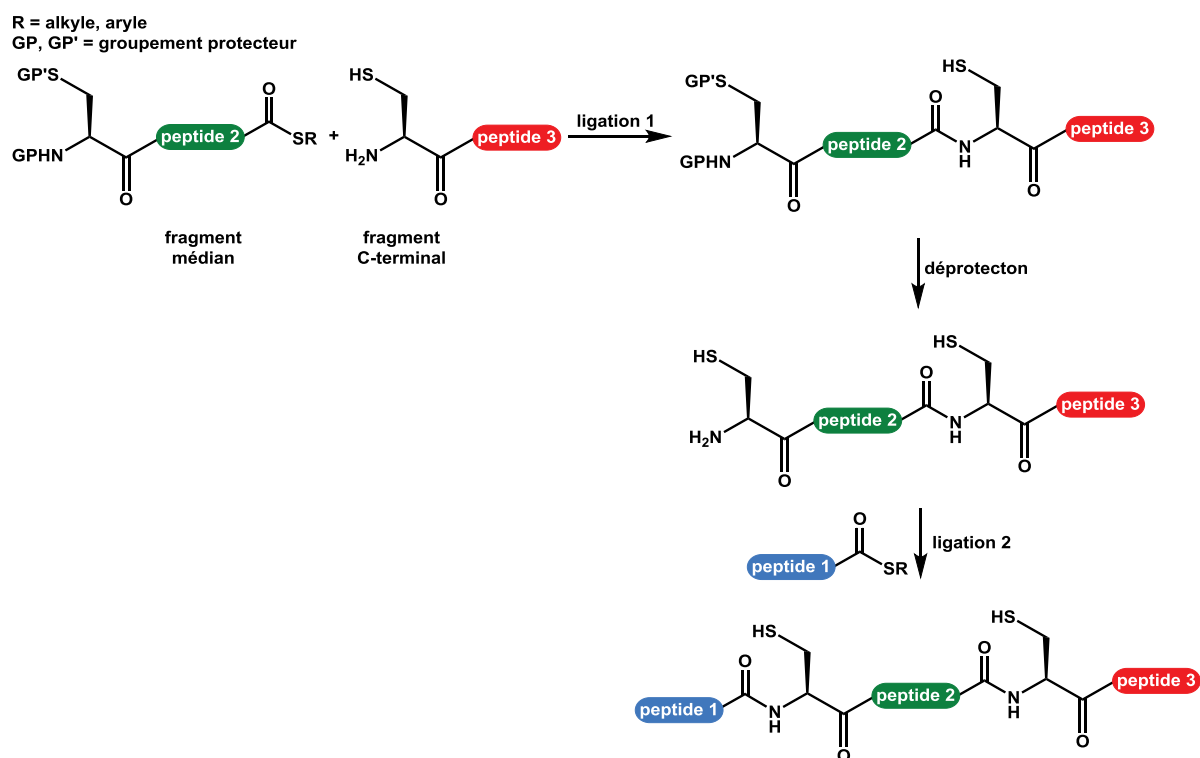
<sup>32</sup> Hackeng T. M.; Griffin J. H.; Dawson P. E., *Proc. Natl. Acad. Sci. Usa*, **1999**, *96*, 10068-10073.

<sup>33</sup> Pollock S. B.; Kent S. B. H., *Chem. Commun.*, **2011**, *47*, 2342-2344.

Deux stratégies différentes sont utilisées : l'assemblage de segments peptidiques du côté C-terminal vers le côté N-terminal (C vers N) ou l'assemblage du côté N-terminal vers le côté C-terminal (N vers C).<sup>34</sup>

### a) Assemblage linéaire de peptides par NCLs multiples de C vers N

Cette stratégie est la plus simple à mettre en place car consistant à utiliser des groupements protecteurs temporaires pour la cystéine N-terminale (figure 14).<sup>34</sup> Les groupements protecteurs sont installés sur le groupement thiol, la fonction amine ou sur les deux fonctions de la cystéine N-terminale. Une première ligation est réalisée entre le fragment C-terminal et le fragment médian (peptide 2 et 3, figure 14). La cystéine N-terminale est ensuite libérée par une étape de déprotection pouvant nécessiter une étape de purification ou être réalisée « *one pot* ». Le fragment N-terminal est ensuite assemblé par une deuxième ligation (peptide 1, figure 14). Ces étapes de déprotection/ligation peuvent être ensuite répétées pour assembler de façon linéaire de très longues séquences peptidiques.

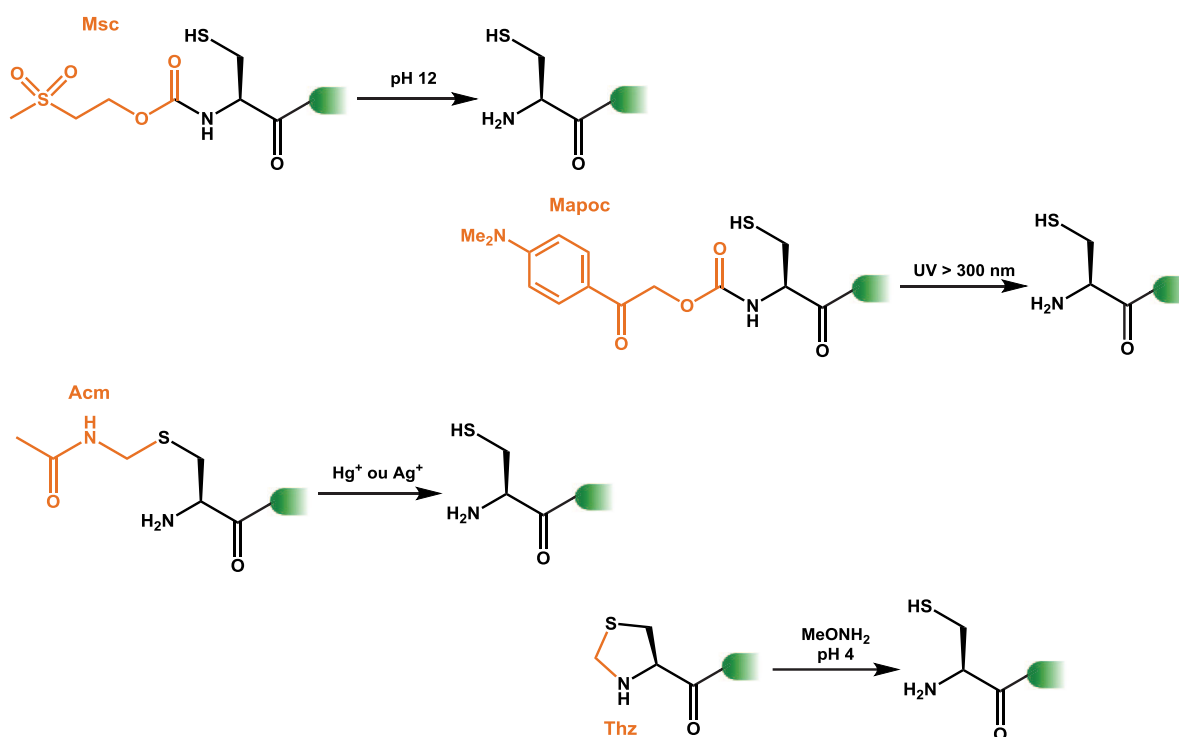


**Figure 14 :** Principe de l'assemblage de peptides par ligations natives multiples de C vers N <sup>34</sup>

Le point clé de cette approche réside dans le choix du groupement protecteur N-terminal. Il doit être stable dans les conditions de couplage, de coupure du peptide de la résine et de déprotection des chaînes latérales, et son élimination après NCL doit pouvoir être réalisée

<sup>34</sup> Raibaut L.; Ollivier N.; Melnyk O., *Chem. Soc. Rev.*, **2012**, *41*, 7001-7015.

dans des conditions bénignes pour les peptides et de préférence en milieu aqueux. La figure 15 regroupe de façon non-exhaustive quelques-uns des groupements protecteurs utilisés ainsi que les conditions pour les éliminer. Par exemple, les groupements 2-méthylsulfonyléthoxycarbone (Msc)<sup>35</sup> et 4-(diméthylamino)phén-acyloxy-carbone (Mapoc)<sup>36</sup> sont installés sur l'amine N-terminale alors que le groupement acétamidométhyle (Acm)<sup>37</sup> est installé sur le thiol de la cystéine N-terminale. Rose et ses collaborateurs ont décrit la protection de la cystéine terminale sous forme de thiazolidine (Thz).<sup>38,39</sup> Cette dernière approche est très utilisée.



**Figure 15 :** Groupements protecteurs pour l'assemblage de C vers N <sup>35-39</sup>

Le groupement photo-labile Mapoc et la thiazolidine ne nécessitent pas forcément une étape de purification après leur déprotection, évitant la perte possible de produit due à la manipulation des peptides, ce qui permet la conduite « *one pot* » de la deuxième ligation.<sup>36,38,39</sup>

### ***b) Assemblage linéaire de peptides par NCLs multiples de N vers C***

Cette stratégie d'assemblage du côté N-terminal vers le côté C-terminal fonctionne de la manière suivante : le fragment médian comporte soit un précurseur désactivé de fonction

<sup>35</sup> Muir T. W.; Dawson P. E.; Kent S. B. H., *Methods Enzymol.*, **1997**, 289, 266-298.

<sup>36</sup> Ueda S.; Fujita M.; Tamamura H.; Fujii N.; Otake A., *ChemBioChem*, **2005**, 6, 1983-1986.

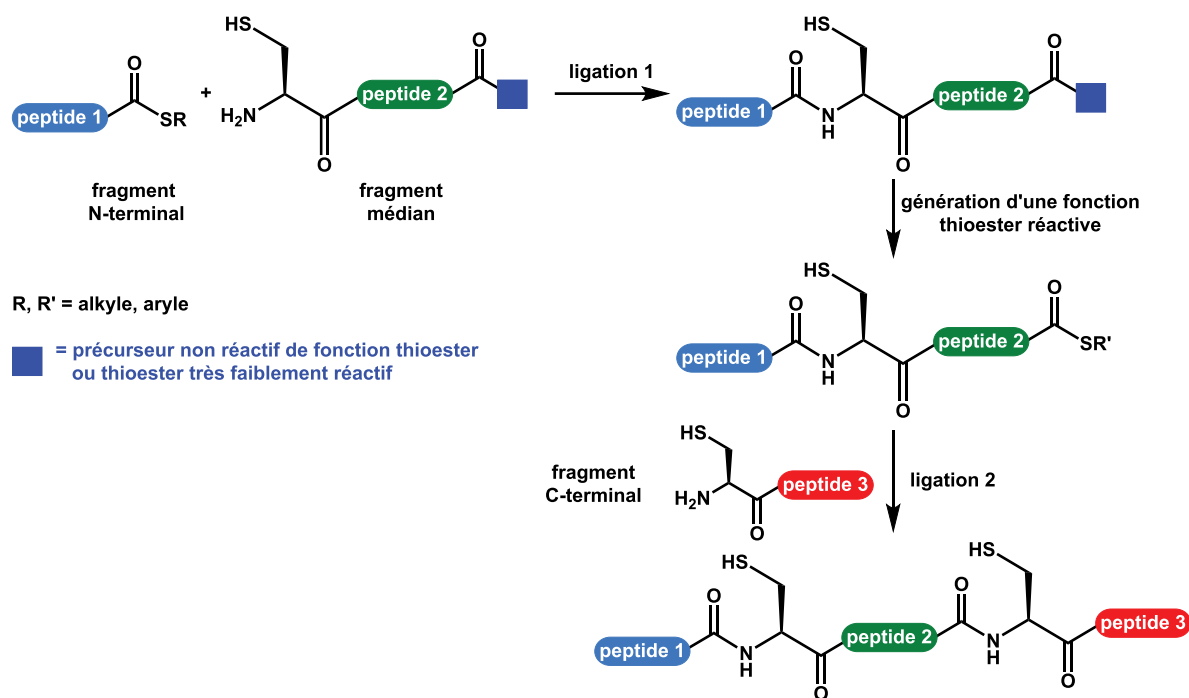
<sup>37</sup> Bang D.; Chopra N.; Kent S. B. H., *J. Am. Chem. Soc.*, **2004**, 126, 1377-1383.

<sup>38</sup> Villain M.; Vizzanova J.; Rose K., *Chem. Biol.*, **2001**, 8, 673-679.

<sup>39</sup> Bang D.; Kent S. B. H., *Angew. Chem. Int. Ed.*, **2004**, 43, 2534-2538.

thioester, soit un thioester très faiblement réactif, permettant une première ligation sélective entre le fragment N-terminal et ce fragment médian (peptide 1 et 2, figure 16). La fonction thioester réactive est ensuite formée pour permettre la réalisation d'une deuxième ligation.<sup>34</sup>

Le concept des NCLs successives de N vers C est décrit ici et certains des précurseurs désactivés de fonction thioester utilisés seront présentés ultérieurement en section 1.5, lors de la description des méthodes de synthèse de peptides thioester.



**Figure 16 :** Principe de l'assemblage de peptides par ligations natives multiples de N vers C <sup>34</sup>

Il est à noter que la vitesse de ligation peut être accélérée par ajout de thiol exogène en excès pour catalyser l'étape de transthioestérification.<sup>40</sup> L'acide 4-mercaptophénylacétique (MPAA) a, par exemple, été décrit comme thiol efficace pour catalyser la NCL;<sup>41</sup> le thioester formé par le MPAA est très activé car ce thiol aryle forme un très bon groupement partant. De plus, le MPAA est très soluble dans l'eau à pH neutre et il n'est pas malodorant.

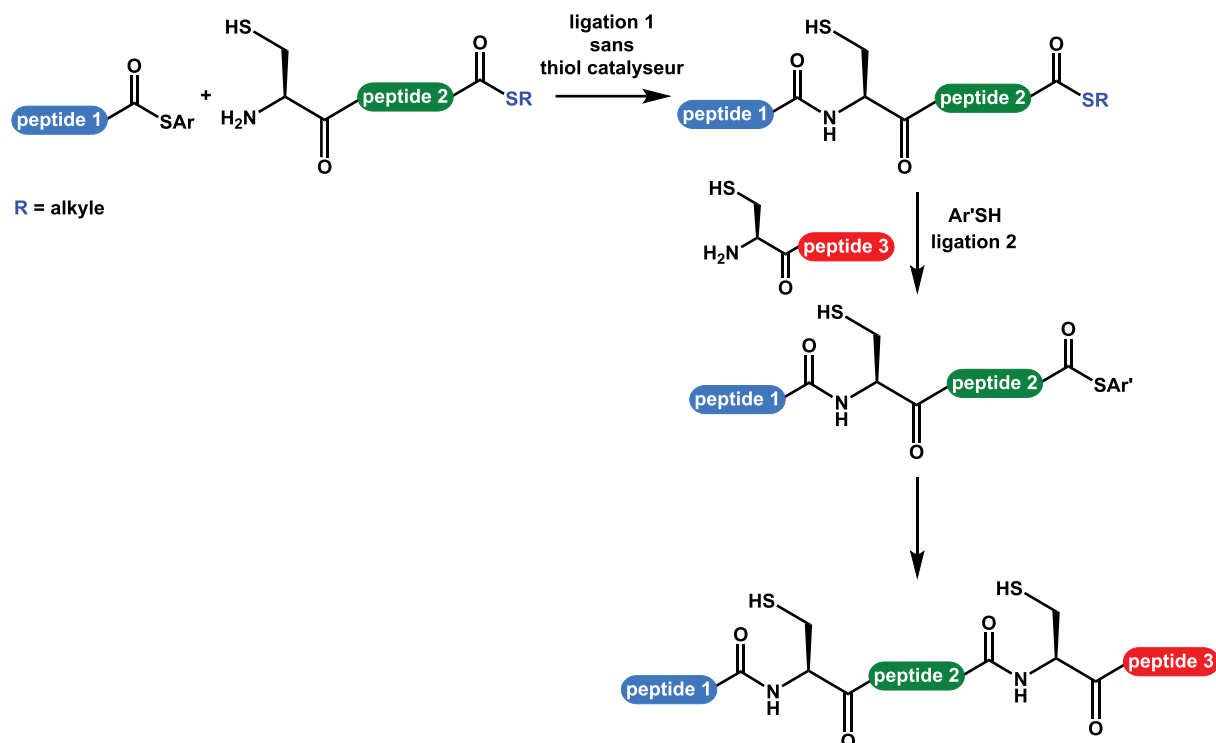
Plusieurs méthodes d'assemblage de N vers C reposant sur l'utilisation de peptides thioester montrant des différences de réactivité ont été décrites, ce principe a été nommé ligation sous contrôle cinétique (*kinetically controlled ligation* - KCL). Par exemple, Kent et ses collaborateurs ont décrit l'exploitation de la différence de réactivité entre un thioester d'aryle (peptide 1, figure 17) et un thioester d'alkyle (peptide 2, figure 17) pour des ligations

<sup>40</sup> Dawson P. E.; Churchill M. J.; Ghadiri M. R.; Kent S. B. H., *J. Am. Chem. Soc.*, **1997**, *119*, 4325-4329.

<sup>41</sup> Johnson E. C. B.; Kent S. B. H., *J. Am. Chem. Soc.*, **2006**, *128*, 6640-6646.



successives.<sup>42</sup> En l'absence de thiol aromatique du type MPAA comme catalyseur, seul le thioester d'aryle (peptide 1) est réactif. Après cette première ligation, un thiol exogène est ajouté au milieu pour catalyser la seconde ligation.



**Figure 17 :** Ligation chimique native sous contrôle cinétique (KCL) <sup>42</sup>

Les approches de C vers N et N vers C peuvent être combinées pour la synthèse convergente de protéines de grande taille. En 2014, l'équipe de Kay a décrit, par exemple, la synthèse de l'enzyme 4-hydroxy-tetrahydrodipicolinate synthase (DapA, 312 résidus) par NCL convergentes de huit fragments, ce qui en fait la protéine la plus longue synthétisée par voie chimique à l'heure actuelle.<sup>43</sup>

A noter que des méthodes de ligations natives successives sur support solide ont également été développées (de C vers N et de N vers C). L'utilisation du support solide permet d'éviter les étapes de purifications intermédiaires conduisant à l'obtention de chaînes peptidiques de grande taille. Un point clé de la stratégie est le développement de bras permettant le greffage sélectif du premier fragment sur le support solide qui puisse résister aux conditions de ligation ainsi qu'aux conditions de coupure de ce bras efficaces pour la libération

<sup>42</sup> Bang D.; Pentelute B. L.; Kent S. B. H., *Angew. Chem. Int. Ed.*, **2006**, *45*, 3985-3988.

<sup>43</sup> Weinstock M. T.; Jacobsen M. T.; Kay M. S., *Proc. Natl. Acad. Sci. Usa*, **2014**, *111*, 11679-11684.

du produit ciblé.<sup>44-49</sup> Ces approches, bien que très prometteuses, ne seront pas décrites en détail car elles sortent du champ de cette thèse.

### 1.3.5 Développements pour la généralisation de la NCL et méthodes apparentées

Une des limitations majeures de la NCL est la présence obligatoire d'une cystéine à la jonction entre les deux fragments. Or, la cystéine fait partie des acides aminés les moins abondants dans les protéines (fréquence d'apparition : 1,5 - 2 %),<sup>50</sup> ce qui restreint le nombre de cibles potentielles directement synthétisables par NCL « classique ».

De nombreux développements ont été réalisés pour étendre la portée de la NCL à la synthèse de la plus grande diversité de cibles possible en s'affranchissant de la nécessité d'une cystéine au point de jonction. J'ai choisi d'exposer ici brièvement, sans être exhaustif, les méthodes les plus représentatives et les plus innovantes conceptuellement.

#### a) Désulfuration du produit de ligation

En 2001, Yan et Dawson ont décrit la désulfuration catalytique de la cystéine d'une protéine synthétisée par NCL, par hydrogénolyse en présence d'un catalyseur hétérogène à base de Ni<sup>0</sup> ou de Pd<sup>0</sup>, conduisant à une alanine au site de ligation (figure 18).<sup>51</sup> Des conditions pour la réalisation de la ligation et de la désulfuration « *one-pot* » ont également été décrites, limitant la diminution de rendement due à une purification intermédiaire.<sup>52,53</sup> Les cystéines présentes dans les séquences peptidiques doivent néanmoins être protégées, ce qui rajoute une étape possiblement délicate de déprotection finale.

---

<sup>44</sup> Canne L. E.; Botti P.; Simon R. J.; Chen Y.; Dennis E. A.; Kent S. B. H., *J. Am. Chem. Soc.*, **1999**, *121*, 8720-8727.

<sup>45</sup> Brik A.; Keinan E.; Dawson P. E., *J. Org. Chem.*, **2000**, *65*, 3829-3835.

<sup>46</sup> Cotton G. J.; Muir T. W., *Chem. Biol.*, **2000**, *7*, 253-261.

<sup>47</sup> Bang D.; Kent S. B. H., *Proc. Natl. Acad. Sci. Usa*, **2005**, *102*, 5014-5019.

<sup>48</sup> Raibaut L.; Adihou H.; Desmet R.; Delmas A. F.; Aucagne V.; Melnyk O., *Chem. Sci.*, **2013**, *4*, 4061-4066.

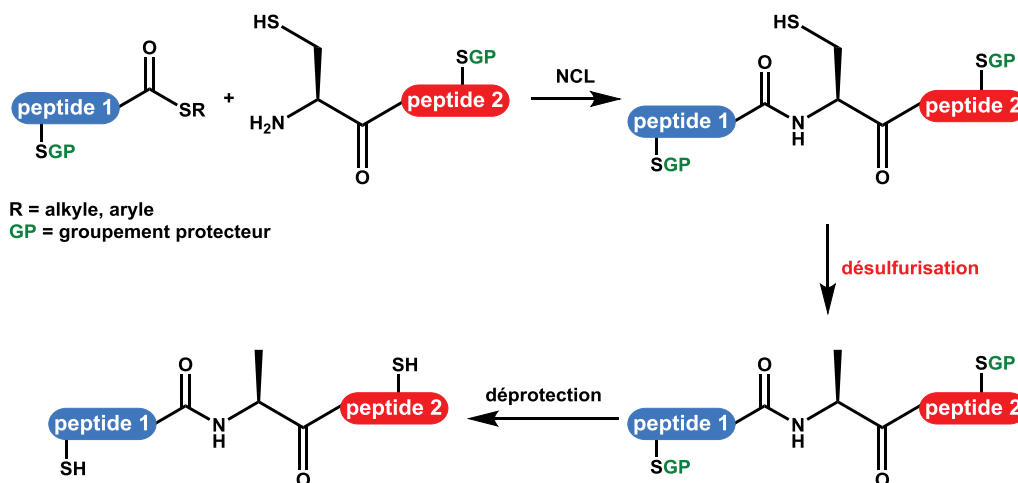
<sup>49</sup> Jbara M.; Seenaiah M.; Brik A., *Chem. Commun.*, **2014**, *50*, 12534-12537.

<sup>50</sup> Trinquier G.; Sanejouand Y-H., *Protein Eng.*, **1998**, *11*, 153-159.

<sup>51</sup> Yan L. Z.; Dawson P. E., *J. Am. Chem. Soc.*, **2001**, *123*, 526-533.

<sup>52</sup> Ma J.; Zeng J.; Wan Q., *Top. Curr. Chem.*, **2015**, *363*, 57-101.

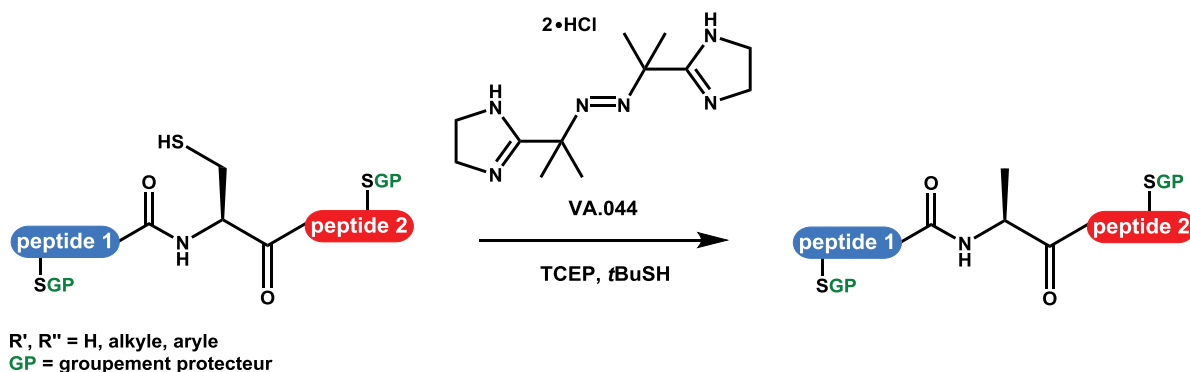
<sup>53</sup> Malins L. R.; Payne R. J., *Top. Curr. Chem.*, **2015**, *362*, 27-87.



**Figure 18 :** Principe de la désulfurisation du produit de la NCL <sup>51</sup>

Même si des conditions relativement douces ont été développées, l'utilisation de métaux pour la désulfurisation est associée à plusieurs limitations : cela peut causer l'épimérisation des alcools secondaires des thréonines, et cette stratégie peut être incompatible avec certains acides aminés (tryptophane, méthionine) ou avec certaines fonctions chimiques comme les thioesters. De plus, la possibilité d'interactions non-spécifiques entre la protéine synthétisée et le métal utilisé peut être source de pertes importantes de produit.

Wan et Danishefsky ont décrit une méthode de désulfurisation par voie radicalaire, sans métaux et plus douce, en utilisant un analogue hydrosoluble de l'AIBN, le VA.044 (dihydrochlorate de 2,2'-azobis[2-(2-imidazolin-2-yl)propane]) en présence de *tris*-(2-carboxyéthyl)phosphine (TCEP) et de *tert*-butanethiol (*t*BuSH) (figure 19).<sup>54</sup> Le TCEP est un agent réducteur qui est, par ailleurs, largement utilisé comme additif pour la NCL pour prévenir la formation de ponts disulfure intra- et intermoléculaires *via* une oxydation par l'oxygène.<sup>55</sup>

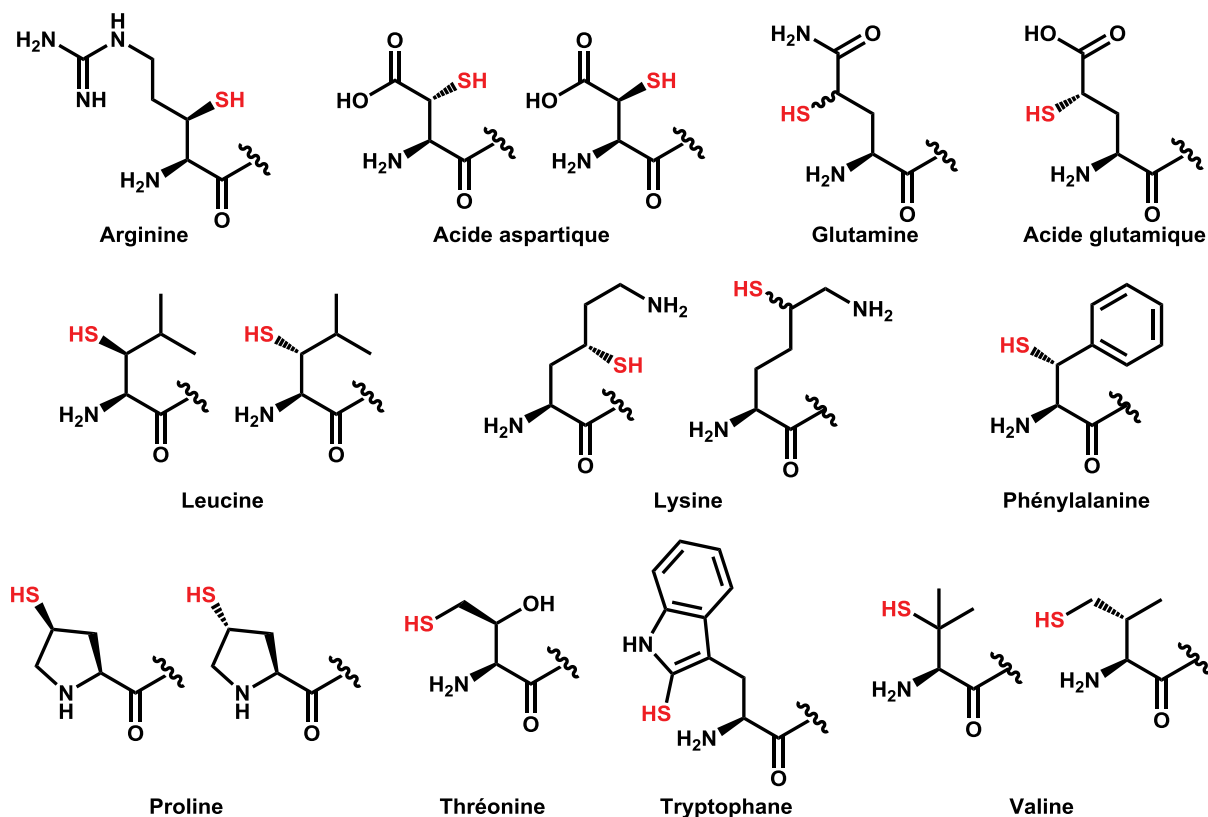


**Figure 19 :** Désulfurisation radicalaire à l'aide de VA.044 en présence de TCEP et de *t*BuSH <sup>54</sup>

<sup>54</sup> Wan Q.; Danishefsky S. J., *Angew. Chem. Int. Ed.*, **2007**, 46, 9248-9252.

<sup>55</sup> Burns J. A.; Butler J. C.; Moran J.; Whitesides G. M., *J. Org. Chem.*, **1991**, 56, 2648-2650.

Ces stratégies de désulfuration ont été également étendues à de nombreux autres résidus par utilisation de dérivés d'acides aminés  $\beta$ -,  $\gamma$ - ou  $\delta$ -thiolés, élargissant grandement les possibilités de cette méthode (figure 20).<sup>52,53</sup> Il faut noter que dans le cas de dérivés  $\delta$ -thiolés, la formation de l'état de transition à 7 chaînons (moins favorable que la formation d'un état de transition à 5 ou 6 chaînons) lors de la ligation, peut conduire à une cinétique de réaction lente.



**Figure 20 :** Dérivés  $\beta$ -,  $\gamma$ - ou  $\delta$ -thiolés d'acides aminés utilisés<sup>52,53</sup>

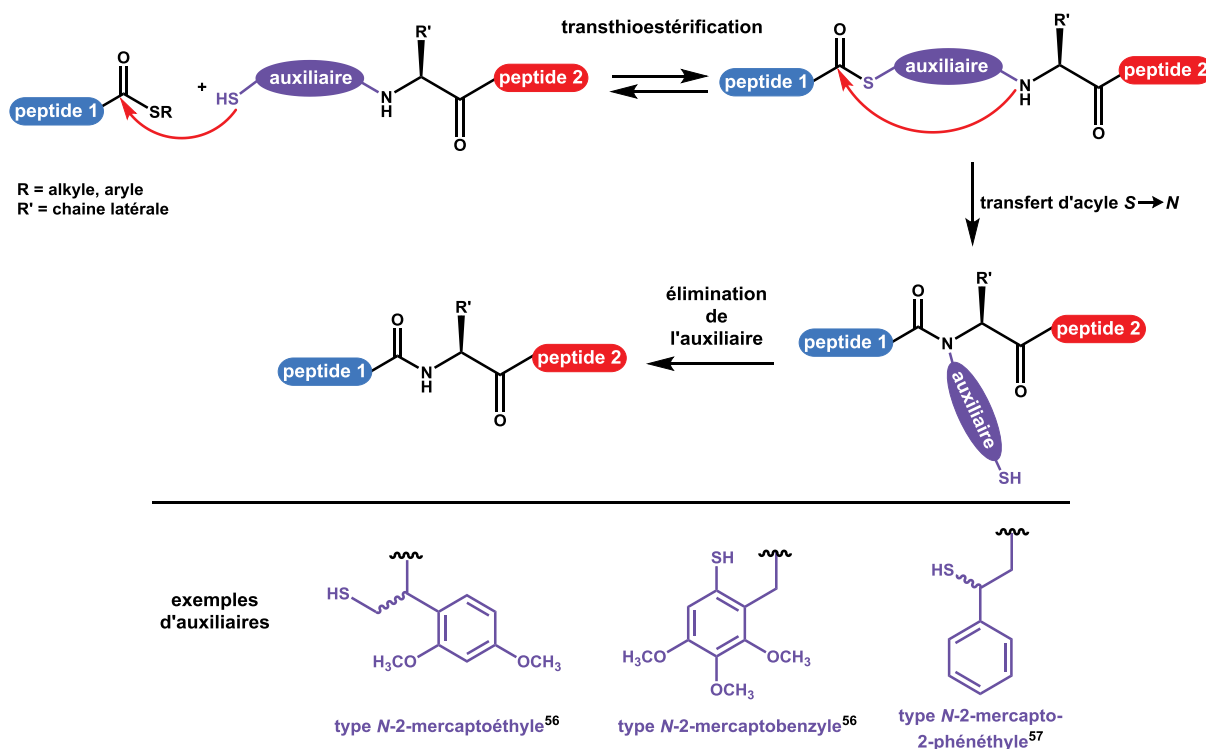
Cette approche est actuellement très en vogue pour la synthèse de protéines ne contenant pas de cystéine dans leur séquence, même si la plupart des analogues thiolés des acides aminés nécessaires pour cette stratégie ne sont pas encore disponibles commercialement et que leur synthèse nécessite de nombreuses étapes.

### b) Ligations natives avec transfert d'acyle assistées par des auxiliaires thiolés

Plusieurs méthodes reposant sur l'utilisation d'auxiliaires munis d'un groupement thiol ont été développées. Ces méthodes combinent, comme la NCL, une étape de transthioestérification suivie d'un transfert d'acyle intramoléculaire  $S \rightarrow N$ .

#### ▪ Auxiliaires N-terminaux

L'auxiliaire thiolé peut être introduit en N-terminal comme substituant de l'amine (figure 21).<sup>56,57</sup> L'auxiliaire est éliminé après la ligation.



**Figure 21 :** Principe de la NCL assistée par des auxiliaires N-terminaux <sup>56,57</sup>

Hormis un exemple très récent,<sup>57</sup> ces stratégies sont majoritairement limitées à des jonctions entre deux acides aminés stériquement très peu encombrés. Dans le cas d'acides aminés encombrés, le transfert d'acyle intramoléculaire sur l'amine secondaire est très lent. De plus, les conditions de coupure de l'auxiliaire peuvent compromettre l'intégrité des protéines synthétisées; le développement d'auxiliaires plus généralisables et dont la coupure après ligation serait également plus douce est sans doute encore nécessaire pour développer pleinement le potentiel de cette approche. L'élimination de l'auxiliaire de type N-2-mercapto-2-

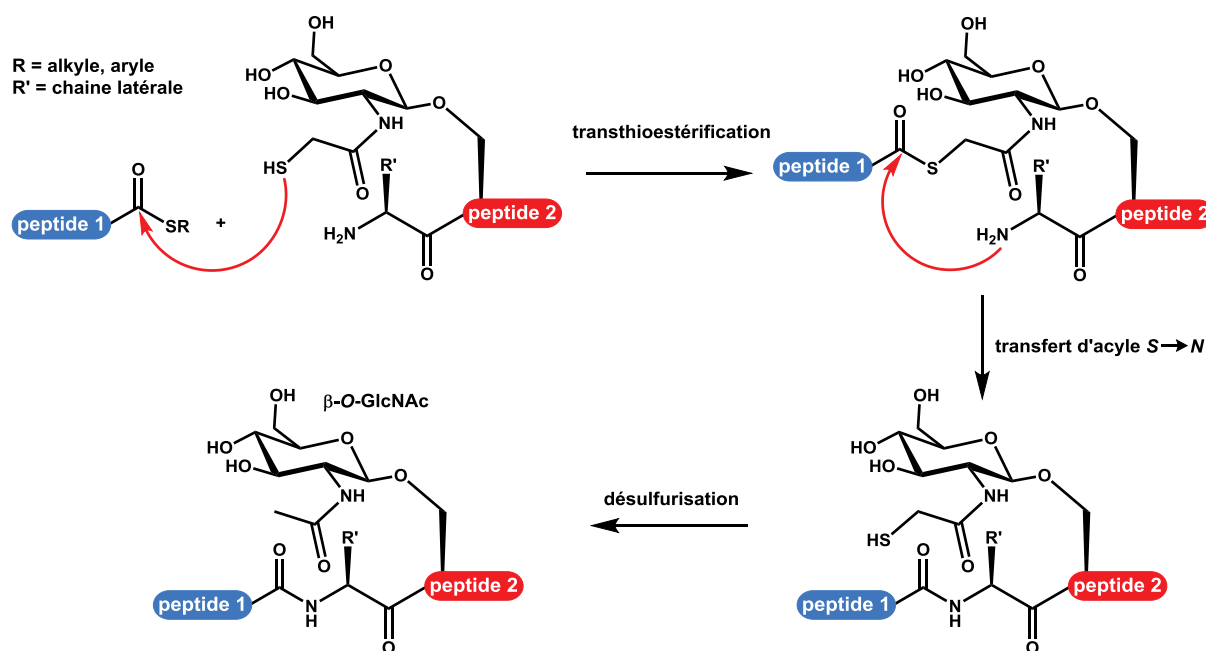
<sup>56</sup> Offer J., *Biopolymers*, **2010**, 94, 530-541.

<sup>57</sup> Loibl S. F.; Harpaz Z.; Seitz O., *Angew. Chem. Int. Ed.*, **2015**, DOI: 10.1002/anie.201505274.

phénéthyle (figure 21) décrit très récemment, est réalisée par voie radicalaire dans des conditions relativement douces, répondant *a priori* à ce problème.<sup>57</sup>

- Auxiliaires portés par la chaîne latérale
- Ligation native assistée par des sucres (SAL)

En 2006, Wong et ses collaborateurs ont développé une méthode élégante de synthèse de glycoprotéines par ligation assistée par un sucre sans nécessité d'une cystéine au site de ligation (*sugar-assisted ligation*, SAL).<sup>58</sup> La fonction thiol est introduite sur la partie saccharidique d'un acide aminé glycosylé (figure 22).



**Figure 22 :** Ligation native assistée par des sucres (SAL) <sup>58</sup>

De façon intéressante, la réaction est plus efficace quand le sucre est placé sur l'avant dernier acide aminé du fragment C-terminal plutôt que sur le dernier, alors que l'intermédiaire formé lors du transfert d'acyle *S*→*N* est constitué d'un cycle plus grand (typiquement 14 ou 15 chaînons). Dans la méthode initiale, le groupement thiol est placé, sous forme d'un β-mercaptoacétamide, sur la position C2 d'une *N*-acétylglucosamine (GlcNAc) et une étape de désulfuration finale est nécessaire pour régénérer le groupement *N*-acétyl natif (figure 22). L'approche est donc incompatible avec l'utilisation de peptides comportant des cystéines ayant leur groupement thiol libre. Pour y remédier, Wong et ses collaborateurs ont décrit l'utilisation d'un sucre modifié portant un groupement 2-mercaptoacétate en position C3 permettant la

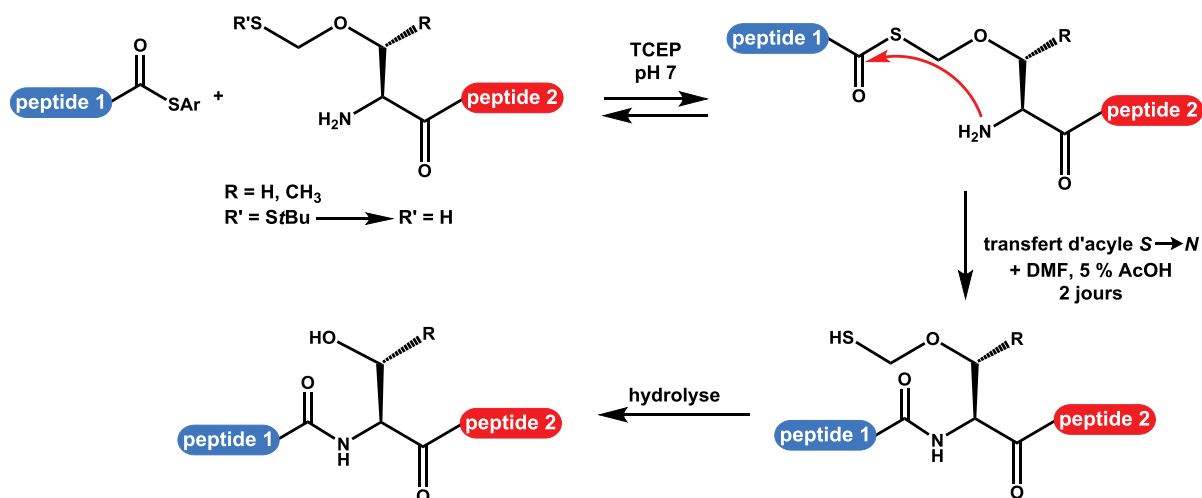
<sup>58</sup> Brik A.; Yang Y. Y.; Ficht S.; Wong C. H., *J. Am. Chem. Soc.*, **2006**, 128, 5626-5627.

transthioestérification, ce groupement pouvant être éliminé en fin de synthèse par une hydrazinolyse qui n'affecte pas les cystéines libres des segments peptidiques.<sup>59</sup>

Cette méthode reste néanmoins très complexe à mettre en œuvre et n'a pas été utilisée en dehors des preuves de concept.

#### - Utilisation d'un groupement mercaptométhyle pour la ligation native

En 2010, Hojo et ses collaborateurs ont développé un nouvel auxiliaire thiolé porté par la chaîne latérale d'une sérine ou thréonine sous la forme d'un thiohémicétal permettant des réactions de ligation native aux sites Xaa-Ser ou Xaa-Thr (figure 23).<sup>60</sup> La fonction thiohémicétale étant très labile, l'auxiliaire est stabilisé par protection sous forme de pont disulfure qui, lors d'une ligation, sera d'abord libéré *in situ* par le TCEP pour permettre l'étape de transthoestérification. Le transfert d'acyle  $S \rightarrow N$  qui s'ensuit fait intervenir un état de transition à 7 chaînons. Le groupement mercaptométhyle est ensuite éliminé par hydrolyse spontanée sans nécessiter d'étape de déprotection supplémentaire (figure 23). Cette stratégie a été appliquée avec succès à la préparation d'une toxine glycopeptidique issue du venin d'un cône marin, la contulakine-G.<sup>60</sup>



**Figure 23 :** Utilisation d'un groupement mercaptométhyle pour une ligation native à la Ser ou Thr<sup>60</sup>

Un problème de la méthode consiste en l'hydrolyse spontanée du groupement thiohémicétal avant la transthoestérification ou le transfert d'acyle intramoléculaire conduisant à différents co-produits. Enfin, l'installation du groupement thiohémicétal sur les

<sup>59</sup> Ficht S.; Payne R. J.; Brik A.; Wong C. H., *Angew. Chem. Int. Ed.*, **2007**, 46, 5975-5979.

<sup>60</sup> Hojo H.; Ozawa C.; Katayama H.; Ueki A.; Nakahara Y.; Nakahara Y., *Angew. Chem. Int. Ed.*, **2010**, 49, 5318-5321.

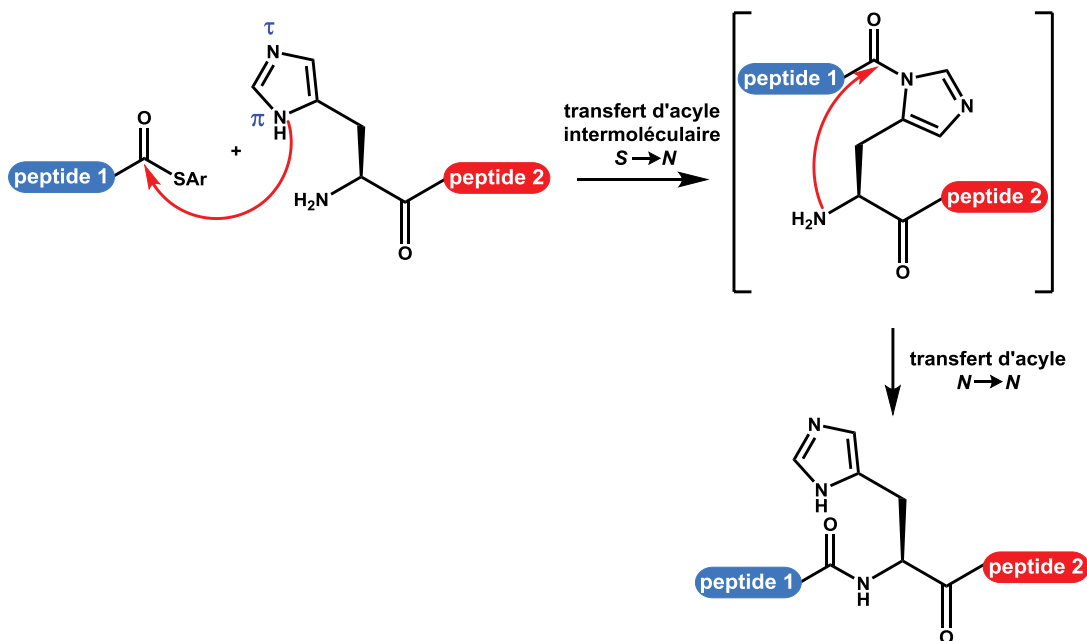
résidus sérine ou thréonine nécessite plusieurs étapes de synthèses organiques pouvant représenter un obstacle pour les non-spécialistes voulant appliquer cette méthode.

### c) Méthodes apparentées à la NCL

Plusieurs autres méthodes apparentées à la NCL ont été développées, reposant également sur une première étape de transfert d'acyle intermoléculaire entre les deux fragments pour rapprocher la fonction amine N-terminale d'un des partenaires et un dérivé activé du carboxyle C-terminal de l'autre partenaire. Néanmoins, hormis dans un cas, elles diffèrent de la NCL dans le fait que le nucléophile réalisant l'étape de transfert intermoléculaire n'est pas un groupement thiol. J'ai choisi de présenter brièvement les plus représentatives pour illustrer ce domaine foisonnant de la science des peptides et protéines.

#### ▪ Ligation native assistée par une histidine

En 1997, Zhang et Tam ont décrit une méthode de ligation à l'histidine utilisant les propriétés nucléophiles du noyau imidazole pour opérer un transfert d'acyle  $S \rightarrow N$  intermoléculaire (figure 24).<sup>61</sup> Les rendements obtenus sont toutefois très faibles. Ceci pourrait s'expliquer par la compétition entre les deux azotes du noyau imidazole ( $\pi$  et  $\tau$ ), l'acylation de l'azote  $\tau$  produisant un intermédiaire non favorable d'un point de vue géométrique. Par ailleurs, la réaction est incompatible avec la présence d'autres groupements nucléophiles sur les segments peptidiques; le développement de cette méthode n'a pas été poursuivi.



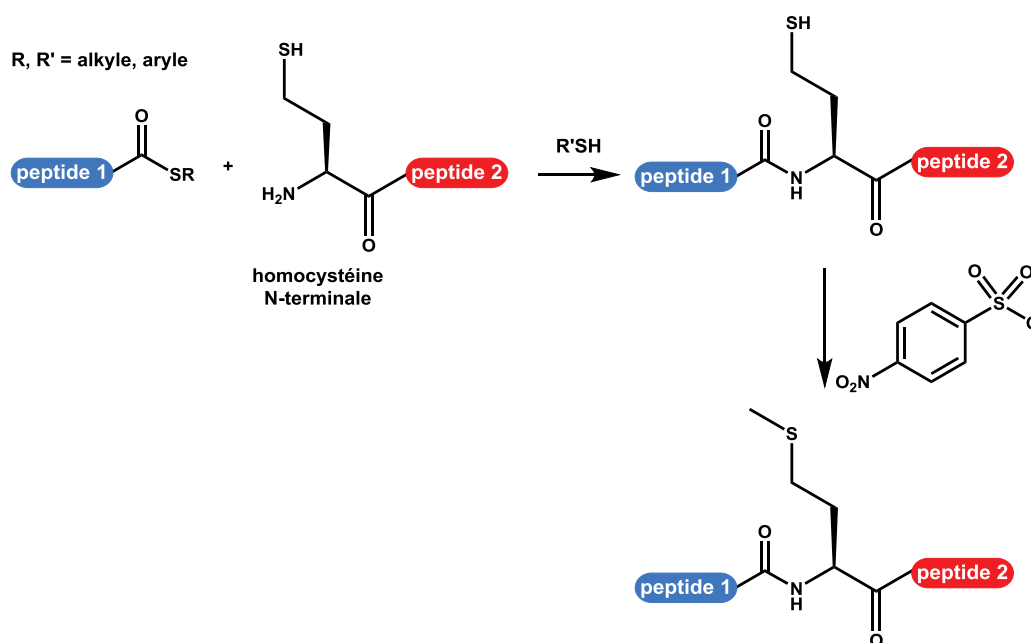
**Figure 24 :** Ligation native assistée par une histidine <sup>61</sup>

<sup>61</sup> Zhang L.; Tam J. P., *Tetrahedron Lett.*, **1997**, 38, 3-6.



- Formation d'une méthionine au site de ligation

En 1998, Tam et Yu ont également décrit une NCL réalisée en remplaçant la cystéine N-terminale du partenaire C-terminal par une homocystéine (figure 25).<sup>62</sup> La transthioestérfication s'opère normalement avec le thiol de la chaîne latérale de l'homocystéine et le réarrangement par transfert d'acyle  $S \rightarrow N$  implique un état de transition à 6 chaînons. Après ligation, l'homocystéine peut être méthylée par réaction avec le *para*-nitrobenzènesulfonate de méthyle pour former une méthionine au site de ligation. Si la méthode peut sembler attrayante, les auteurs ont néanmoins décrit plusieurs réactions secondaires ayant lieu lors de cette réaction de méthylation dont la rupture de la liaison amide entre l'homocystéine et l'acide aminé qui le précède. De plus, les chaînes latérales des résidus cystéine nécessitent d'être protégées.



**Figure 25 :** Ligation native assistée par une homocystéine suivie de la formation d'une méthionine<sup>62</sup>

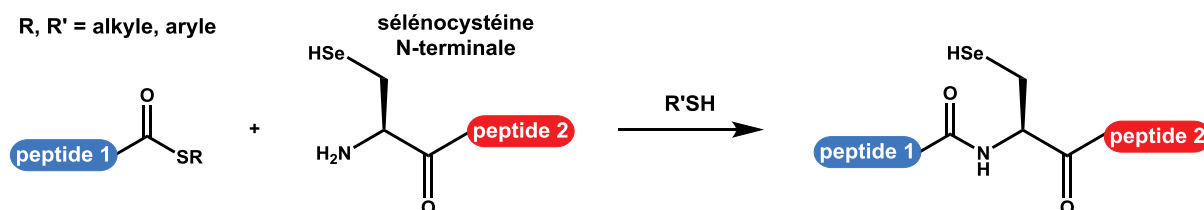
- Ligation native assistée par une sélénocystéine

L'utilisation de sélénocystéinyl-peptide en remplacement du cystéinyl-peptide classique a également été décrite comme variation de la NCL (figure 26).<sup>63</sup> Le pKa du groupement séléinol est d'environ 5 à comparer avec le pKa du groupement thiol de la cystéine, d'environ 8. Le groupement séléinol est donc un meilleur groupement partant et est majoritairement déprotonné à pH neutre, faisant de lui un excellent nucléophile à ce pH. L'utilisation de la sélénocystéine conduit à des ligations nettement plus rapides et cette méthode est

<sup>62</sup> Tam J. P.; Yu Q., *Biopolymers*, **1998**, 46, 319-327.

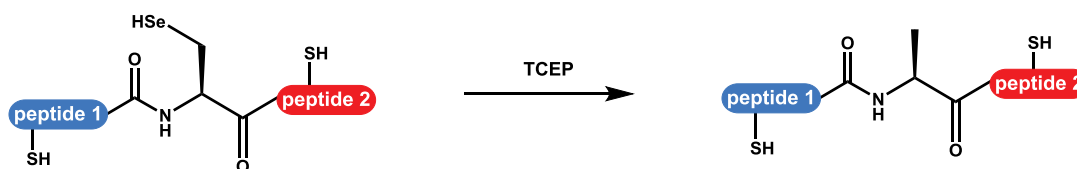
<sup>63</sup> Gieselmann M. D.; Xie L.; Van Der Donk W. A., *Org. Lett.*, **2001**, 3, 1331-1334.

particulièrement adaptée à des partenaires thioester présentant un acide aminé encombré en C-terminal.<sup>64</sup>



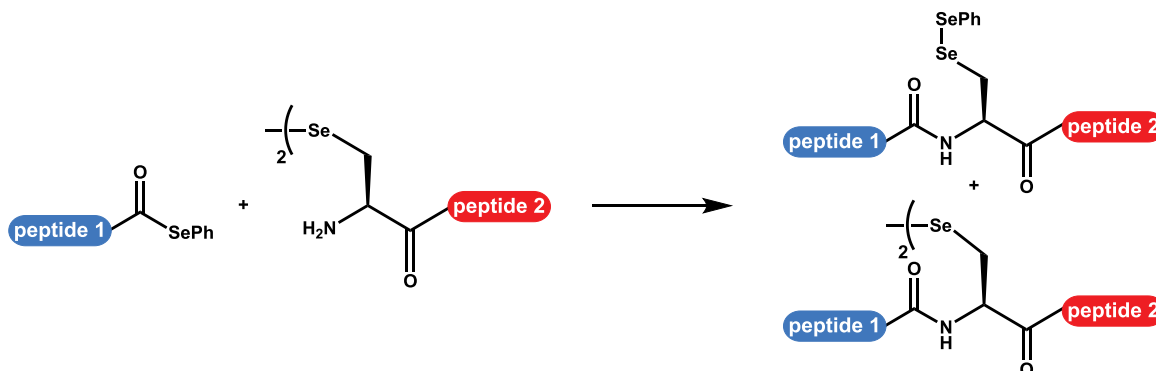
**Figure 26 :** Ligation à la sélénocystéine <sup>63</sup>

L'utilisation du TCEP dans cette ligation est à proscrire car elle peut conduire à une désélenisation du produit. Dawson et ses collaborateurs ont exploité cette particularité en décrivant l'utilisation du TCEP pour la désélenisation sélective post-NCL d'une sélénocystéine conduisant à l'obtention d'une alanine au point de ligation sans nécessiter de protection des cystéines des fragments peptidiques (figure 27).<sup>65</sup>



**Figure 27 :** Désélenisation sélective du produit de ligation <sup>65</sup>

Très récemment, Payne et ses collaborateurs ont décrit une ligation n'utilisant non pas un peptide thioester mais un séléniothioester en complément d'un sélénocystéinyl-peptide (figure 28).<sup>66</sup> Cette méthode ne nécessite pas l'utilisation de thiol exogène, ni de réducteur comme le TCEP et les cinétiques observées sont extrêmement rapides (réaction terminée après 10 min).



**Figure 28 :** Ligation à la sélénocystéine <sup>66</sup>

<sup>64</sup> Malins L. R.; Mitchell N. J.; Payne R. J., *J. Pept. Sci.*, **2014**, *20*, 64-77.

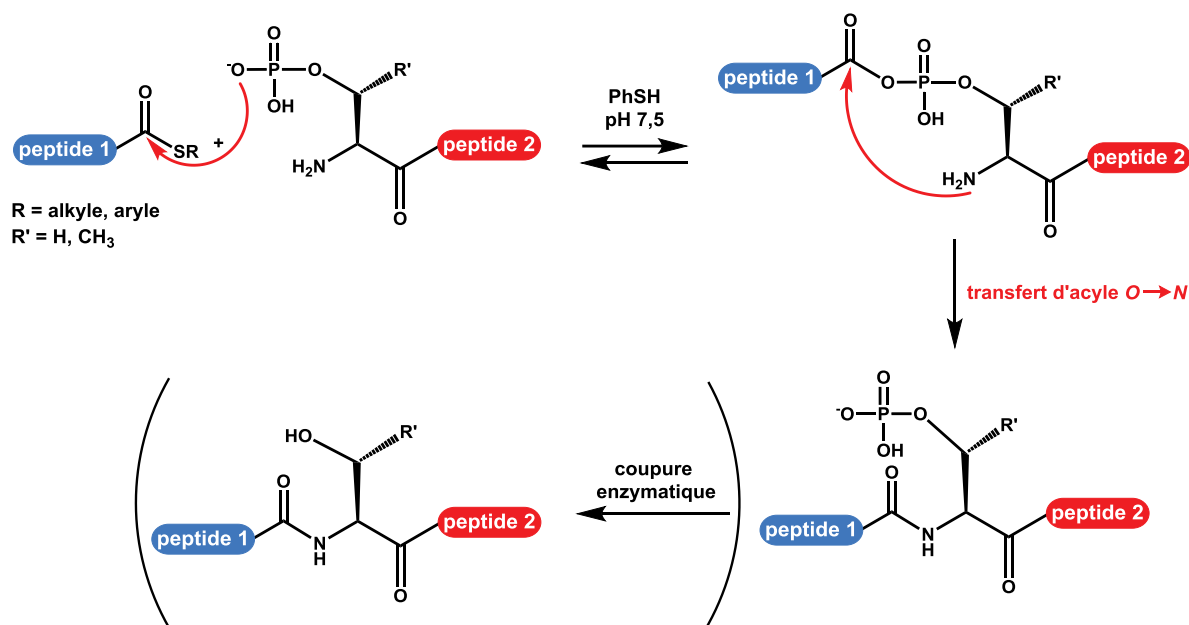
<sup>65</sup> Metanis N.; Keinan E.; Dawson P. E., *Angew. Chem. Int. Ed.*, **2010**, *49*, 7049-7053.

<sup>66</sup> Mitchell N. J.; Malins L. R.; Liu X.; Thompson R. E.; Chan B.; Radom L.; Payne R. J., *J. Am. Chem. Soc.*, **2015**, DOI : 10.1021/jacs.5b07237.

En dépit de son potentiel, cette approche est peu utilisée, notamment à cause de la forte réactivité des groupements séléniols pouvant être source de plusieurs réactions secondaires dont la formation de ponts diséléniure par oxydation à l'air, compliquant la manipulation de ces peptides.

- Ligation native assistée par un groupement phosphate

Thomas et Payne ont mis au point une méthode de ligation native reposant sur la réactivité d'un peptide comportant une phosphosérine ou une phosphothréonine N-terminale (figure 29).<sup>67</sup> Le caractère nucléophile du groupement phosphate est *a priori* suffisant à un pH d'environ 7 pour permettre la transestérification avec la fonction thioester formant un phosphate d'acyle très réactif. Le transfert d'acyle  $O \rightarrow N$  passe vraisemblablement par un état de transition à 7 chaînons, même si le mécanisme n'est pas démontré. Le produit obtenu est un phosphopeptide. L'utilisation de l'enzyme phosphatase alcaline permet d'éliminer le groupement phosphate et d'obtenir une sérine ou thréonine au site de ligation.



**Figure 29 :** Ligation native assistée par un groupement phosphate <sup>67</sup>

Les cinétiques de réaction observées avec cette méthode sont néanmoins très lentes et cette méthode n'a pas été utilisée en dehors de preuve de concept utilisant des peptides modèles.

<sup>67</sup> Thomas G. L.; Payne R. J., *Chem. Commun.*, **2009**, 45, 4260-4262.

Toutes ces méthodes sont des alternatives intéressantes à la NCL pour étendre la boîte à outils disponible pour les chimistes des protéines de synthèse, même si hormis de rares exceptions (désulfuration, ligation à la sélénocystéine, utilisation d'auxiliaires ou d'un groupement mercaptométhyle), ces méthodes se sont arrêtées au stade de preuves de concept.

Il est important de noter l'utilisation commune des peptides thioester pour toutes ces réactions, illustrant ainsi leur rôle central dans la synthèse de longs peptides et de protéines par voie chimique.

### 1.4 Autres stratégies de ligation native

Bien que la NCL soit la stratégie de ligation native faisant référence actuellement, plusieurs autres stratégies ont été décrites. Ces méthodologies diffèrent de la NCL car elles ne sont pas basées sur un transfert d'acyle intermoléculaire pour lier chimiosélectivement de façon covalente les deux fragments. J'expose ces stratégies ici, brièvement, avant de traiter le cœur de mon sujet de thèse en section 1.5.

#### 1.4.1 Ligation de Staudinger native

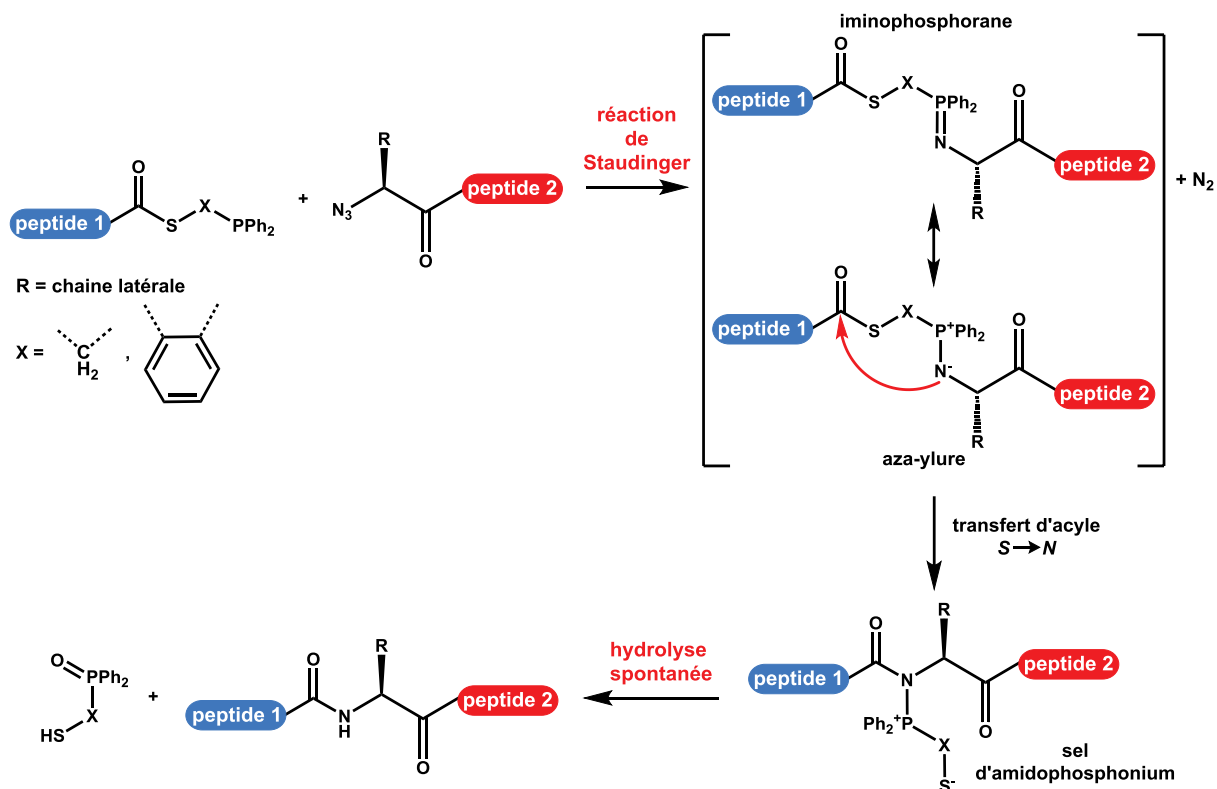
En 2000, les équipes de Bertozzi et Raines ont développé de façon indépendante une méthode de ligation native reposant sur l'utilisation de la réaction de Staudinger.<sup>68-70</sup> La ligation consiste en la réaction entre un phosphino-thioester C-terminal et un azoture N-terminal (figure 30). L'azote de l'intermédiaire aza-ylure formé opère une attaque nucléophile sur le carbone électrophile du thioester conduisant à un transfert d'acyle intramoléculaire  $S \rightarrow N$  pour former la liaison amide passant par un état de transition à 5 ou 6 chaînons. L'hydrolyse spontanée de l'amidophosphonium formé permet d'obtenir l'obtention d'un produit de ligation sans trace du groupement phosphoré.

---

<sup>68</sup> Saxon E.; Bertozzi C., *Science*, **2000**, 287, 2007-2010.

<sup>69</sup> Nilsson B. L.; Kiessling L. L.; Raines R. T., *Org. Lett.*, **2000**, 2, 1939-1941.

<sup>70</sup> Soellner M. B.; Nilsson B. L.; Raines R. T., *J. Am. Chem. Soc.*, **2006**, 128, 8820-8828.



**Figure 30 :** Principe de la ligation de Staudinger native <sup>68-70</sup>

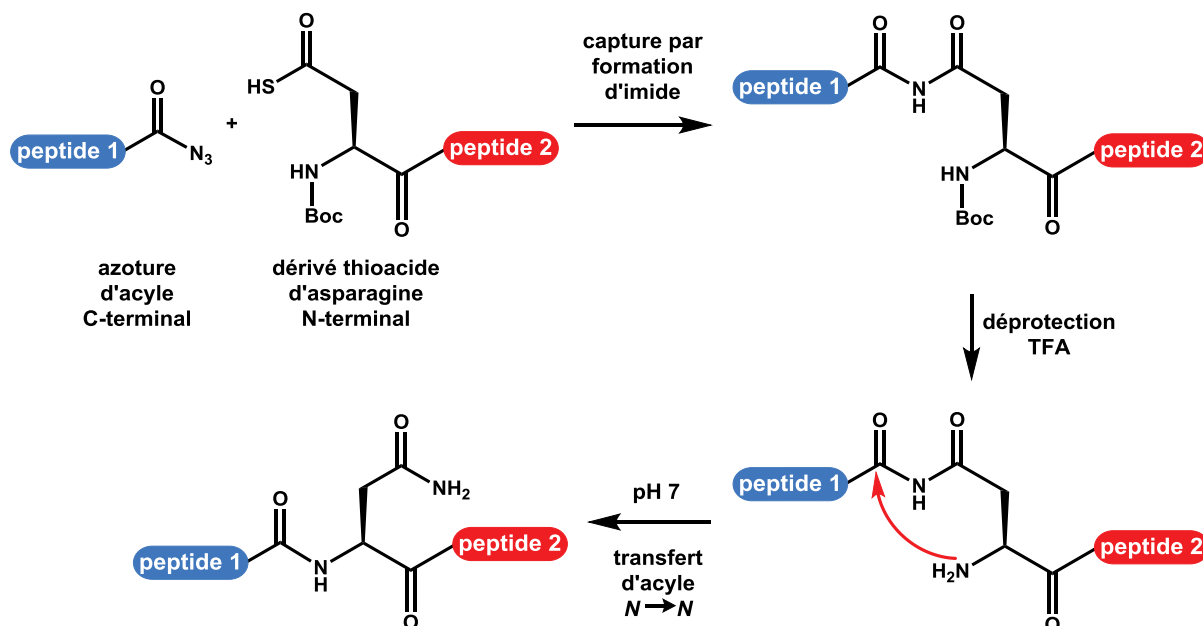
Cette stratégie est théoriquement applicable à n'importe quels acides aminés au site de ligation et la chimiosélectivité élevée de la réaction entre une phosphine et un azoture permet l'utilisation de peptides complètement déprotégés. Toutefois, la phosphine est une fonction très réactive pouvant s'oxyder en présence de trace d'oxygène, et l'hydrolyse du phosphonium de l'aza-ylure intermédiaire est en compétition avec le transfert d'acyle  $S \rightarrow N$ . Dans une étude mécanistique, Raines et ses collaborateurs ont déterminé que la concentration idéale des peptides devait être de plus de 100 mM pour éviter cette réaction secondaire, cette concentration n'est sans doute pas réaliste pour une application en dehors de peptides modèles.<sup>70</sup>

### 1.4.2 Ligation native par capture sous forme d'imide puis transfert d'acyle $N \rightarrow N$

En 2013, Melnyk et ses collaborateurs ont décrit une méthode ligation à l'asparagine faisant intervenir un peptide azoture d'acyle C-terminal et un peptide comportant un dérivé thioacide *N*-protégé d'asparagine *N*-terminal (figure 31).<sup>71</sup> La première étape repose sur une capture d'acyle par réaction entre l'azoture d'acyle et le thioacide, conduisant à la formation d'une fonction imide. Après déprotection de l'amine du dérivé d'asparagine, la liaison amide est

<sup>71</sup> Mhida R.; Boll E.; Fécourt F.; Ermolenko M.; Ollivier N.; Sasaki K.; Crich D.; Delpech B.; Melnyk O., *Bioorg. Med. Chem.*, **2013**, *21*, 3479-3485.

formée par un transfert d'acyle intramoléculaire  $N \rightarrow N$ , favorisé par l'électrophilie du carbonyle de l'imide, passant par un état de transition à 6 chaînons.



**Figure 31 :** Principe de la ligation native par capture sous forme d'imide puis transfert d'acyle  $N \rightarrow N$ <sup>71</sup>

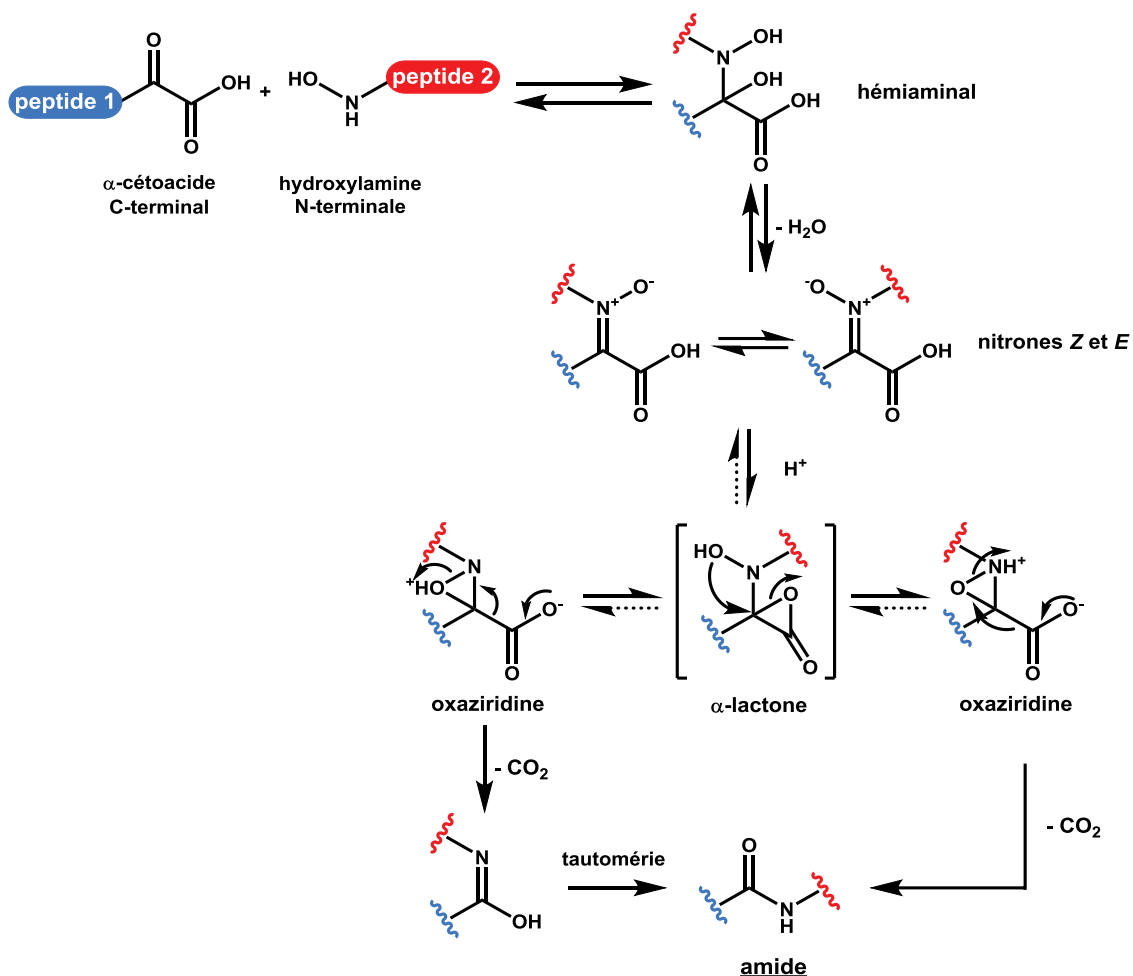
Lors de l'étude de la capture sous forme d'imide avec des peptides modèles, des produits secondaires consistant en les imides symétriques (peptide1-imide-peptide1 et peptide2-imide-peptide2) ont été observés en quantités importantes par rapport à l'imide asymétrique désiré. Leur formation est due à l'échange rapide des fonctions azide et thioamide. Le développement de cette stratégie s'est limité à la preuve de concept.

### 1.4.3 Ligation KAHA

En 2006, Bode et ses collaborateurs ont décrit une nouvelle méthode de ligation native conceptuellement très éloignée des méthodes existantes. Cette stratégie consiste à coupler un peptide comportant un  $\alpha$ -cétoacide C-terminal avec un peptide comportant une hydroxylamine N-terminale (figure 32), cette ligation a été nommée «  $\alpha$ -ketoacid-hydroxylamine ligation » (ligation KAHA).<sup>72</sup> Cette stratégie fonctionne avec des peptides complètement déprotégés et ne requiert pas de réactifs additionnels; de plus, l'eau et le dioxyde de carbone sont les seuls co-produits de la réaction. Le mécanisme complexe de la réaction est résumé par la formation d'un hémiaminal suivie de la perte d'une molécule d'eau conduisant à une nitroène; finalement, la

<sup>72</sup> Bode J. W.; Fox R. M.; Baucom K. D., *Angew. Chem. Int. Ed.*, **2006**, 45, 1248-1252.

liaison amide est formée par une suite de réarrangements incluant une décarboxylation (figure 32).<sup>73</sup>



**Figure 32 :** Mécanisme de la ligation KAHA <sup>72-73</sup>

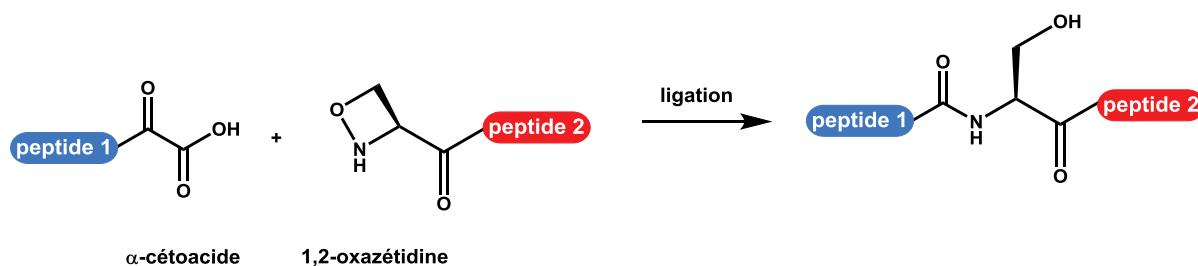
La version initiale de cette réaction était limitée aux milieux organiques (DMF ou DMSO) et nécessitait des concentrations élevées en peptides (20 à 100 mM), ce qui représentait une limitation majeure de la stratégie.<sup>73</sup>

Plusieurs améliorations successives ont été décrites,<sup>73-75</sup> aboutissant finalement très récemment à une méthode *a priori* efficace à de plus faibles concentrations.<sup>76</sup> Cette méthode consiste en la ligation d'un peptide  $\alpha$ -cétoacide C-terminal et d'un peptide contenant une oxazétidine N-terminale pour conduire à un produit de ligation avec une sérine à la jonction des deux fragments (figure 33). La tension de cycle élevée de l'oxazétidine augmente la réactivité de ce dérivé d'hydroxylamine conduisant à des ligations rapides.

<sup>73</sup> Pusterla I.; Bode J. W., *Angew. Chem. Int. Ed.*, **2012**, 51, 513-516.

<sup>74</sup> Pattabiraman V. R.; Ogunkoya A. O.; Bode J. W., *Angew. Chem. Int. Ed.*, **2012**, 51, 5114-5118.

<sup>75</sup> Wucherpfennig T. G.; Rohrbacher F.; Pattabiraman V. R.; Bode J. W., *Angew. Chem. Int. Ed.*, **2014**, 53, 1-5.



**Figure 33** : Ligation KAHA avec une oxazétidine N-terminale <sup>76</sup>

Toutefois, les peptides  $\alpha$ -cétoacide C-terminaux et les oxazétidines N-terminales sont difficiles à obtenir car nécessitant de nombreuses étapes de synthèses.

Malgré cela, cette stratégie semble être une approche très prometteuse, illustrée par son application à la synthèse de plusieurs cibles.<sup>76</sup>

#### 1.4.4 Ligation native par formation d'une oxazolidine

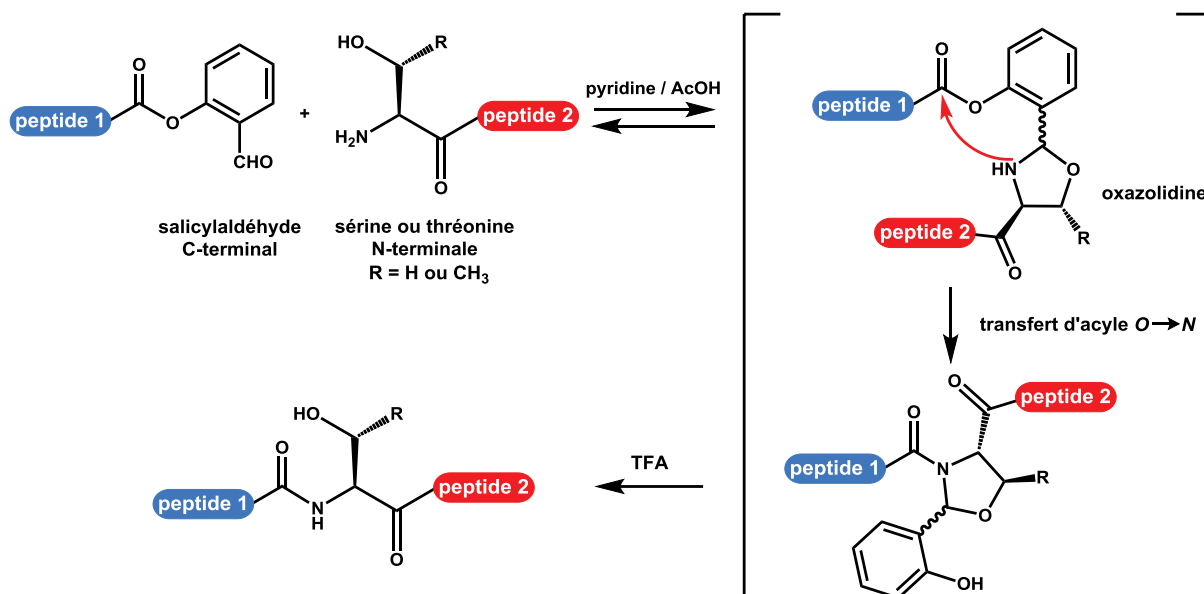
En 2010, Li et ses collaborateurs, s'inspirant des travaux précédents de Tam,<sup>77</sup> ont décrit une ligation faisant intervenir une sérine ou une thréonine N-terminale (*serine/threonine ligation* - STL) réagissant avec un ester de salicylaldéhyde C-terminal (figure 34).<sup>78</sup> Cette réaction est réalisée dans mélange pyridine/acide acétique. La sérine ou la thréonine N-terminale est réversiblement liée avec le groupement aldéhyde du motif C-terminal pour former un intermédiaire oxazolidine. Cet intermédiaire se réarrange irréversiblement par transfert d'acyle  $O \rightarrow N$  pour former la liaison amide *via* un état de transition à 6 chaînons; le *N,O*-acétal est éliminé dans un second temps par acidolyse à l'aide de TFA (figure 34).

<sup>76</sup> Pusterla I.; Bode J. W., *Nat. Chem.*, **2015**, 7, 668-672.

<sup>77</sup> Liu C. F.; Tam J. P., *Proc. Natl. Acad. Sci. Usa*, **1994**, 91, 6584-6588.

<sup>78</sup> Li X.; Lam H. Y.; Zhang Y.; Chan C. K., *Org. Lett.*, **2010**, 12, 1724-1727.





**Figure 34 :** Principe de la ligation native par formation d'une oxazolidine <sup>78</sup>

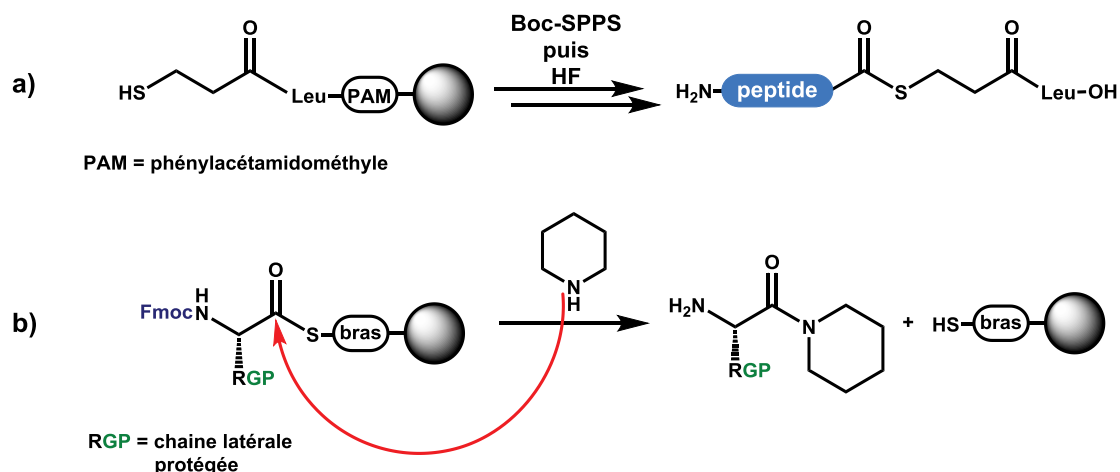
Cette méthode est incompatible avec des esters de salicylaldehyde d'acide aspartique, d'acide glutamique, de cystéine ou de lysine. De plus, les cinétiques observées peuvent être très lentes en fonction de l'acide aminé C-terminal. Néanmoins, cette réaction représente sans doute une alternative viable à la NCL.

### 1.5 Synthèse de peptides thioester par stratégie Fmoc

Nous venons de voir que les peptides C $\alpha$ -thioester sont essentiels dans la réaction de NCL et ses extensions mais aussi, à titre plus anecdotique, dans la ligation de Staudinger native ou encore la ligation thioester assistée par des ions argent.

De tels peptides thioester peuvent être préparés facilement par SPPS en stratégie Boc (figure 35a) mais comme il a été déjà évoqué auparavant, cette technique tend à être de moins en moins utilisée au profit de la stratégie Fmoc, notamment pour des problèmes de sécurité.

Or, la synthèse de peptides C $\alpha$ -thioester en stratégie Fmoc est compliquée par l'incompatibilité de la fonction thioester avec la pipéridine utilisée à chaque cycle pour éliminer le groupement Fmoc. En effet, la pipéridine est suffisamment nucléophile pour attaquer le carbone électrophile du carbonyle de la fonction thioester (figure 35b). La difficulté de synthèse des peptides thioester en stratégie Fmoc représente une autre limitation majeure de la NCL. Ce point constitue la partie centrale de mon sujet de thèse.



**Figure 35 :** a) Synthèse de peptides thioester par Boc-SPPS; b) sensibilité de la fonction thioester lors de l'utilisation de pipéridine en Fmoc-SPPS

Pour pallier ce problème, des équipes ont tenté de développer des conditions d'élimination du groupement Fmoc utilisant des réactifs beaucoup moins nucléophiles que la pipéridine. Par exemple, Guo et ses collaborateurs ont décrit l'utilisation du 1,8-diazabicyclo[5.4.0]undéc-7-ène (DBU), une base non nucléophile, en présence d'HOBt permettant la déprotection de l'amine terminale du fragment peptidique en cours d'élongation.<sup>79</sup> Toutefois, aucune des techniques développées n'empêche totalement la coupure de la fonction thioester après des cycles d'élongation répétés.

Une solution beaucoup plus générale réside dans l'introduction de la fonction thioester C-terminale après élongation de la chaîne peptidique. De très nombreuses techniques reposant sur ce principe ont été développées pour permettre la synthèse de peptides thioester par Fmoc-SPPS. Néanmoins, cette solution est conceptuellement difficile car le carboxyle C-terminal est habituellement lié à la résine par l'intermédiaire d'un bras. La suite de cette section a pour vocation de donner un aperçu large, sans se vouloir exhaustif, des méthodes les plus représentatives d'un point de vue historique et en mettant plus particulièrement en avant les méthodes les plus généralisables.

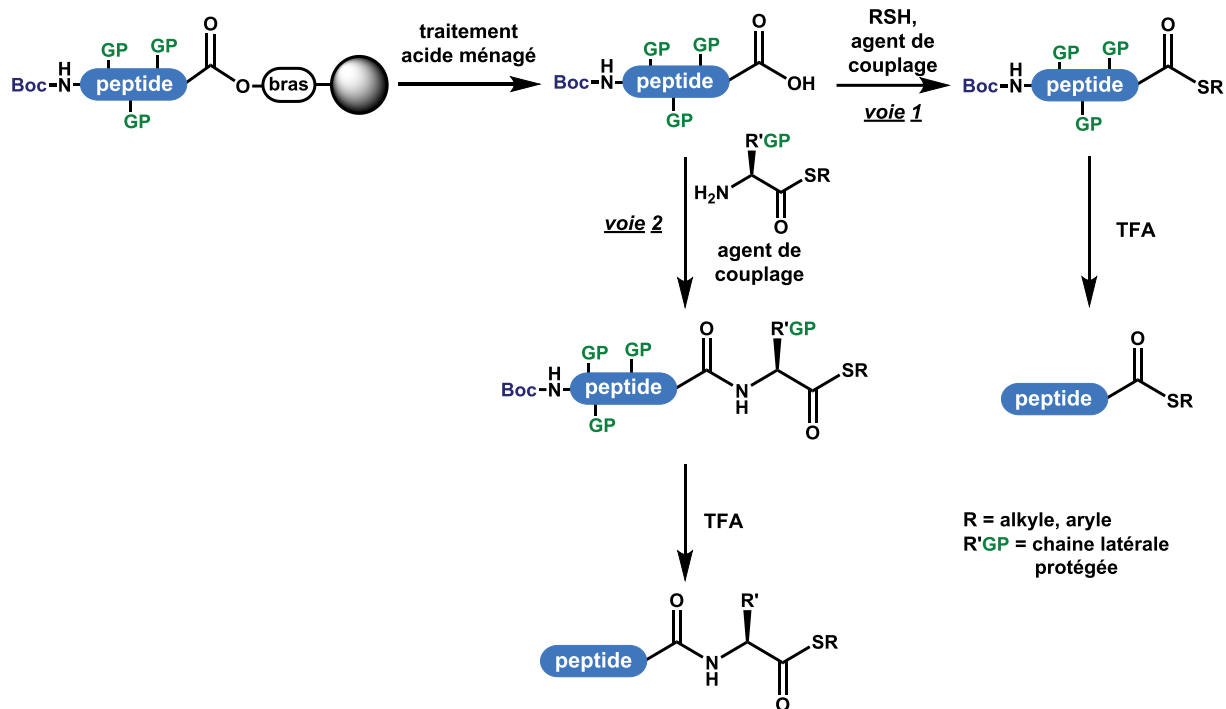
### 1.5.1 Introduction de la fonction thioester sur un peptide protégé en solution

L'approche qui est probablement la plus intuitive est la thioestérification en solution de peptides dont toutes les fonctions autres que le carboxyle C-terminal sont protégées. Le peptide est greffé sur une résine fonctionnalisée par un bras hautement acido-labile (type 2-chlorotrytile ou SASRIN<sup>TM</sup>).<sup>80,81</sup> En fin d'élongation, un traitement acide doux permet la coupure du peptide de

<sup>79</sup> Bu X. Z.; Xie G. Y.; Law C. W.; Guo Z. H., *Tetrahedron Lett.*, **2002**, 43, 2419-2422.

<sup>80</sup> Futaki S.; Sogawa K.; Maruyama J.; Asahara T.; Niwa M.; Hojo H., *Tetrahedron Lett.*, **1997**, 38, 6237-6240.

la résine sans affecter les groupements protecteurs des chaînes latérales. La fonction thioester peut être alors formée en solution, par couplage soit avec un thiol (voie 1, figure 36), soit avec un  $\alpha$ -amino thioester dérivé d'un acide aminé (voie 2, figure 36). Les groupements protecteurs acido-labiles des chaînes latérales sont ensuite éliminés par un traitement au TFA.



**Figure 36 :** Introduction d'une fonction thioester sur un peptide protégé en solution <sup>80,81</sup>

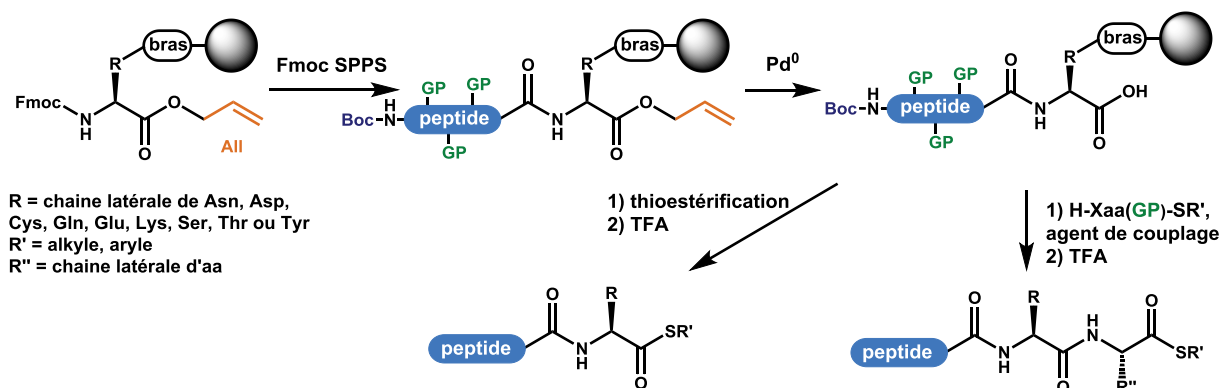
Cette méthode simple à mettre en œuvre est toutefois très limitée, pour des raisons déjà évoquées lors de la description de l'approche de couplage en solution de peptides protégés : de tels fragments peptidiques montrent une solubilité très limitée et l'activation du carboxyle du résidu C-terminal peut entraîner son épimérisation.

### 1.5.2 Greffage par une chaîne latérale

Une approche évitant les problèmes de solubilité inhérents à l'utilisation de peptides protégés repose sur l'introduction de la fonction thioester sur un peptide protégé encore lié au support solide ayant servi à sa synthèse. Une des possibilités consiste à greffer un peptide par le biais de la chaîne latérale d'un de ses acides aminés. L'acide carboxylique C-terminal est protégé par un groupement protecteur orthogonal, tel qu'un groupement allyle, durant l'élongation (figure 37). Après l'assemblage de la chaîne peptidique, le groupement allyle est éliminé à l'aide d'un traitement au palladium (0) puis la fonction thioester est introduite par couplage avec un thiol ou un  $\alpha$ -amino thioester. Le peptide est ensuite décroché et déprotégé par traitement au

<sup>81</sup> Yamamoto N.; Tanabe Y.; Okamoto R.; Dawson P. E.; Kajihara Y., *J. Am. Chem. Soc.*, **2008**, *130*, 501-510.

TFA. Initialement, cette stratégie a été utilisée par greffage d'un peptide par la chaîne latérale d'une sérine ou d'une thréonine.<sup>82</sup> La stratégie a ensuite été étendue à l'ancrage par la chaîne latérale de plusieurs autres acides aminés (Asp, Asn, Cys, Gln, Glu, Lys ou Tyr).<sup>83-85</sup>



**Figure 37 :** Introduction de la fonction thioester d'un peptide greffé sur résine par sa chaîne latérale <sup>82-85</sup>

La thioestérification peut engendrer l'épimérisation de l'acide aminé C-terminal, le couplage d'un acide aminé thioester en position C-terminale réduit ce risque mais ne l'évite pas totalement. Un autre inconvénient de cette approche est la gamme restreinte d'acides aminés pouvant être greffés par leur biais de leur chaîne latérale.

### 1.5.3 Greffage par un amide du squelette peptidique

Une stratégie, très similaire à l'approche décrite précédemment, a été développée dans laquelle le peptide est greffé à la résine par le biais d'un amide de son squelette, à l'aide d'un bras de type *backbone amide linker* (BAL, figure 38).<sup>86,87</sup>

<sup>82</sup> Nakamura K.; Hanai N.; Kanno M.; Kobayashi A.; Ohnishi Y.; Ito Y.; Nakahara Y., *Tetrahedron Lett.*, **1999**, 40, 515-518.

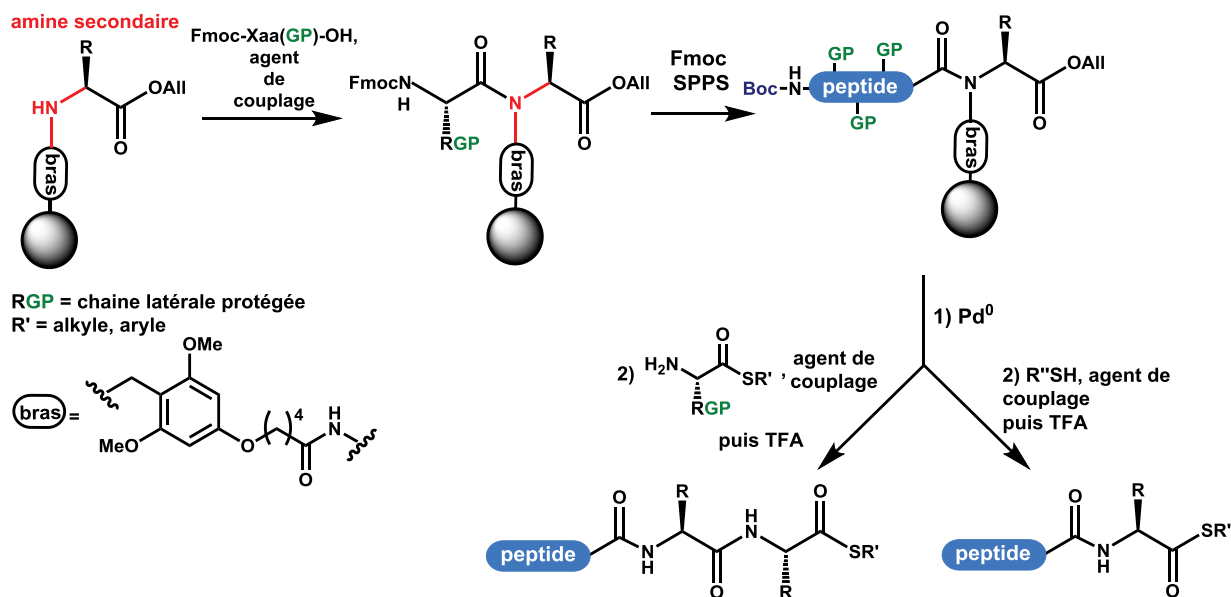
<sup>83</sup> Li L.; Wang P., *Tetrahedron Lett.*, **2007**, 48, 29-32.

<sup>84</sup> Lelièvre D.; Barta P.; Aucagne V.; Delmas A. F., *Tetrahedron Lett.*, **2008**, 49, 4016-4019.

<sup>85</sup> Ficht S.; Payne R. J.; Guy R. T.; Wong C. H., *Chem. Eur. J.*, **2008**, 14, 3620-3629.

<sup>86</sup> Alsina J.; Yokum T. S.; Albericio F.; Barany G., *J. Org. Chem.*, **1999**, 64, 8761-8769.

<sup>87</sup> Gross C. M.; Lelièvre D.; Woodward C. K.; Barany G., *J. Peptide Res.*, **2005**, 65, 395-410.



**Figure 38 :** Introduction d'une fonction thioester sur un peptide greffé par un amide de son squelette<sup>86,87</sup>

Cette stratégie est plus généralisable que l'approche de greffage par la chaîne latérale car le greffage par le squelette d'un peptide est indépendant de sa séquence. En revanche, le couplage entre un acide aminé et l'amine secondaire encombrée, peu nucléophile, pour initier l'élongation est très difficile; ce couplage est donc limité à des acides aminés très faiblement encombrés (Gly, Ala). Ceci représente une limitation majeure de cette stratégie.

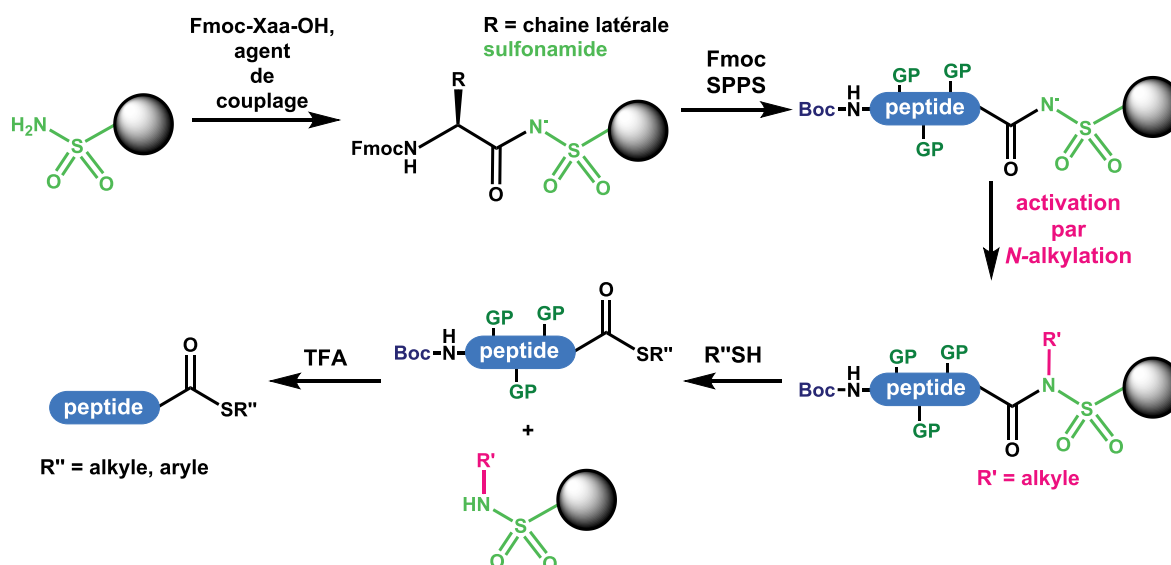
### 1.5.4 Utilisation de *safety-catch linkers*

#### a) Bras sulfonamide de Kenner

Le principe du *safety-catch linker* (ou bras à cran de sécurité) a été introduit pour la SPPS par Kenner et ses collaborateurs en 1971.<sup>88</sup> Par ce principe, le peptide est greffé à la résine par son extrémité C-terminale à la résine *via* un bras sulfonamide en formant une fonction de type imide. Ce bras est stable tout au long de l'élongation car l'atome d'azote de cette fonction est déprotoné dans les conditions de Fmoc-SPPS, évitant ainsi l'attaque d'un nucléophile sur l'imide. En fin de synthèse, la *N*-alkylation du sulfonamide permet de générer un dérivé activé du carboxyle de l'acide aminé C-terminal pouvant réagir avec divers nucléophiles pour conduire de façon concomitante à la coupure du peptide de la résine dans des conditions douces.

<sup>88</sup> Kenner G. W.; McDermott J. R.; Sheppard R. C., *J. Chem. Soc. D*, **1971**, 636-637.

Cette approche a été appliquée à la synthèse de peptides thioester.<sup>89-91</sup> Dans ce cas, l'emploi d'un thiol comme nucléophile permet l'obtention d'un peptide thioester qui est finalement déprotégé par traitement au TFA (figure 39).<sup>90,91</sup>



**Figure 39 :** Synthèse de peptides thioester par utilisation du bras sulfonamide <sup>90,91</sup>

Toutefois, le couplage du premier acide aminé avec la résine sulfonamide est très lent et peut être accompagné d'épimérisation, et les conditions d'alkylation du sulfonamide peuvent entraîner des réactions secondaires comme l'alkylation de la fonction thioéther des méthionines.

### b) Bras hydrazide

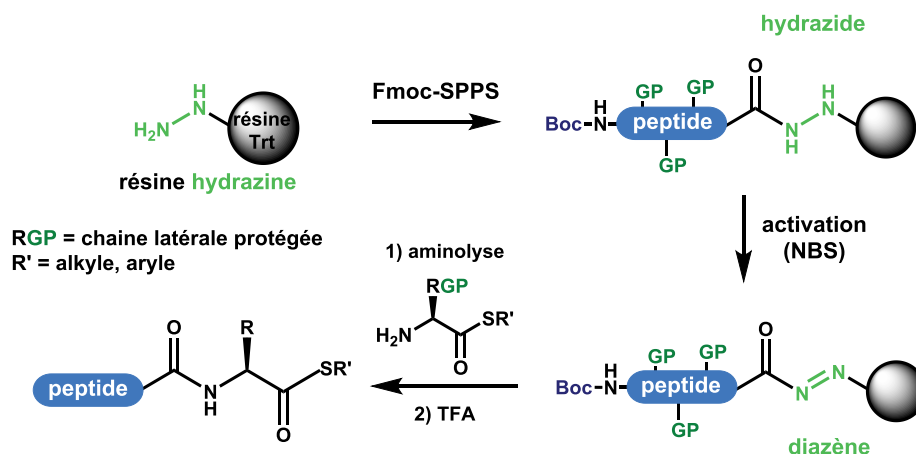
Une autre approche reposant sur un principe de *safety catch linker* consiste en l'utilisation d'une résine hydrazine. Après l'élongation, la fonction hydrazide C-terminale est oxydée à l'aide de *N*-bromosuccinimide (NBS) formant ainsi un *N*-acyle diazène électrophile qui peut réagir avec un  $\alpha$ -amino thioester, ce qui conduit au décrochage du peptide (figure 40).<sup>92</sup> La déprotection du peptide est finalement opérée en solution à l'aide de TFA.

<sup>89</sup> Backes B. J.; Ellman J. A., *J. Org. Chem.*, **1999**, *64*, 2322-2330.

<sup>90</sup> Ingenito R.; Bianchi E.; Fattori D.; Pessi A., *J. Am. Chem. Soc.*, **1999**, *121*, 11369-11374.

<sup>91</sup> Shin Y.; Winans K. A.; Backes B. J.; Kent S. B. H.; Ellman J. A.; Bertozzi C. R., *J. Am. Chem. Soc.*, **1999**, *121*, 11684-11689.

<sup>92</sup> Camarero J. A.; Hackel B. J.; de Yoreo J. J.; Mitchell A. R., *J. Org. Chem.*, **2004**, *69*, 4145-4151.

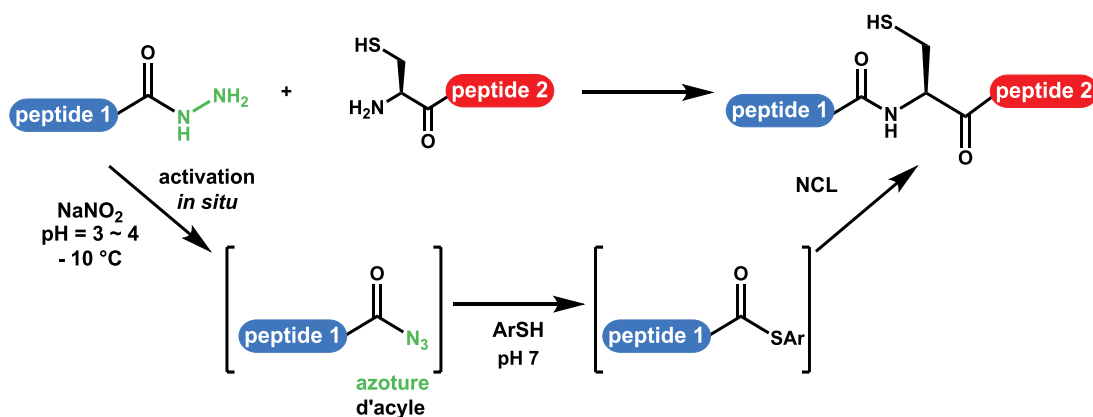


**Figure 40** : Synthèse de peptides thioester via un bras hydrazide <sup>92</sup>

A noter que la thiolysse directe du composé activé sous forme de diazène ne peut pas être réalisée car l'utilisation de thiol peut entraîner la réduction de cette fonction. Cette approche est simple à mettre en œuvre mais les conditions d'oxydation ne sont pas compatibles avec les fonctions thioéther et thiol des méthionines et des cystéines.

### 1.5.5 Peptides hydrazide comme précurseurs de thioester

En 2011, L. Liu et ses collaborateurs ont apporté une modification importante non apparentée au principe de *safety catch linker* de l'approche des peptides hydrazide permettant de les utiliser directement en NCL.<sup>93</sup> L'extrémité hydrazide d'un peptide complètement déprotégé peut être convertie en azoture d'acyle par oxydation en solution puis l'ajout d'un thiol au milieu réactionnel permet la formation *in situ* du peptide thioester dans des conditions de NCL (figure 41).



**Figure 41** : Utilisation directe de peptides hydrazide comme précurseurs de thioester en NCL <sup>93</sup>

<sup>93</sup> Fang G. M.; Li Y. M.; Shen F.; Huang Y. C.; Li J. B.; Lin Y.; Cui H. K.; Liu L., *Angew. Chem. Int. Ed.*, **2011**, 50, 7645-7649.

Les conditions d'oxydation, faisant appel au nitrite de sodium à basse température, sont particulièrement douces. La préparation des peptides hydrazide par Fmoc-SPPS est aisée et cette stratégie a été appliquée à de nombreuses cibles. Par ailleurs, les peptides hydrazide sont des précurseurs désactivés de peptides thioester pouvant être activés au moment voulu pour une ligation. Cette propriété importante a permis l'application de cette approche à l'assemblage de fragments multiples par NCLs séquentielles de N vers C (voir section 1.3.4). De plus, l'approche peut également être appliquée à des NCLs séquentielles de C vers N, ce qui rend possible une stratégie de synthèse par ligations multiples convergentes.<sup>94</sup> Toutefois, l'étape d'oxydation par le nitrite de sodium peut conduire à plusieurs réactions secondaires à partir de l'azoture d'acyle comme son hydrolyse, la formation d'un amide C-terminal ou encore un réarrangement de Curtius formant un isocyanate qui peut réagir avec un thiol pour conduire à une fonction carbamothioate, limitant ainsi cette méthode.<sup>95</sup> Malgré cela, il faut noter que cette approche est une des méthodes les plus utilisées aujourd'hui (voir, par exemple, les références 43, 95-97).<sup>43,95-97</sup>

### 1.5.6 Bras 3,4-diaminobenzoyle (Dbz)

En 2008, Dawson et ses collaborateurs ont introduit un nouveau bras de type 3,4-diaminobenzoyle (Dbz) reposant sur un principe de *safety catch*.<sup>98</sup> L'assemblage de la chaîne sur un support fonctionnalisé par ce bras conduit à l'obtention d'un peptide *o*-aminoanilide (figure 42). Après élongation, l'amine libre du bras est acylée à l'aide de chloroformiate de 4-nitrophényle. Cette étape est suivie d'une cyclisation par *N*-acylation intramoléculaire pour conduire à l'activation de l'amide sous forme de fonction imide *via* la formation d'une *N*-acyl-benzimidazolone (Nbz). Le peptide obtenu après traitement au TFA peut être alors converti en peptide thioester de façon à être isolé, ou alors *in situ* lors de son utilisation directe dans une NCL.

---

<sup>94</sup> Fang G-M.; Wang J-X.; Liu L., *Angew. Chem. Int. Ed.*, **2012**, *51*, 10347-10350.

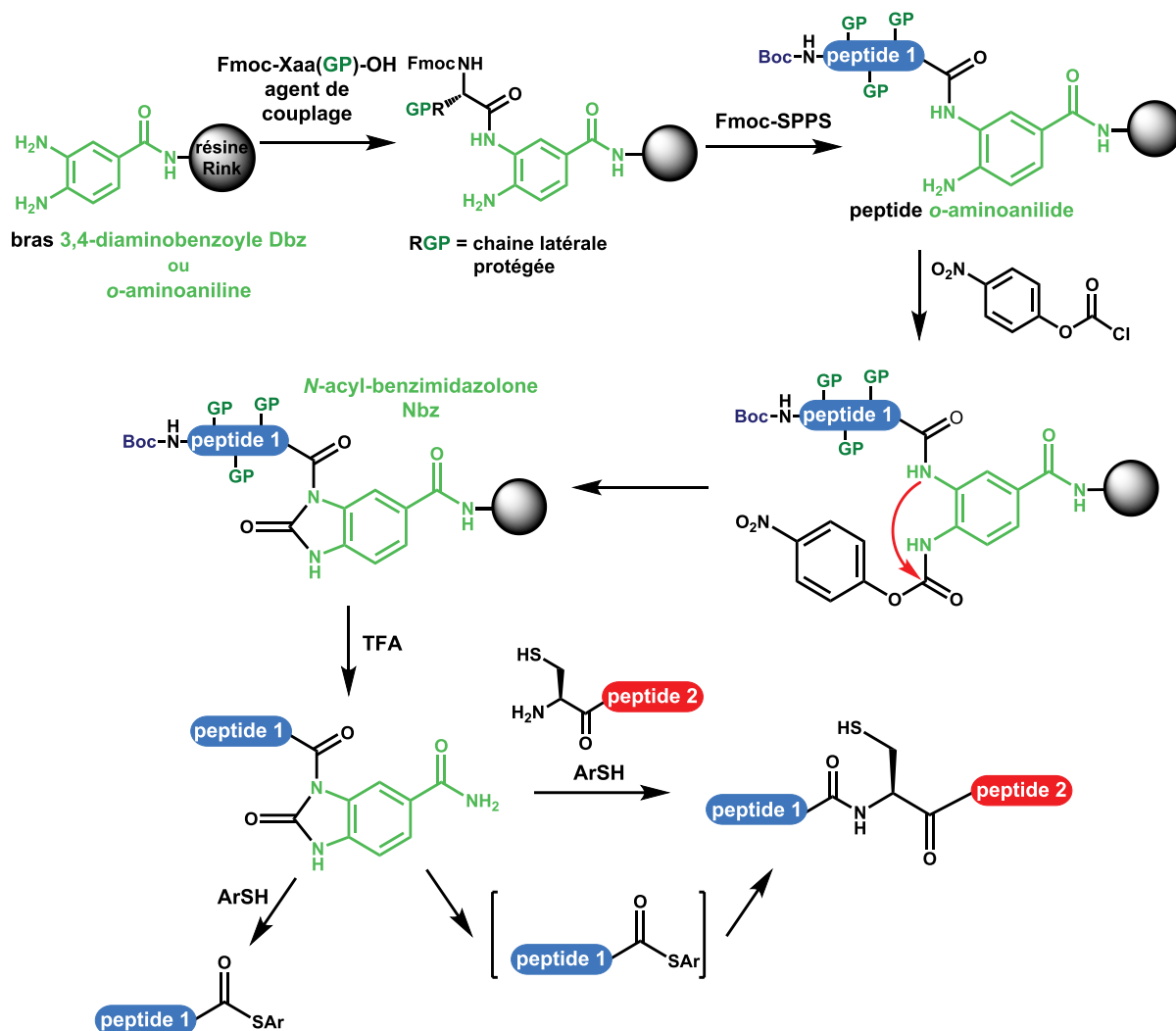
<sup>95</sup> Mong S. K.; Vinogradov A. A.; Simon M. D.; Pentelute B. L., *ChemBioChem*, **2014**, *15*, 721 – 733.

<sup>96</sup> Reif A.; Siebenhaar S.; Tröster A.; Schmälzlein M.; Lechener C.; Velisetty P.; Gottwald K.; Pöhner C.; Boos I.; Schubert V.; Rose-John S.; Unverzagt C., *Angew. Chem. Int. Ed.*, **2014**, *53*, 12125-12131.

<sup>97</sup> Seenayah M.; Jbara M.; Mali S. M.; Brik A., *Angew. Chem. Int. Ed.*, **2015**, *54*, 12374-12378.

<sup>98</sup> Blanco-Canosa J. B.; Dawson P. E., *Angew. Chem. Int. Ed.*, **2008**, *47*, 6851–6855.





**Figure 42 :** Synthèse de peptides *Nbz* et utilisation directe en NCL <sup>98</sup>

Cette approche est elle aussi une des méthodes les plus utilisées de nos jours car la synthèse du précurseur de peptides thioester peut être réalisée aisément (voir, par exemple, les références 99-101).<sup>99-101</sup> Néanmoins, l'utilisation de cette voie pour la synthèse de peptides riches en glycine est limitée car lors d'un couplage, la *N*-Fmoc-glycine peut réagir avec la seconde amine du bras, conduisant à son acylation, ce qui empêche la formation de la fonction imide activée du motif *Nbz*. La réalisation d'étapes d'acétylation après chaque couplage (*capping*) pendant l'élongation peut conduire au même phénomène, ce qui limite grandement cette stratégie.

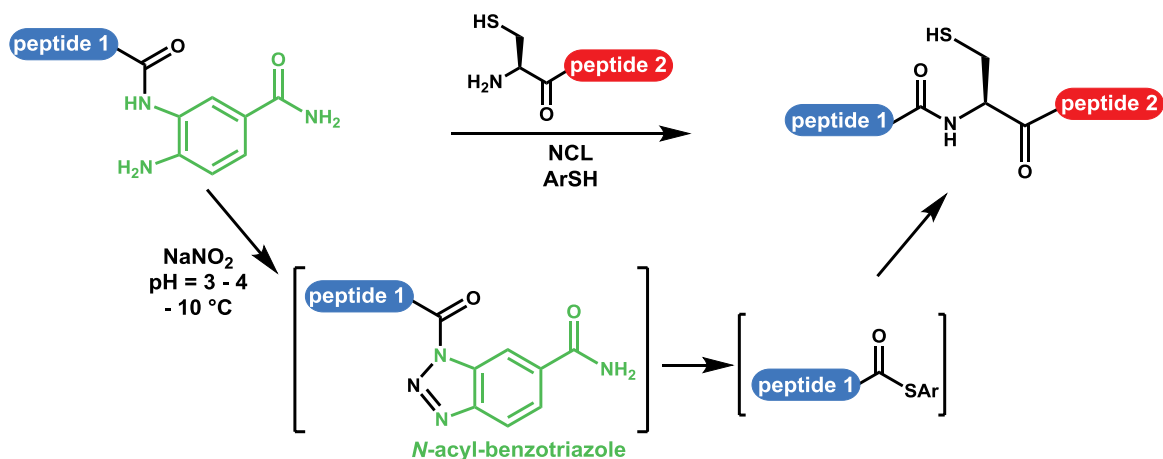
En 2015, l'équipe de L. Liu a apporté une modification à cette stratégie. L'activation en solution d'un peptide *o*-aminoanilide à l'aide de nitrite de sodium conduit à la formation d'un

<sup>99</sup> Lindgren J.; Ekblad C.; Abrahamson L.; Karlstrom A. E., *ChemBioChem*, **2012**, *13*, 1024-1031.

<sup>100</sup> Fauvet B.; Butterfield S. M.; Fuks J.; Brik A.; Lashuel H. A., *Chem. Commun.*, **2013**, *49*, 9254-9256.

<sup>101</sup> Okamoto R.; Mandal K.; Ling M.; Luster A. D.; Kajihara Y.; Kent S. B. H., *Angew. Chem. Int. Ed.*, **2014**, *53*, 5188-5193.

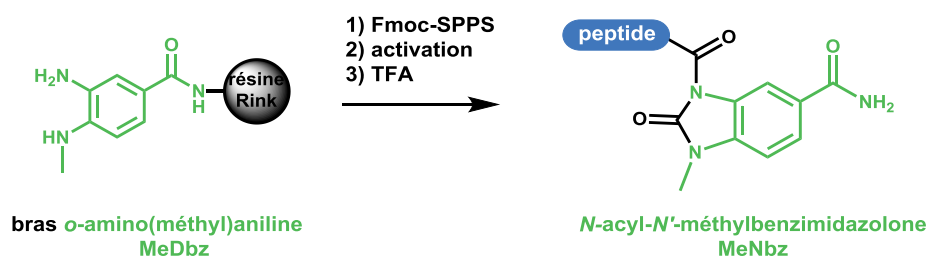
peptide *N*-acyl-benzotriazole qui est converti *in situ* en peptide thioester dans des conditions de NCL (figure 43).<sup>102</sup> Comme les peptides hydrazide, les peptides *o*-aminoanilide utilisés par cette stratégie peuvent être considérés comme des précurseurs désactivés de peptides thioester grâce à leur activation possible en solution. Cette stratégie a donc pu être appliquée à des ligations successives de N vers C.<sup>102</sup>



**Figure 43 :** Activation en solution de peptides *o*-aminoanilide et utilisation directe en NCL <sup>102</sup>

Néanmoins, l'approche de Liu ne résout pas le problème d'acylation parasite de la seconde fonction amine du bras durant l'élongation.

En 2015 également, Dawson et ses collaborateurs ont apporté une amélioration à leur méthode en utilisant une version *N*-méthylée (MeDbz) du bras permettant de réduire la nucléophilie de la seconde fonction amine et ainsi d'éviter son acylation parasite au cours de la SPPS (figure 44).<sup>103</sup>



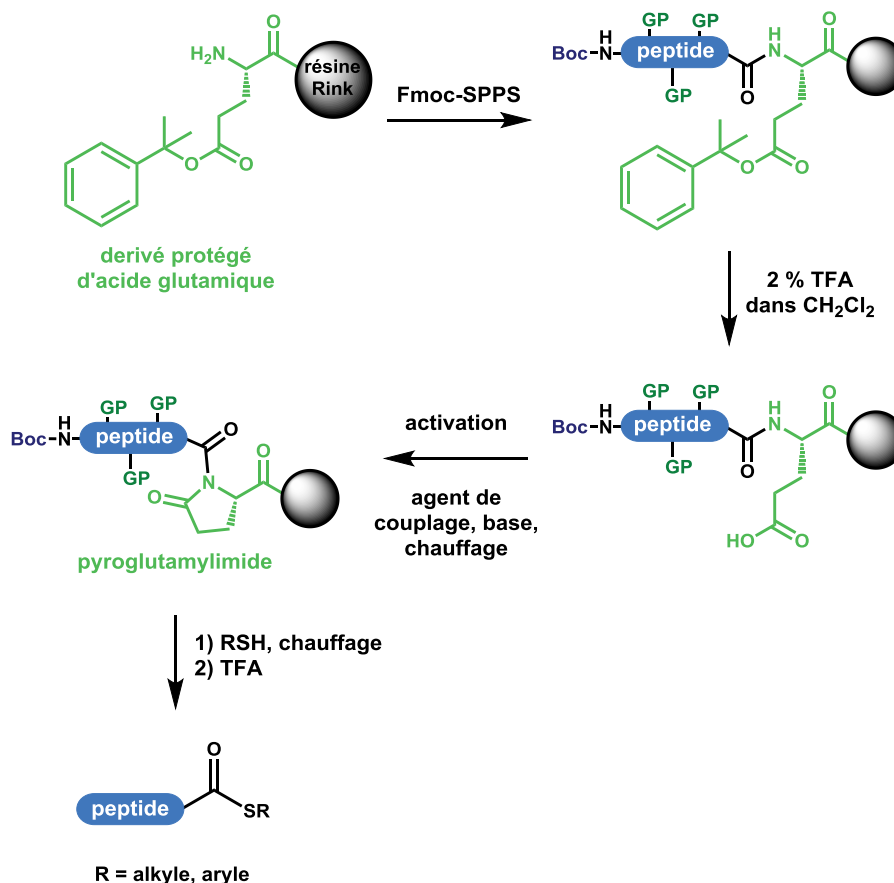
**Figure 44 :** Synthèse de peptides MeNbz via un bras *o*-amino(méthyl)aniline (MeDbz) <sup>103</sup>

<sup>102</sup> Wang J. X.; Fang G. M.; He Y.; Qu D. L.; Yu M.; Hong Z. Y.; Liu L., *Angew. Chem. Int. Ed.*, **2015**, 54, 2194-2198.

<sup>103</sup> Blanco-Canosa J. B.; Nardone B.; Albericio F.; Dawson P. E., *J. Am. Chem. Soc.*, **2015**, 137, 7197-7209.

### 1.5.7 Bras pyroglutamylimide

Jensen et ses collaborateurs ont décrit une méthode, apparentée à celle de Dawson, d'activation de l'amide C-terminal d'un peptide sous forme d'imide.<sup>104</sup> La méthode repose sur l'utilisation d'un dérivé protégé d'acide glutamique en tant que bras dont l'activation conduit à la formation d'un motif pyroglutamylimide par *N*-acylation intramoléculaire. Le traitement du peptide pyroglutamylimide à l'aide d'un thiol permet d'obtenir un peptide thioester et le peptide est finalement déprotégé par un traitement au TFA (figure 45).



**Figure 45 :** Synthèse de peptides thioester par formation d'un pyroglutamylimide <sup>104</sup>

Le motif pyroglutamylimide est cependant peu réactif et sa thiolysé nécessite le chauffage prolongé du milieu réactionnel pouvant être source de dégradation des peptides synthétisés. Cette approche ne semble pas être généralisable.

<sup>104</sup> Tofteng A. P.; Sorensen K. K.; Conde-Frieboes K. W.; Hoeg-Jensen T.; Jensen K. J., *Angew. Chem. Int. Ed.*, **2009**, *48*, 7411-7414.

### 1.5.8 Transfert d'acyle intramoléculaire $O \rightarrow S$ ou $N \rightarrow S$

Depuis le début des années 2000, de très nombreuses méthodes exploitant un transfert d'acyle intramoléculaire  $O \rightarrow S$  ou  $N \rightarrow S$  grâce à un « dispositif moléculaire de thioestérification » C-terminal ont été développées.<sup>105,106</sup> Ces dispositifs sont de types  $\beta$ -mercapto ester (transfert d'acyle  $O \rightarrow S$ ) ou  $\beta$ -mercapto amide  $N,N$ -disubstitué (transfert d'acyle  $N \rightarrow S$ ) sauf dans de rares exceptions (figure 46). Dans la plupart des cas, ils servent de lien entre le peptide et le support solide *via* un bras acido-labile. Après coupure du bras et déprotection du peptide, le groupement thiol du dispositif joue le rôle d'un nucléophile intramoléculaire lors d'un réarrangement qui forme la fonction thioester en passant par un état de transition cyclique à 5 chaînons.

Dans le cas des peptides  $\beta$ -mercapto ester, le transfert d'acyle  $O \rightarrow S$  peut avoir lieu à pH neutre de façon quasi-irréversible. Ces peptides peuvent donc être directement utilisés en NCL, la fonction thioester étant formée *in situ* (figure 46a), ce qui minimise ainsi les étapes post-SPPS nécessaires et constitue un avantage majeur de cette approche.

L'emploi d'un dispositif amide plutôt qu'ester rend ce réarrangement par transfert d'acyle beaucoup moins favorable d'un point de vue thermodynamique vu la très faible électrophilie de la fonction amide causée par la délocalisation du doublet non-liant de l'atome d'azote. De plus, le transfert d'acyle  $N \rightarrow S$  d'un dispositif  $\beta$ -mercapto amide  $N$ -monosubstitué en conformation *trans* est impossible d'un point de vue géométrique, la conformation *cis* est *a priori* cruciale pour que le thiol puisse attaquer le carbonyle électrophile de la liaison amide. Dans le cas d'un amide  $N,N$ -disubstitué, l'isomère *trans* est déstabilisé du fait d'un encombrement stérique important, ce qui favorise un équilibre avec l'isomère *cis* (figure 46b), permettant ainsi la formation de la fonction thioester. Cependant, de manière générale, ce transfert d'acyle  $N \rightarrow S$  est réversible en raison de la nucléophilie de la fonction amine et n'est pas favorisé thermodynamiquement à pH neutre. Les peptides portant ces dispositifs doivent, dans la plupart des cas, être thioestérifiés en milieu acide en présence d'un excès de thiol exogène préalablement à leur utilisation en NCL. Les conditions acides permettent de protoner la fonction amine et de déplacer l'équilibre.

Quelques méthodes utilisant des peptides contenant un dispositif  $\beta$ -mercapto amide C-terminal pouvant opérer dans les conditions de NCL ont été décrites, rendant ainsi possibles les réactions « *one pot* » (formation de la fonction thioester et ligation) (figure 46c). Les peptides répondant à ce critère ont été appelés « peptides crypto-thioester ». Il est à noter que les

<sup>105</sup> Mende F.; Seitz O., *Angew. Chem. Int. Ed.*, **2011**, 50, 1232-1240.

<sup>106</sup> Kawakami T., *Top. Curr. Chem.*, **2015**, 362, 102-135.

cinétiques observées lors de NCLs utilisant ces peptides sont généralement plus lentes que lors de réactions utilisant des peptides thioester pré-formés.

Ces méthodes sont particulièrement attractives car elles sont potentiellement généralisables. Ce domaine est encore en forte expansion aujourd'hui et constitue la partie centrale de mon sujet de thèse.

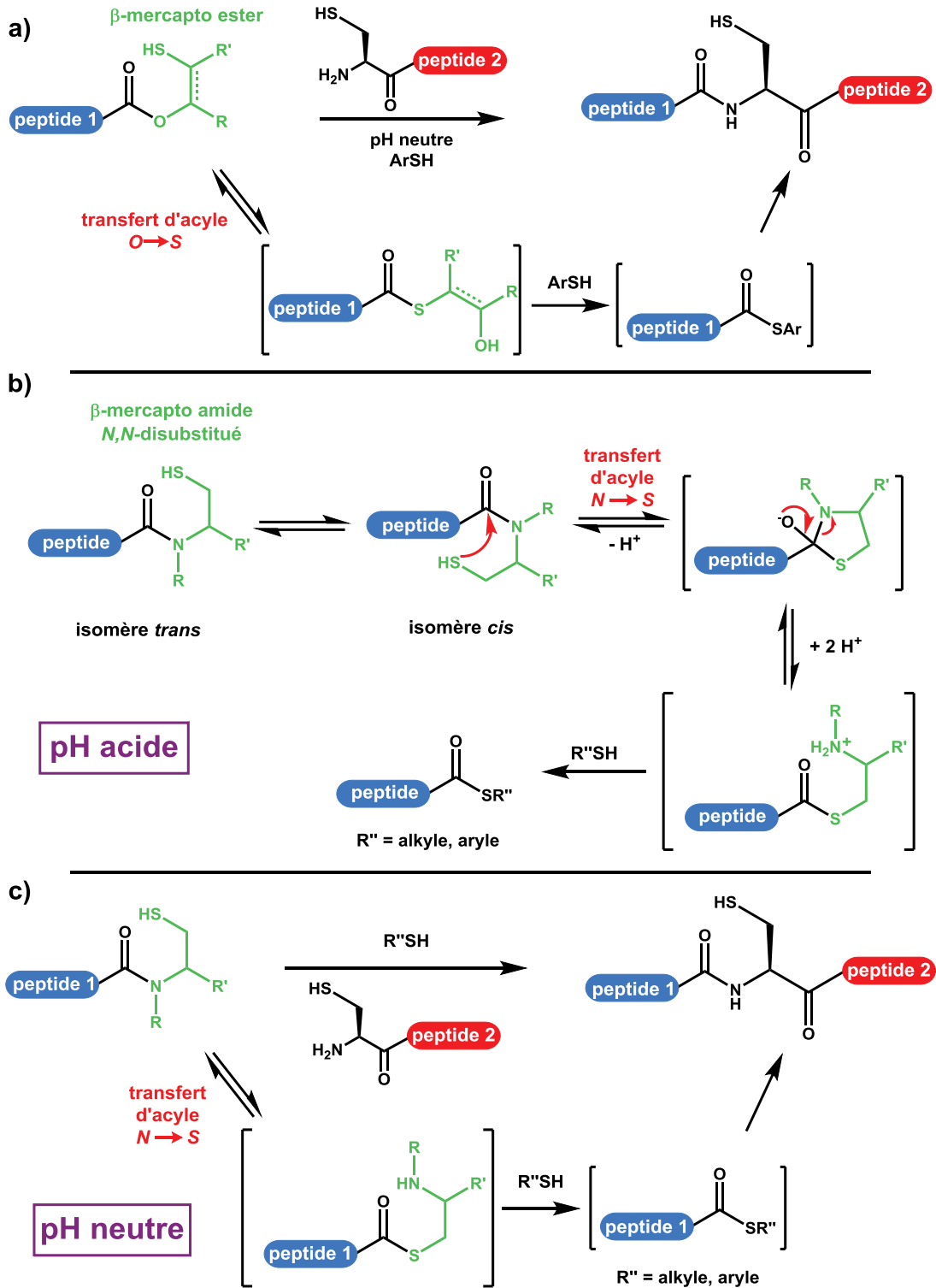


Figure 46 : Dispositifs de thioestérification par transfert d'acyle intramoléculaire  $O \rightarrow S$  ou  $N \rightarrow S$

Une description quasi-exhaustive des méthodes utilisant un dispositif de thioestérification intramoléculaire est développée dans la suite, en trois parties : les méthodes de transfert d'acyle  $O \rightarrow S$ , les méthodes de transfert d'acyle  $N \rightarrow S$  nécessitant une thioestérification en milieu acide et enfin celles impliquant des peptides crypto-thioester.

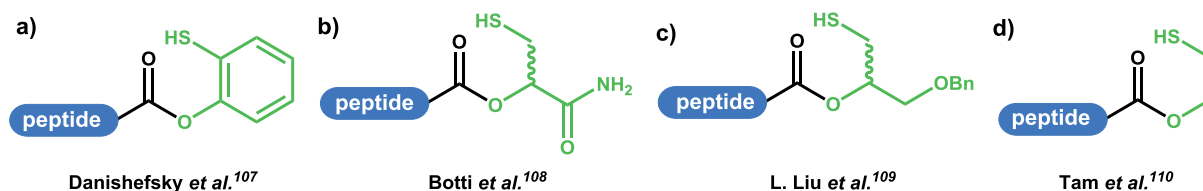
### a) Transfert d'acyle $O \rightarrow S$

Le premier dispositif permettant l'obtention d'une fonction thioester par transfert d'acyle intramoléculaire  $O \rightarrow S$  a été décrit en 2004 par Danishefsky et ses collaborateurs.<sup>107</sup> Ce dispositif, consistant en un ester de 2-mercaptophénol (figure 47a), doit toutefois être installé après SPPS du fait de l'instabilité des esters de phénol dans les conditions utilisées pour l'élongation. De plus, les esters de phénol sont rapidement hydrolysés dans les conditions de ligation.

La même année, Botti et ses collaborateurs ont décrit l'utilisation d'un dispositif de type ester de 2-mercapto-1-carboxamide-éthyle présentant l'avantage d'être stable lors de l'élongation (figure 47b).<sup>108</sup> Néanmoins, ces esters subissent également une réaction secondaire d'hydrolyse lors de la ligation.

L'équipe de L. Liu a réussi à réduire cette hydrolyse indésirable lors de la NCL à un taux inférieur à 10 %, en utilisant un peptide comportant un motif ester de 2-mercaptoéthyle fonctionnalisé par un groupement donneur d'électrons ( $-CH_2OBn$ , figure 47c).<sup>109</sup>

Tam et Eom ont décrit l'utilisation d'un dispositif très simple consistant en un ester de 2-mercaptoéthyle (figure 47d).<sup>110</sup> Ce dispositif peut servir de lien entre le peptide et le support solide lors de la SPPS, par greffage *via* le groupement thiol. Toutefois, cet ester est également instable dans les conditions de NCL et la formation de la fonction thioester doit être réalisée de façon préalable à l'aide d'un thiol exogène en milieu acide, perdant ainsi l'avantage majeur de l'approche de transfert d'acyle  $O \rightarrow S$ .



**Figure 47 :** Dispositifs de thioestérification par transfert d'acyle  $O \rightarrow S$ <sup>107-110</sup>

<sup>107</sup> Warren J. D.; Miller J. S.; Keding S. J.; Danishefsky S. J., *J. Am. Chem. Soc.*, **2004**, *126*, 6576-6578.

<sup>108</sup> Botti P.; Villain M.; Manganiello S.; Gaertner H., *Org. Lett.*, **2004**, *6*, 4861-4864.

<sup>109</sup> Zheng J. S.; Cui H. K.; Fang G. M.; Xi W. X.; Liu L., *ChemBioChem*, **2010**, *11*, 511-515.

<sup>110</sup> Eom K. D.; Tam J. P., *Org. Lett.*, **2011**, *13*, 2610-2613.

L'hydrolyse parasite récurrente des peptides ester dans les conditions de la NCL limite grandement la portée de ces stratégies. Le développement de peptides portant un dispositif de thioestérification par transfert d'acyle intramoléculaire  $O \rightarrow S$  dont la fonction ester serait stable à l'hydrolyse dans des conditions de NCL constituerait une grande avancée.

### *b) Transfert d'acyle $N \rightarrow S$ en milieu acide*

Contrairement à une fonction ester, la fonction amide n'est pas facilement hydrolysable, ce qui représente un avantage majeur des méthodes de transfert d'acyle  $N \rightarrow S$  par rapport aux méthodes de transfert d'acyle  $O \rightarrow S$ .

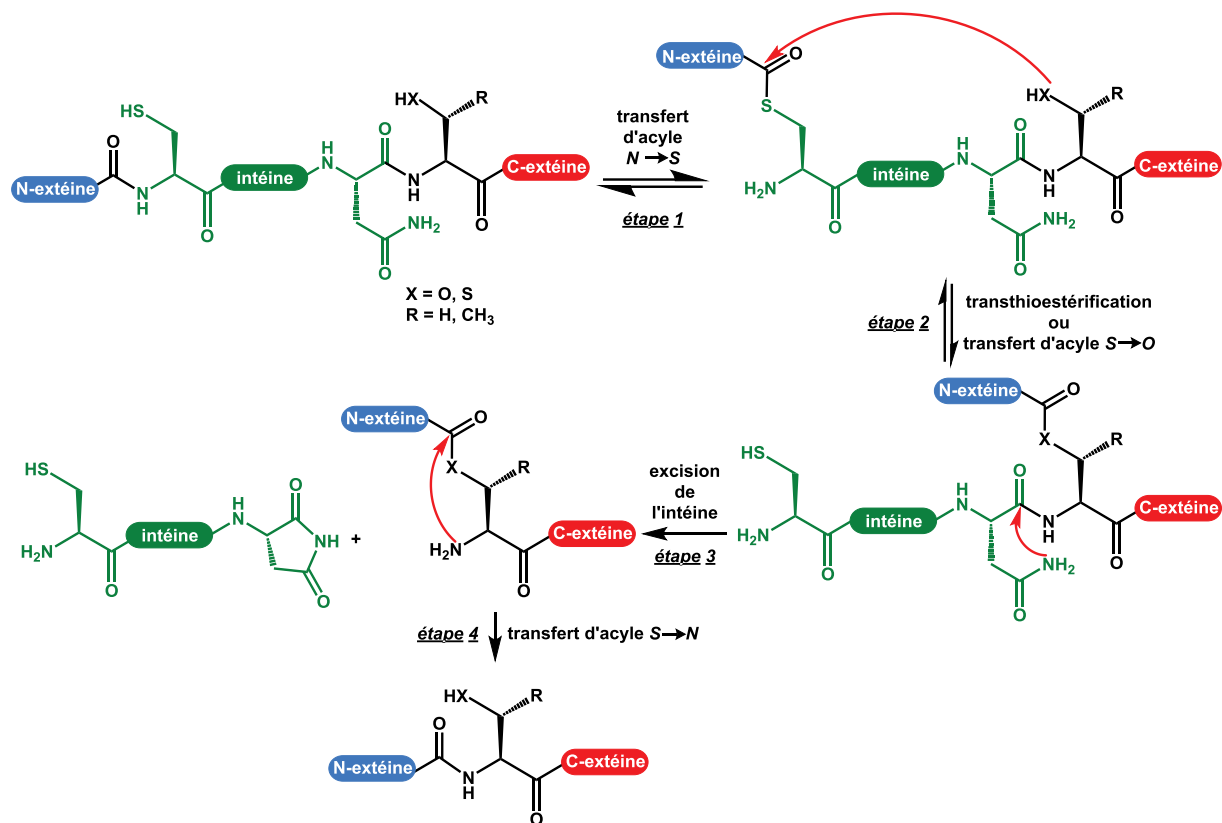
#### ▪ Les intéines

Le mécanisme de transfert d'acyle intramoléculaire  $N \rightarrow S$ , qui est le mécanisme inverse de la NCL, a été observé en 1990 dans des protéines bactériennes possédant un domaine autocatalytique appelé « intéine ».<sup>111</sup> Les intéines participent à une modification post-traductionnelle (épissage protéique) consistant en la coupure et la formation de liaisons peptidiques.<sup>112</sup> Les intéines sont capables de s'exciser de la protéine hôte en joignant leurs fragments externes appelés N- et C-extéines (figure 48). Certains acides aminés hautement conservés participent à cette réaction (Cys C-terminale, Asn N-terminale et une His interne de l'intéine, Cys, Ser ou Thr en position N-terminale de la C-extéine).<sup>112</sup> La première étape de l'excision de l'intéine consiste en la formation d'une fonction thioester entre la N-extéine et l'intéine par un transfert d'acyle  $N \rightarrow S$  (étape 1, figure 48). Une transthioestérification ou un transfert d'acyle  $S \rightarrow O$  a ensuite lieu entre le thioester précédemment formé et le thiol ou l'alcool du résidu N-terminal de la C-extéine (étape 2, figure 48). L'excision de l'intéine procède ensuite par formation d'un succinimide (étape 3, figure 48). Enfin, un transfert d'acyle  $O$  ou  $S \rightarrow N$  conduit à la formation de la liaison peptide entre la N- et la C-extéine par un mécanisme similaire à celui de la NCL (étape 4, figure 48).<sup>112</sup>

---

<sup>111</sup> Mills K. V.; Johnson M. A.; Perler F. B., *J. Biol. Chem.*, **2014**, 289, 14498-14505.

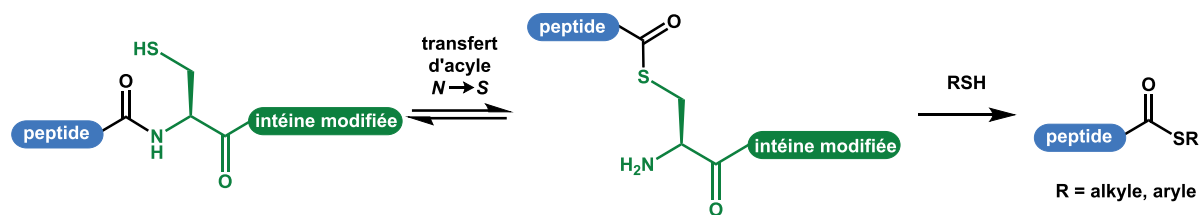
<sup>112</sup> Shah N. H.; Muir T. W., *Chem. Sci.*, **2014**, 5, 446-461.



**Figure 48 :** Mécanisme d'excision des intéines <sup>112</sup>

Cette machinerie autoenzymatique a d'ailleurs été mise à profit dans une approche de synthèse de protéines par NCL combinant des fragments préparés par voie chimique et des fragments préparés par voie recombinante. Cette stratégie a été nommée « *expressed protein ligation* » (EPL).<sup>113</sup>

La synthèse de peptides thioester par voie recombinante utilise des intéines modifiées de façon à inhiber l'étape de transthioestérification ou le transfert d'acyle  $S \rightarrow O$  (figure 49).<sup>112,113</sup>



**Figure 49 :** Synthèse de peptides thioester par EPL <sup>113</sup>

Cette approche élégante permet la combinaison des avantages des synthèses de protéines par voie chimique et par voie recombinante mais sa mise en œuvre peut être ardue.

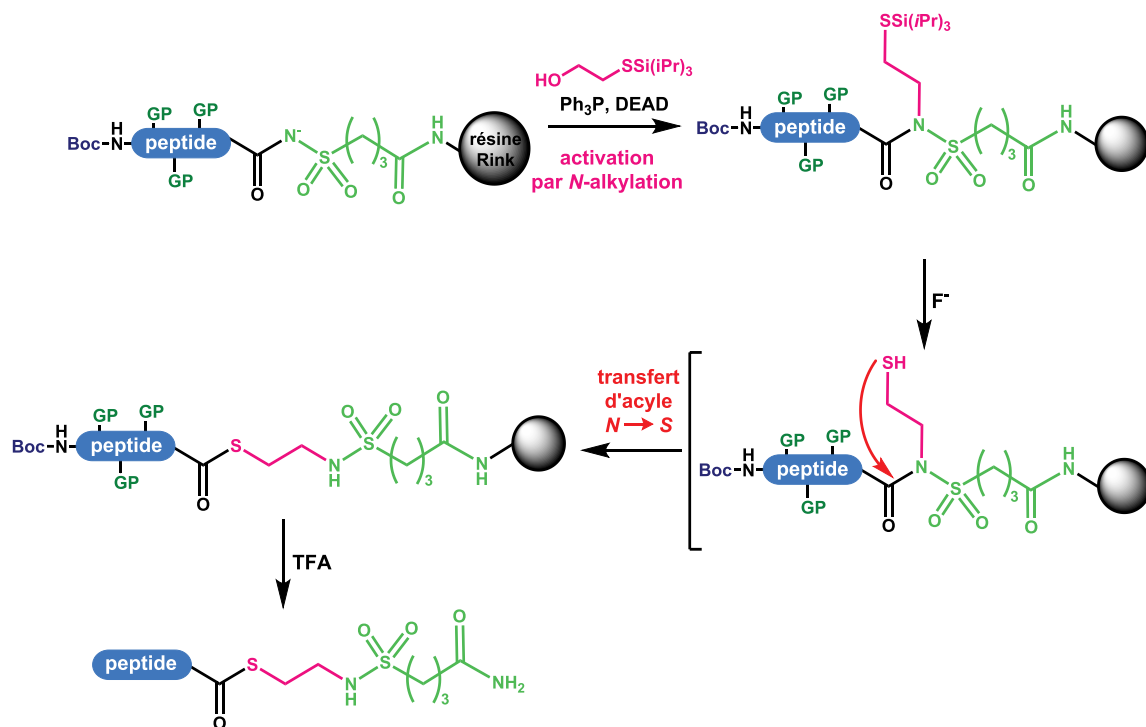
<sup>113</sup> Muir T. W.; Sondhi D.; Cole P. A., *Proc. Natl. Acad. Sci. Usa*, **1998**, 95, 6705-6710.



Comme cela a été dit précédemment, ce mécanisme de transfert d'acyle  $N \rightarrow S$  est exploité dans de nombreuses méthodes de synthèse de peptides thioester et ce domaine est encore actuellement en vive expansion.

- Bras sulfonamide activé

En 2005, l'équipe de Melnyk a développé une variante intramoléculaire de l'utilisation du *safety catch linker* sulfonamide (décrite en 1.5.4).<sup>114</sup> Après élongation, la fonction  $N$ -acyle sulfonamide est activée par  $N$ -alkylation *via* une réaction de Mitsunobu à l'aide d'un dérivé  $S$ -protégé du 2-mercaptoéthanol (figure 50). Le groupement thiol est déprotégé par traitement avec des ions fluorure et peut ensuite intervenir dans le transfert d'acyle intramoléculaire  $N \rightarrow S$ . Si cet exemple peut être considéré comme précurseur des méthodes contemporaines de synthèse de thioester par transfert d'acyle  $N \rightarrow S$ , il représente un cas particulier, car l'espèce réactive est un dérivé électroniquement activé de type imide plutôt qu'un amide.



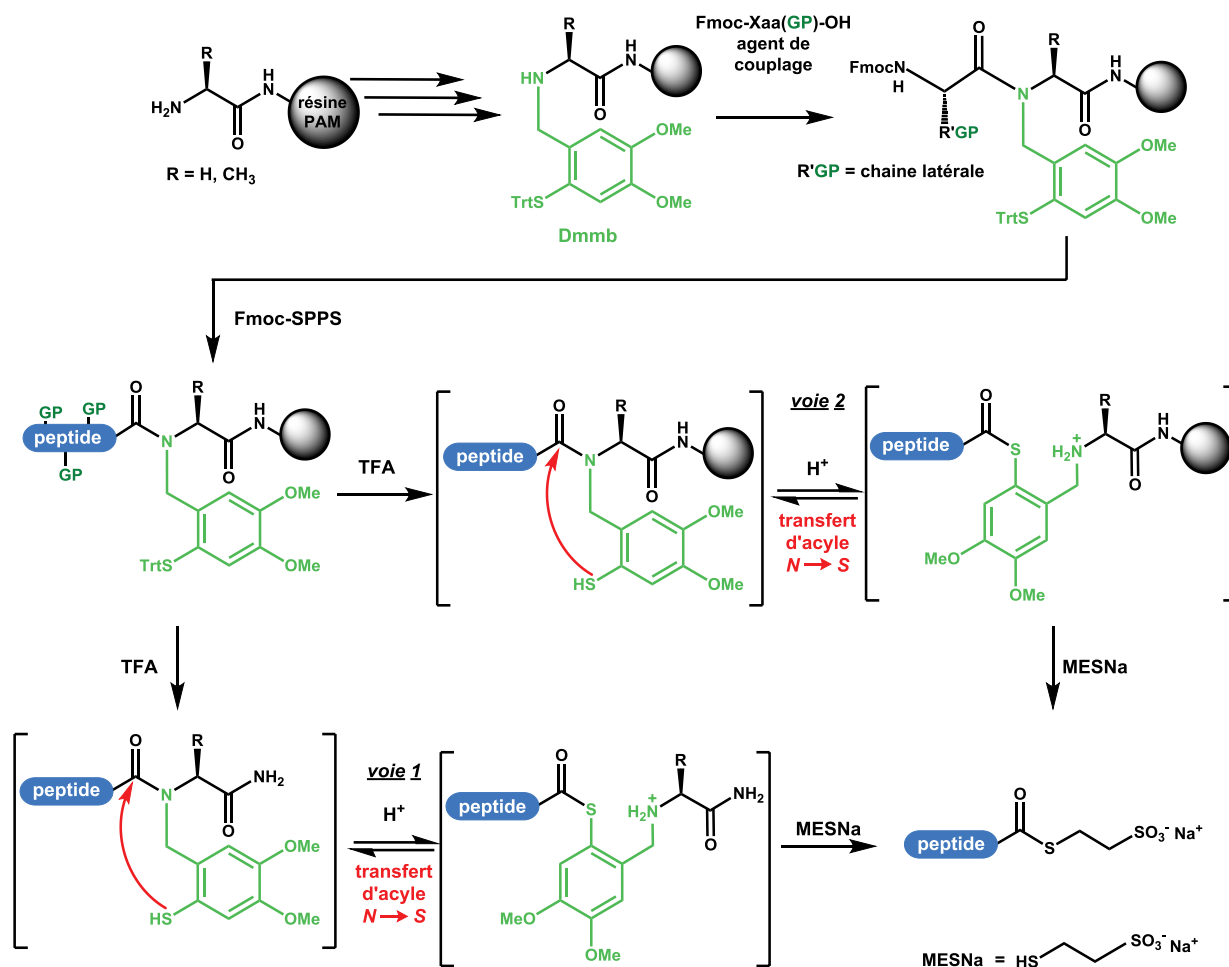
**Figure 50 :** Utilisation du bras sulfonamide<sup>114</sup>

Les inconvénients de cette méthode sont inhérents à l'utilisation du bras sulfonamide (installation difficile du premier acide aminé C-terminal, activation difficile). De plus, le développement de cette stratégie ne semble pas avoir été poursuivi et son application s'est limitée à des peptides modèles.

<sup>114</sup> Ollivier N.; Behr J. B.; El-Mahdi O.; Blanpain A.; Melnyk O., *Org. Lett.*, **2005**, 7, 2647-2650.

▪ *N*-4,5-Diméthoxy-2-mercaptobenzyle amide (Dmmb)

En 2005, Vorherr et ses collaborateurs ont mis au point la première méthode de transfert d'acyle intramoléculaire  $N \rightarrow S$  due à un dispositif amide  $N,N$ -disubstitué non activé électroniquement, faisant appel à un groupement *N*-4,5-diméthoxy-2-mercaptobenzyle (Dmmb) ayant été initialement développé sans succès comme auxiliaire pour la NCL (voir section 1.3.5b).<sup>115</sup> Ce premier exemple est le seul cas faisant intervenir un dispositif  $\gamma$ -mercapto amide (figure 51). Le transfert d'acyle  $N \rightarrow S$  peut avoir lieu en solution (voie 1, figure 51) ou sur résine, en utilisant dans ce cas un bras stable au TFA (voie 2, figure 51).



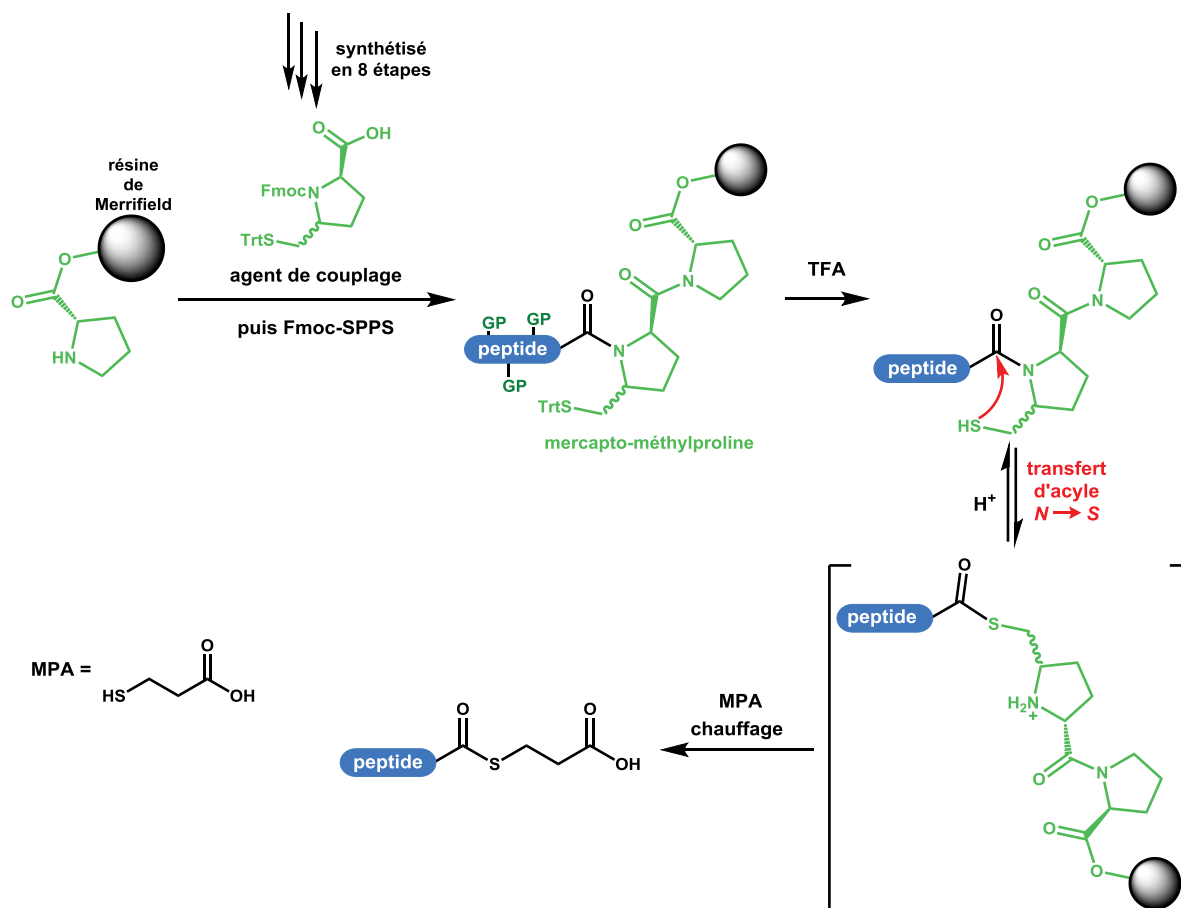
**Figure 51** : Utilisation de l'auxiliaire Dmmb <sup>115</sup>

Cette stratégie comporte plusieurs limitations dont la nécessité de plusieurs étapes de synthèse en solution de l'auxiliaire avant de fonctionnaliser la résine. De plus, le couplage du premier acide aminé sur l'auxiliaire est difficile, restreignant cette approche aux acides aminés C-terminaux présentant un faible encombrement stérique.

<sup>115</sup> Kawakami T.; Sumida M.; Nakamura K.; Vorherr T.; Aimoto S., *Tetrahedron Lett.*, **2005**, 46, 8805-8807.

- 5-Mercaptométhylproline

En 2006, Nakahara et ses collaborateurs ont été les premiers à proposer un dispositif  $\beta$ -mercapto amide  $N,N$ -disubstitué pour un transfert d'acyle intramoléculaire  $N \rightarrow S$ . Leur dispositif se compose d'un motif 5-mercaptométhylproline (figure 52).<sup>116</sup> Après élongation, la formation de la fonction thioester a lieu sur support solide et le peptide thioester peut être isolé après réaction avec un thiol tel que l'acide 3-mercaptopropionique (MPA) en milieu acide (figure 52).



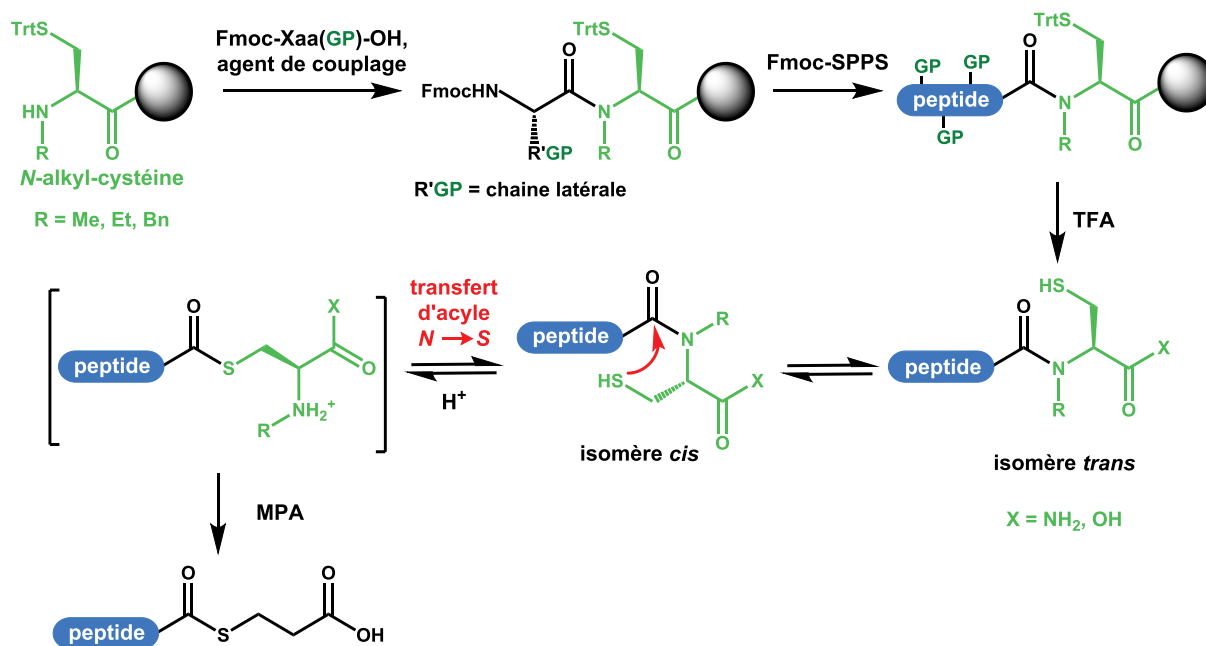
**Figure 52 :** Utilisation d'une 5-mercaptométhylproline <sup>116</sup>

La préparation du motif 5-mercaptométhylproline nécessite plusieurs étapes laborieuses de synthèse en solution. De plus, ce motif est peu réactif, sa thioestérification par réaction avec le MPA demande de très longs temps de réaction ou l'utilisation de conditions sévères de chauffage. Cette méthode est, de ce fait, peu utilisée.

<sup>116</sup> Nagaïke F.; Onuma Y.; Kanazawa C.; Hojo H.; Ueki A.; Nakahara Y.; Nakahara Y., *Org. Lett.*, **2006**, *8*, 4465-4468.

- *N*-Alkyl-cystéine

En 2007, Hojo et ses collaborateurs ont décrit l'utilisation de peptides comportant une *N*-alkyl-cystéine C-terminale.<sup>117</sup> Il faut noter que le groupement alkyle s'installe facilement sur la cystéine par une réaction d'amination réductrice. Après élongation et traitement TFA, le peptide obtenu est traité par du MPA en milieu acide pour obtenir le peptide thioester (figure 53).



**Figure 53 :** Utilisation d'une *N*-alkyl-cystéine <sup>117</sup>

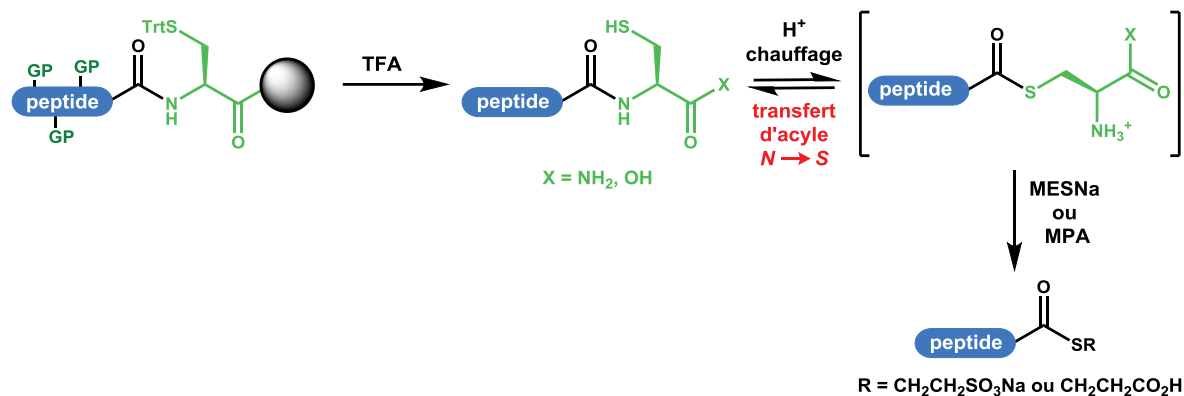
La thioestérification d'un peptide comportant une *N*-alkyl-cystéine C-terminale à l'aide de MPA est très lente. De plus, le couplage du premier acide aminé sur la cystéine *N*-alkylée est très difficile, nécessitant l'emploi d'agents de couplage puissants ou de chauffage.

- Cystéine

En 2009, Macmillan et ses collaborateurs ont exploité un phénomène de transfert d'acyle *N*→*S* spontané observé dans une séquence peptidique non modifiée lors de son chauffage prolongé en milieu acide.<sup>118</sup> La méthode développée consiste en l'utilisation de peptides composés d'une cystéine C-terminale pouvant se comporter comme un dispositif de thioestérification de type  $\beta$ -mercapto amide non disubstitué si elle est précédée d'une glycine, histidine ou d'une seconde cystéine, par réaction avec un thiol en milieu acide à haute température (60-70 °C) (figure 54).

<sup>117</sup> Hojo H.; Onuma Y.; Akimoto Y.; Nakahara Y.; Nakahara Y., *Tetrahedron Lett.*, **2007**, 48, 25-28.

<sup>118</sup> Kang J.; Richardson J.P.; Macmillan D., *Chem. Commun.*, **2009**, 407-409.



**Figure 54 :** Utilisation d'une cystéine C-terminale <sup>118</sup>

L'obtention aisée du peptide précurseur présentant une séquence native (par Fmoc-SPPS classique ou voie recombinante) représente l'atout majeur de cette stratégie.

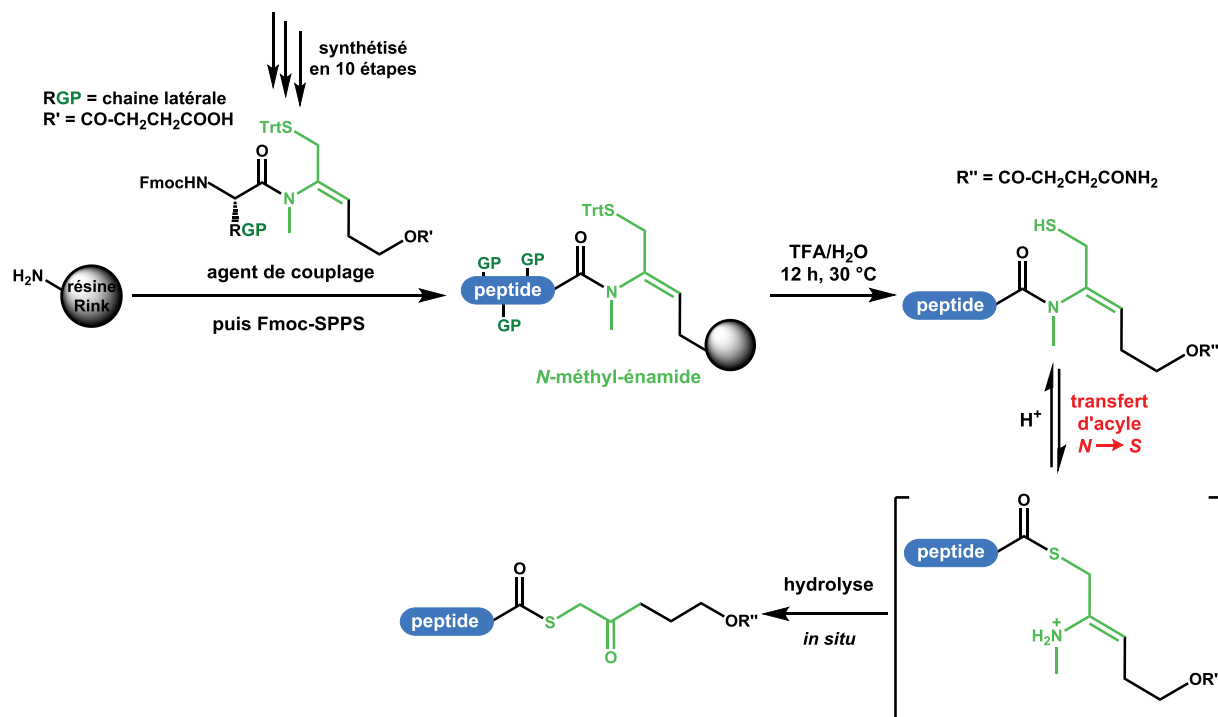
L'applicabilité de cette méthode a par la suite été étendue à d'autres sites (Arg, Leu, Phe, Ser et Val-Cys).<sup>119</sup> Néanmoins, la thioestérification est très lente car la formation de la liaison thioester est beaucoup moins favorisée que dans le cas d'un dispositif  $\beta$ -mercapto amide *N,N*-disubstitué et cette réaction peut nécessiter des conditions réactionnelles sévères (haute température). De plus, cette stratégie reste limitée à une gamme restreinte d'acides aminés C-terminaux et n'a pas été utilisée en dehors du laboratoire où elle a été développée.

- Formation irréversible de la fonction thioester à partir de *N*-méthyl-énamide

L. Liu et ses collaborateurs ont décrit une méthode de transfert d'acyle *N*→*S* rendu irréversible par la conception du dispositif de thioestérification.<sup>120</sup> Ce dispositif C-terminal est de type  $\beta$ -mercapto-*N*-méthyl-énamide (figure 55). La déprotection du peptide, la rupture de son lien avec la résine et le transfert d'acyle *N*→*S* ont lieu de façon concomitante lors d'un traitement avec un mélange TFA/eau. L'énamine formée est hydrolysée *in situ* empêchant la réversibilité du transfert d'acyle.

<sup>119</sup> Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.

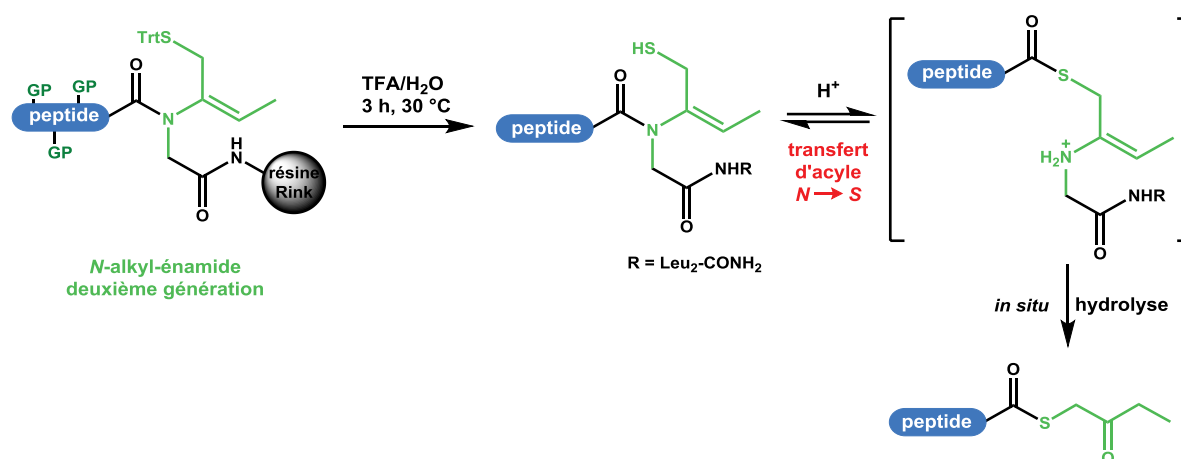
<sup>120</sup> Zheng J. S.; Chang H. N.; Wang F. L.; Liu L., *J. Am. Chem. Soc.*, **2011**, *133*, 11080-11083.



**Figure 55 :** Synthèse de peptides thioester à partir d'un N-méthyl-énamide <sup>120</sup>

Si cette méthode est très élégante, la synthèse en solution du motif N-méthyl-énamide est très fastidieuse. De plus, la formation du thioester est très lente et nécessite un traitement au TFA très prolongé.

En 2014, Zheng et ses collaborateurs ont décrit l'utilisation d'un nouveau dispositif énamide, fonctionnant de la même façon que le précédent, dont la thioestérification peut être réalisée plus rapidement, à l'aide d'un traitement TFA plus court (figure 56).<sup>121</sup>



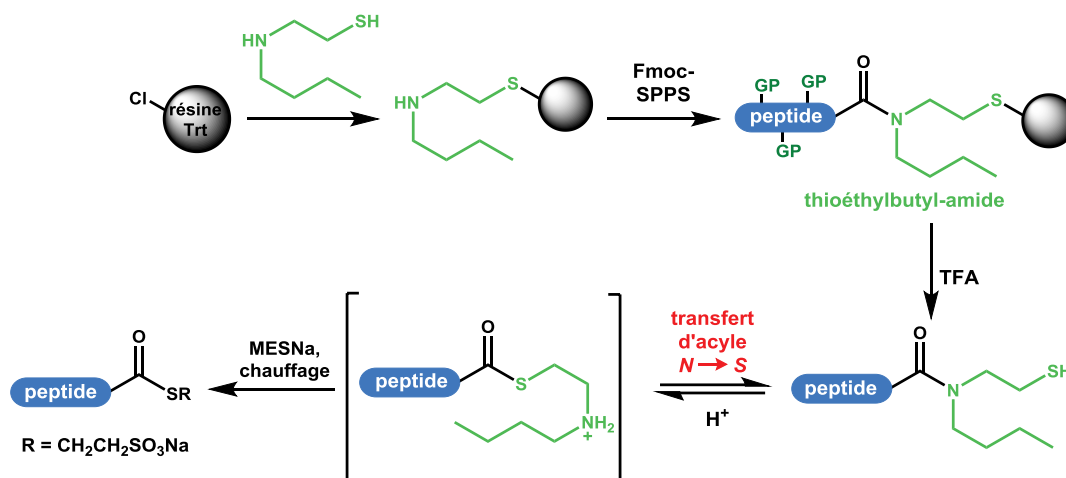
**Figure 56 :** N-alkyl-énamide de seconde génération <sup>121</sup>

<sup>121</sup> Zheng J. S.; Chen X.; Tang S.; Chang H. N.; Wang F. L.; Zuo C., *Org. Lett.*, **2014**, *16*, 4908-4911.

La synthèse en solution du motif énamide de seconde génération nécessite également plusieurs étapes laborieuses, représentant toujours un inconvénient limitant l'application de cette stratégie.

- Thioéthylalkyl-amide (TEA)

En 2013, Tam et ses collaborateurs ont décrit l'utilisation d'un dispositif thioéthylalkyl-amide (TEA) C-terminal.<sup>122</sup> Le motif est greffé par son groupement thiol sur la résine (figure 57). Après élongation et libération du peptide à l'aide de TFA, la thioestérification est réalisée de façon quasi-quantitative par une réaction avec le MESNa en milieu acide.



**Figure 57** : Utilisation d'un thioéthylalkyl-amide <sup>122</sup>

Cette stratégie semble intéressante au vu de la simplicité de synthèse du peptide précurseur par Fmoc-SPPS. Toutefois, les rendements obtenus après traitement TFA sont faibles. De plus, la stratégie n'a pas été appliquée à des acides aminés encombrés stériquement, dont le couplage avec l'amine secondaire du dispositif, pourrait être difficile.

### c) Peptides crypto-thioester

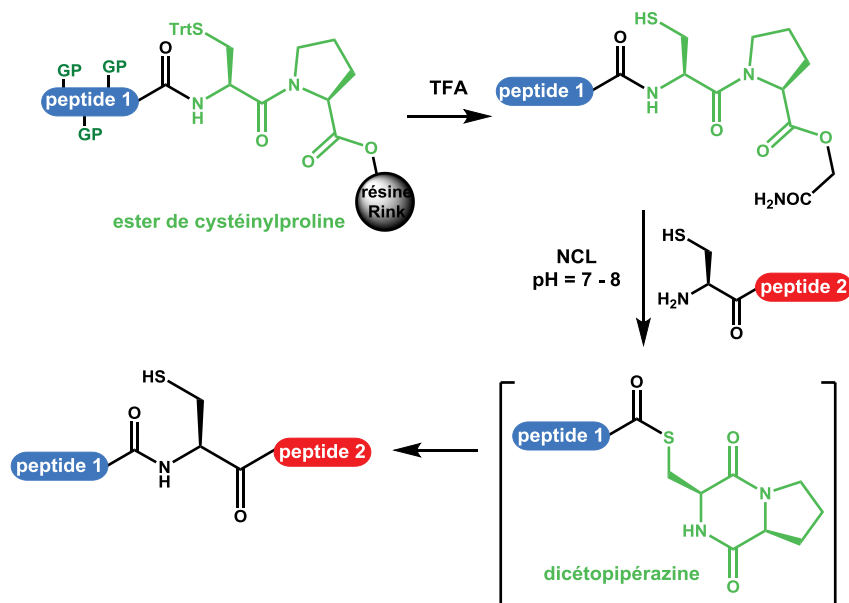
Le développement de méthodes de synthèse de peptides dits « crypto-thioester » utilisant un transfert d'acyle intramoléculaire  $N \rightarrow S$  connaît un fort engouement depuis 2009. Cette approche est particulièrement intéressante car elle minimise les étapes post-SPPS nécessaires avant une NCL. De plus, certains dispositifs peuvent également jouer le rôle de précurseurs désactivés de fonction thioester permettant des NCLs multiples successives de N vers C. Outre les dispositifs présentés dans cette partie, il faut noter que les peptides obtenus par la méthode des hydrazides de L. Liu (section 1.5.5) ou du bras Dbz de Dawson (section 1.5.6)

<sup>122</sup> Taichi M.; Hemu X.; Qiu Y.; Tam J. P., *Org. Lett.*, **2013**, *15*, 2620-2623.

peuvent être également considérés comme des peptides crypto-thioester même si ces méthodes ne reposent pas sur un transfert d'acyle  $N \rightarrow S$ .

- Peptides ester de cystéinyl proline (CPE) et cystéinyl prolyl cystéine (CPC)

Kawakami et Aimoto ont décrit l'utilisation de peptides comportant un motif ester de cystéinyl proline (CPE) en tant que peptides crypto-thioester (figure 58).<sup>123</sup> Ce dispositif n'est pas un amide disubstitué mais la réaction de transfert d'acyle  $N \rightarrow S$  est rendue possible à pH neutre car l'équilibre menant à la formation de la fonction thioester est déplacée par un réarrangement irréversible spontané formant une dicétopipérazine. La thioestérification d'un peptide CPE peut donc avoir lieu *in situ* dans des conditions de NCL.



**Figure 58 :** Utilisation de peptides CPE <sup>123</sup>

Néanmoins, l'acide aminé C-terminal porté par le motif CPE peut être épimérisé lors de la formation du thioester qui nécessite des conditions légèrement basiques.

Kawakami et Aimoto ont également décrit l'utilisation d'un peptide comportant un motif cystéinyl prolyl cystéine C-terminal (CPC) réagissant de la même façon.<sup>124</sup> Cette approche est élégante car le peptide précurseur présente une séquence native et peut être obtenu par voie recombinante. Toutefois, le traitement nécessaire pour obtenir le peptide thioester consiste en un chauffage à 110 °C en milieu acide, conditions pouvant donner lieu à des réactions secondaires.

<sup>123</sup> Kawakami T.; Aimoto S., *Tetrahedron*, **2009**, 65, 3871-3877.

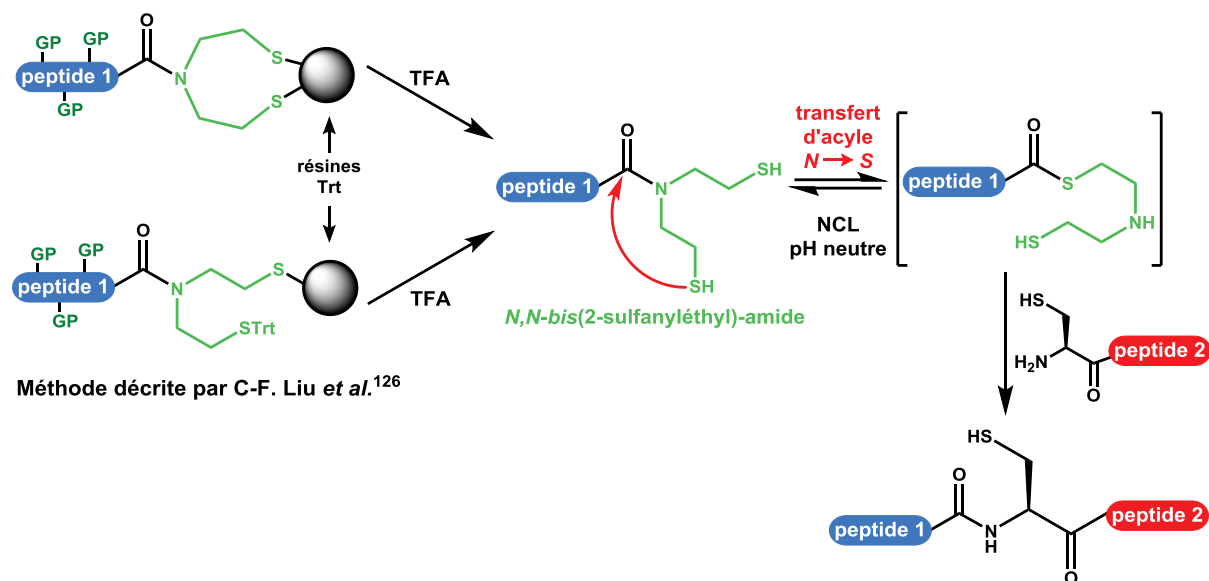
<sup>124</sup> Kawakami T.; Shimizu S.; Aimoto S., *Bull. Chem. Soc. Jpn*, **2010**, 83, 570-574.



▪ Peptides *N,N*-bis(2-sulfanyléthyl)amide (SEA)

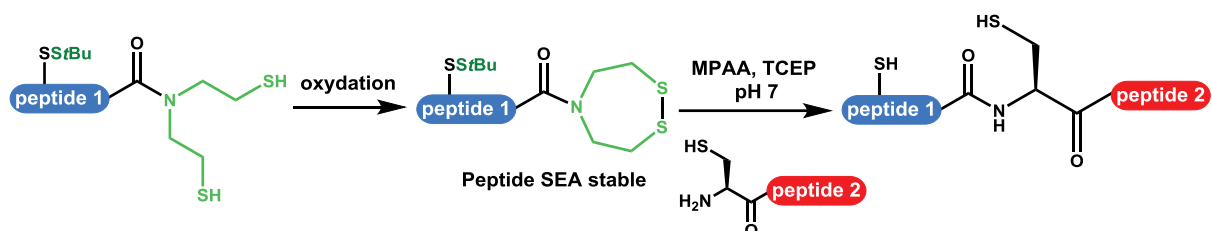
En 2010 et 2011, les équipes de Melnyk<sup>125</sup> puis de C-F. Liu<sup>126</sup> ont indépendamment décrit l'utilisation de peptides comportant un motif *N,N*-bis(2-sulfanyléthyl)-amide (SEA) C-terminal (figure 59). Dans la méthode développée par Liu, le motif est lié à la résine par un de ses groupements thiols; dans la méthode de Melnyk, le motif SEA est lié à la résine par ses deux groupements thiols (figure 59).

Méthode décrite par Melnyk *et al.*<sup>125</sup>



**Figure 59 :** Utilisation de peptides SEA<sup>125,126</sup>

Les peptides SEA obtenus après traitement TFA (figure 59) ne sont pas stables car pouvant être sujets à un transfert d'acyle *N*→*S* précoce dans les conditions de purification, ce qui peut compliquer leur manipulation. Melnyk et ses collaborateurs ont montré qu'il est possible de former le dérivé disulfure cyclique du motif SEA par une étape post-SPPS d'oxydation pour obtenir des peptides stables (figure 60). Les peptides SEA oxydés peuvent être également directement utilisés en NCL en présence de TCEP.

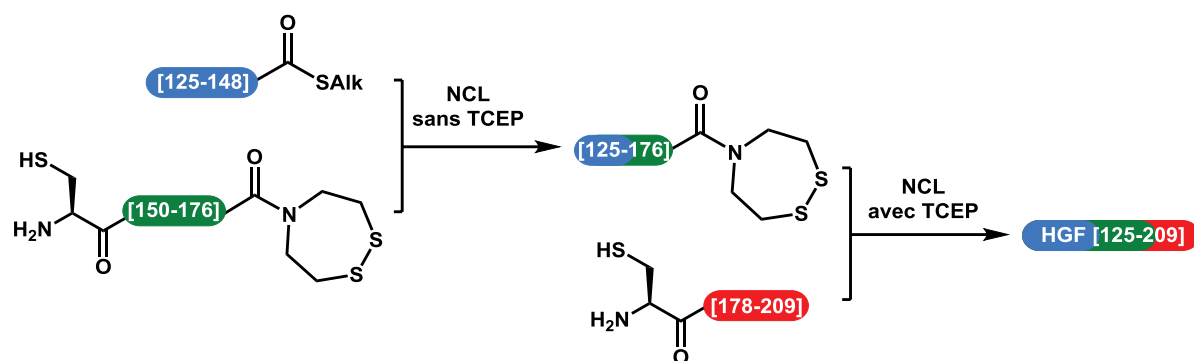


**Figure 60 :** Oxydation post-SPPS du motif C-terminal des peptides SEA<sup>125</sup>

<sup>125</sup> Ollivier N.; Dheur J.; Mhidia R.; Blanpain A.; Melnyk O., *Org. Lett.*, **2010**, *12*, 5238-5241.

<sup>126</sup> Hou W.; Zhang X.; Li F.; Liu C. F., *Org. Lett.*, **2011**, *13*, 386-389.

Melnyk et ses collaborateurs ont utilisé cette possibilité de masquage du dispositif SEA menant à des précurseurs désactivés de peptides thioester pour la synthèse de protéines par NCLs successives de N vers C.<sup>127</sup> Cette stratégie a été notamment appliquée à la synthèse en trois fragments du domaine K1 du facteur de croissance hépatocytaire (HGF [125-209], 85 acides aminés). La première ligation a lieu entre un peptide thioester préalablement formé et un cystéinyl-peptide comportant un motif SEA désactivé du côté C-terminal, dans des conditions strictement déoxygénées et en l'absence de TCEP; la seconde ligation a lieu entre le peptide obtenu par la première ligation et un cystéinyl-peptide dans des conditions de NCL plus classiques en présence de TCEP (figure 61).

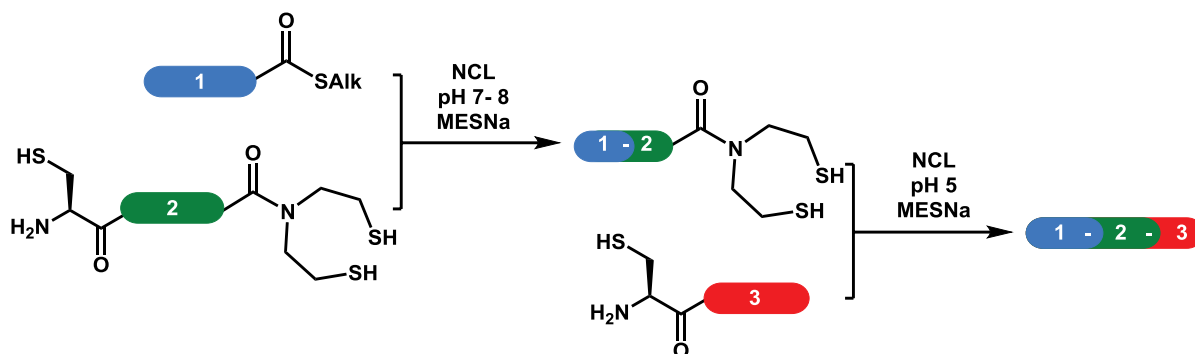


**Figure 61** : Synthèse du domaine K1 d'HGF par NCLs successives de N vers C avec un peptide SEA<sup>127</sup>

C-F. Liu et ses collaborateurs ont également utilisé les peptides SEA pour la synthèse de protéines par NCLs successives en exploitant cette fois-ci la différence de réactivité entre un peptide thioester d'alkyle préformé et un peptide SEA non masqué pour une ligation cinétiquement contrôlée (KCL, voir partie 1.3.4).<sup>128</sup> En effet, à pH neutre, le transfert d'acyle  $N \rightarrow S$  du motif SEA est très lent, la première ligation a donc lieu sélectivement entre le thioester d'alkyle et la cystéine C-terminale du fragment médian (fragment vert, figure 62); la deuxième ligation a lieu à un pH faiblement acide pour conduire au produit de ligation des trois fragments (figure 62).

<sup>127</sup> Ollivier N.; Vicogne J.; Vallin A.; Drobecq H.; Desmet R.; El Mahdi O.; Leclercq B.; Goormachtigh G.; Fajeur V.; Melnyk O., *Angew. Chem. Int. Ed.*, **2012**, *51*, 209-213.

<sup>128</sup> Yang R.; Hou W.; Zhang X.; Liu C. F., *Org. Lett.*, **2012**, *14*, 374-377.



**Figure 62 :** NCLs successives sous contrôle cinétique à partir de peptides SEA <sup>128</sup>

Melnyk et ses collaborateurs ont également décrit des conditions de ligations accélérées utilisant les peptides SEA pour la formation de sites de ligation contenant des acides aminés encombrés,<sup>129</sup> ainsi que des conditions pour la NCL à des jonctions Pro-Cys réputées comme étant difficile à former.<sup>130</sup>

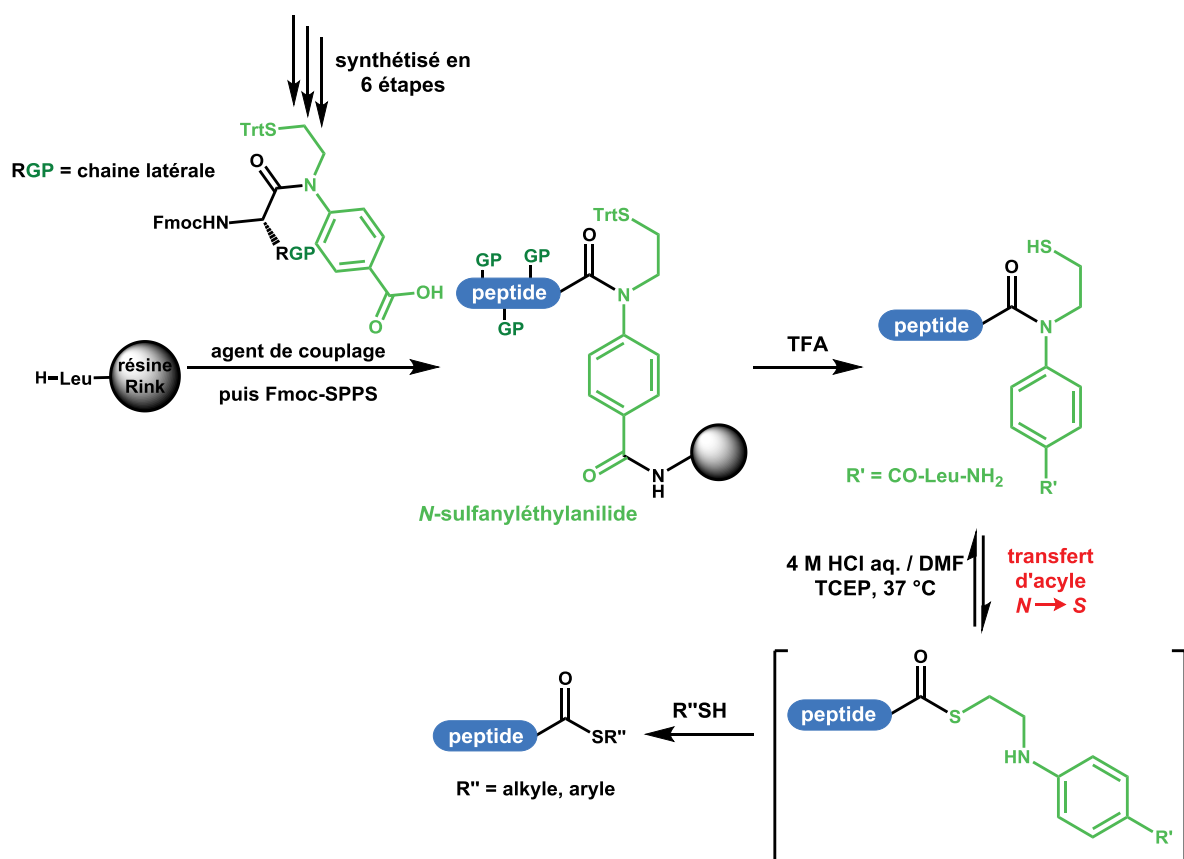
- Peptides *N*-sulfanyléthylanilide (SEAlide)

En 2009, l'équipe d'Otaka a décrit l'utilisation de peptides portant un groupement *N*-sulfanyléthylanilide C-terminal (SEAlide).<sup>131</sup> Le dispositif SEAlide *N*-acylé par le premier acide aminé est synthétisé en solution en 6 étapes puis greffé à la résine. Initialement, après traitement au TFA, la thioestérification par transfert d'acyle *N*→*S* était réalisée en milieu acide en présence de TCEP puis le peptide thioester était isolé après réaction avec un thiol (figure 63).

<sup>129</sup> Dheur J.; Ollivier N.; Melnyk O., *Org. Lett.*, **2011**, *13*, 1560-1563.

<sup>130</sup> Raibaut L.; Seeberger P.; Melnyk O., *Org. Lett.*, **2013**, *15*, 5516-5519.

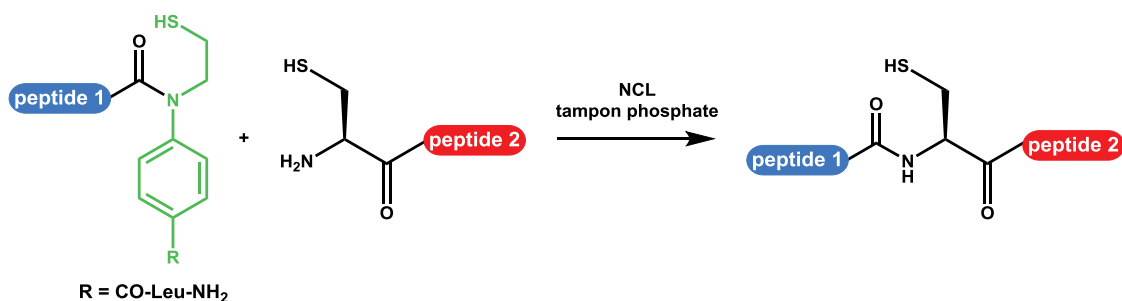
<sup>131</sup> Tsuda S.; Shigenaga A.; Bando K.; Otaka A., *Org. Lett.*, **2009**, *11*, 823-826.



**Figure 63 :** Synthèse de peptides thioester à partir d'un peptide SEALide<sup>131</sup>

Toutefois, la thioestérification en milieu acide conduit à l'épimérisation de l'acide aminé C-terminal lié au dispositif SEALide.

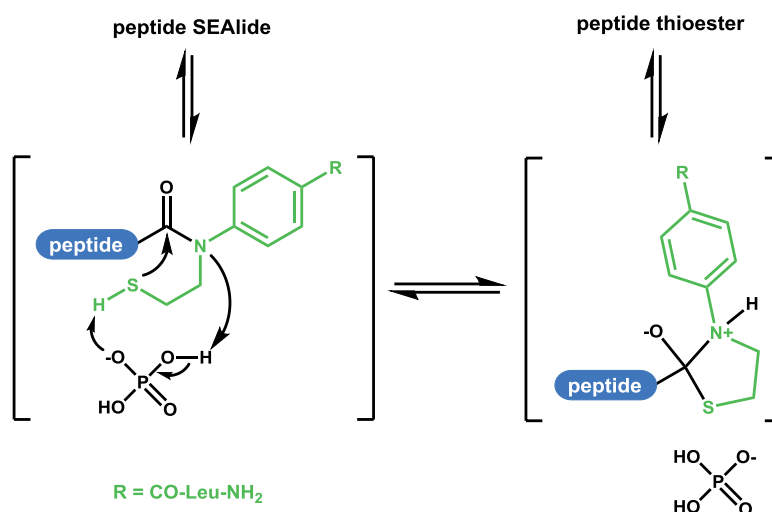
Le motif β-mercapto anilide est plus réactif qu'un β-mercapto-N-alkyl-amide simple grâce à l'activation électronique due à l'anilide. En 2011, Otaka et ses collaborateurs ont montré qu'il était possible d'utiliser directement un peptide SEALide dans une réaction de ligation chimique native sans épimérisation de l'acide aminé C-terminal du peptide précurseur de thioester (figure 64).<sup>132</sup>



**Figure 64 :** Utilisation directe de peptides SEALide en NCL<sup>132</sup>

<sup>132</sup> Sato K.; Shigenaga A.; Tsuji K.; Tsuda S.; Sumikawa Y.; Sakamoto K.; Otaka A., *ChemBioChem*, **2011**, *12*, 1840-1844.

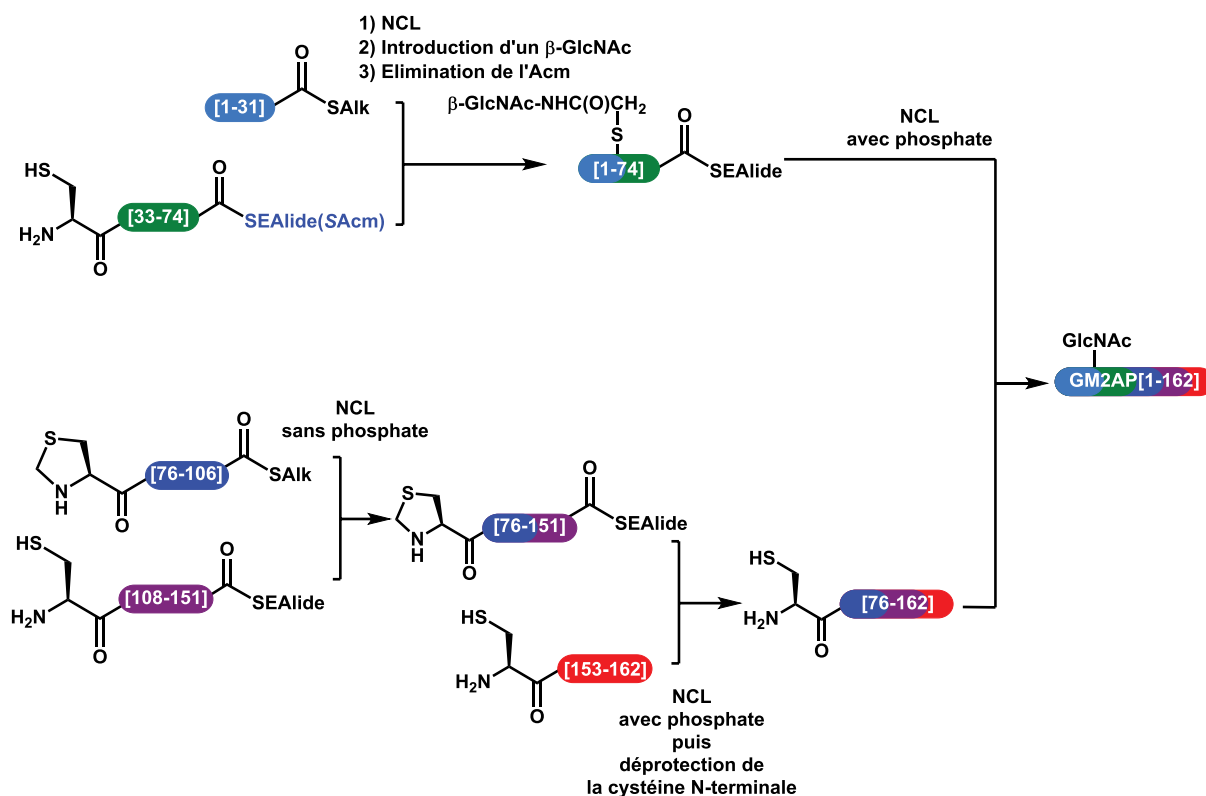
De façon intéressante, la réactivité des peptides SEALide peut être modulée par l'utilisation de solutions tampon de nature différente. En effet, en l'absence de sels de phosphate, la ligation à partir d'un peptide SEALide n'a pas lieu. L'hypothèse proposée par les auteurs est que les sels de phosphate pourraient assister le transfert d'acyle  $N \rightarrow S$  par une catalyse acido-basique dite pseudo-intramoléculaire car le tampon phosphate est présent en très grand excès dans le milieu réactionnel (figure 65).<sup>132</sup>



**Figure 65 :** Hypothèse de catalyse de la thioestérification de peptides SEALide par les sels de phosphate <sup>132</sup>

Otaka et ses collaborateurs ont utilisé cette propriété des peptides SEALide pour la synthèse de protéine par NCLs multiples convergentes sous contrôle cinétique (KCL). La synthèse convergente d'un analogue monoglycosylé de l'activateur protéique GM2 (GM2AP - 162 acides aminés) a notamment été décrite par NCLs successives de 5 fragments avec utilisation de peptides SEALide (figure 66).<sup>133</sup>

<sup>133</sup> Sato K.; Shigenaga A.; Kitakaze K.; Sakamoto K.; Tsuji D.; Itoh K.; Otaka A., *Angew. Chem. Int. Ed.*, **2013**, *52*, 7855-7859.



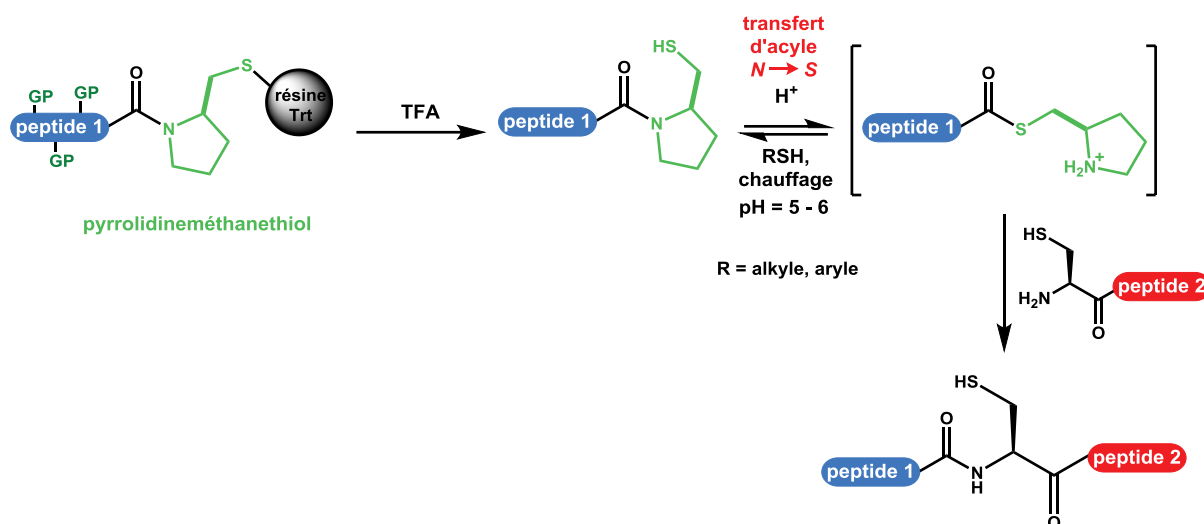
**Figure 66 :** Synthèse convergente d'un analogue de la GM2AP à partir de peptides SEAlide<sup>133</sup>

La synthèse complexe du dispositif SEAlide en solution représente la limitation majeure de cette stratégie.

- Peptides 2-pyrrolidinéméthanthiol (PMT)

L'équipe de C-F. Liu a décrit l'utilisation de peptides portant un motif 2-pyrrolidinéméthanthiol (PMT).<sup>134</sup> Le peptide est lié à la résine par le groupement thiol de ce motif (figure 67). La thioestérification du peptide obtenu après traitement TFA et la ligation sont réalisées « *one pot* » à un pH faiblement acide.

<sup>134</sup> Yang R.; Qi L.; Liu Y.; Ding Y.; Sheng Yi Kwek M.; Liu C. F., *Tetrahedron Lett.*, **2013**, 54, 3777-3780.



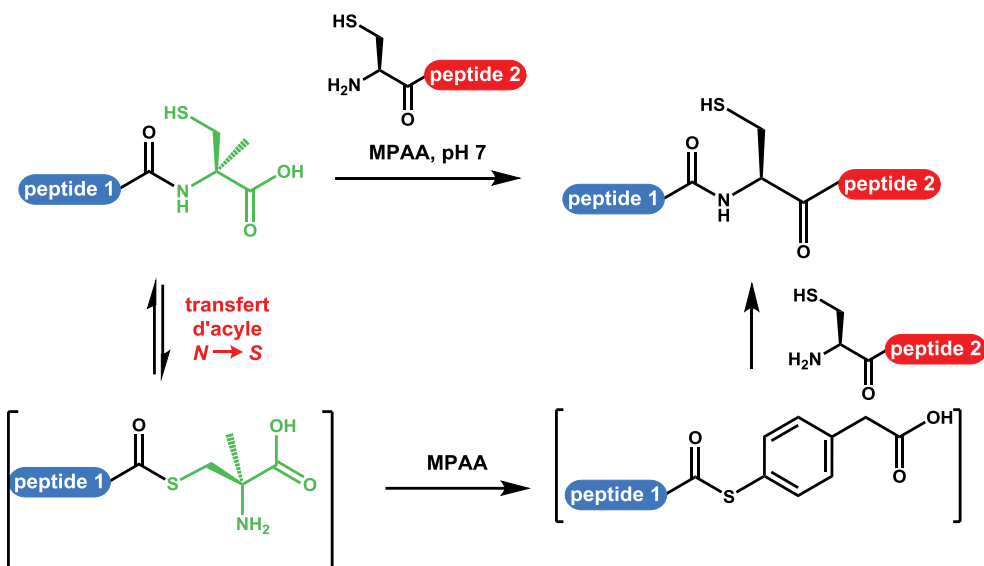
**Figure 67 :** Utilisation de peptides 2-pyrrolidinéméthanethiol <sup>134</sup>

Toutefois, les conditions utilisées sont sévères : chauffage à 42 °C pendant 7 h, pH légèrement acide non-optimal pour la NCL et des concentrations très importantes en peptides : de 15 à 35 mM. De telles concentrations sont atteignables pour de petits fragments peptidiques mais très difficiles à mettre en pratique pour des fragments plus longs utilisés pour la synthèse de petites protéines. Cette méthode a été peu utilisée.

- α-Méthyl-cystéine C-terminale

Offer et ses collaborateurs ont récemment décrit l'utilisation de peptides comportant une α-méthyl-cystéine acide C-terminale.<sup>135</sup> Dans les conditions de NCL, les peptides portant ce dérivé de cystéine terminal peuvent être transformés en peptides thioester par transfert d'acyle  $N \rightarrow S$  (figure 68). Une hypothèse plausible pour expliquer la formation facilitée de la fonction thioester repose sur une formation de l'état de transition cyclique à 5 chaînons lors du transfert d'acyle  $N \rightarrow S$  qui serait favorisée par l'effet Thorpe-Ingold dû à la présence des groupements *gem*-dialkyles de la cystéine C-terminale.

<sup>135</sup> Burlina F.; Papageorgiou G.; Morris C.; White P. D.; Offer J., *Chem. Sci.*, **2014**, 5, 766-770.

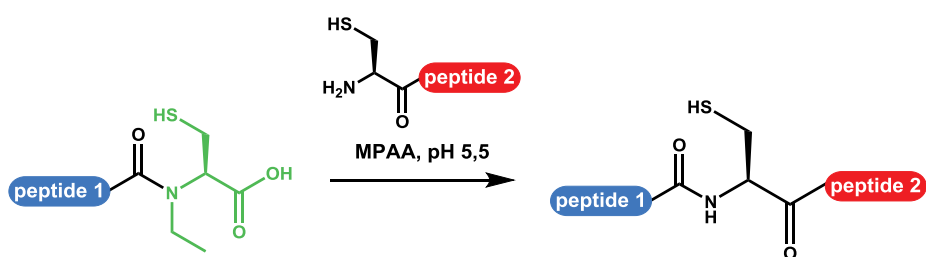


**Figure 68 :** Utilisation de peptides contenant une  $\alpha$ -méthyl-cystéine C-terminale <sup>135</sup>

Toutefois, les cinétiques observées pour des NCLs faisant intervenir ces peptides précurseurs sont lentes et ces réactions nécessitent l'emploi de très fortes concentrations de MPAA.

#### ▪ N-Alkyl-cystéine

En 2015, l'équipe d'Hojo a décrit l'utilisation directe de peptides comportant une N-éthyl-cystéine C-terminale dans une NCL sans avoir à préformer le thioester en milieu acide (figure 69).<sup>136</sup> L'utilisation d'une cystéine acide C-terminale conduit à une NCL significativement plus rapide que l'utilisation d'une cystéine amide C-terminale. Les ligations peuvent être réalisées en 12 h à 37 °C et à pH 5,5 mais les cinétiques observées lors d'une réaction à un pH proche de la neutralité optimal pour la NCL sont très lentes.



**Figure 69 :** Utilisation directe de peptides N-alkyl-cystéine en NCL <sup>136</sup>

#### ▪ Cystéine

En 2015, Macmillan et ses collaborateurs ont décrit l'utilisation directe de peptides de séquence native comportant une cystéine acide C-terminale mais seulement pour la cyclisation

<sup>136</sup> Asahuna Y.; Nabeshima K.; Hojo H., *Tetrahedron Lett.*, **2015**, 56, 1370-1373.



de petits peptides par NCL sans avoir à préformer le thioester en milieu acide.<sup>119</sup> La cyclisation de peptides par NCL intramoléculaire est connue pour être plus rapide que la ligation entre deux fragments en raison d'un phénomène d'augmentation de la concentration locale des fonctions réactives. Néanmoins, les cinétiques de ligation de cette stratégie sont lentes malgré les conditions sévères utilisées (pH 5,8 et chauffage à 60 °C pendant 48 h).

### 1.6 Bilan

Au cours des dernières décennies, la synthèse de peptides et protéines par voie chimique s'est imposée comme une alternative viable et une stratégie complémentaire à la production de protéines par voie recombinante. Deux avancées majeures ont contribué au succès de cette approche; la première repose sur l'introduction de la synthèse peptidique sur phase solide (SPPS) dans les années soixante, permettant un accès rapide à des peptides de tailles modérées (50 acides aminés en routine). La seconde avancée majeure correspond au développement de réactions de ligation entre des peptides totalement déprotégés en milieu aqueux conduisant à l'obtention de chaîne peptidique de grande taille. La ligation chimique native a révolutionné ce domaine au début des années 1990 mais ne s'est pas encore démocratisée réellement dans les laboratoires, notamment à cause de la difficulté de préparation du partenaire thioester par SPPS par la stratégie Fmoc très largement utilisée.

La synthèse par Fmoc-SPPS de peptides thioester, partenaires clés dans plusieurs stratégies de ligation, a fait l'objet du développement de nombreuses méthodes mais aucune de ces méthodes n'est encore utilisée universellement en raison d'inconvénients souvent communs aux différentes stratégies. La formation de la fonction thioester C-terminale après élongation reposant sur l'utilisation d'un dispositif de thioestérification par transfert d'acyle intramoléculaire  $N \rightarrow S$  semble prometteuse, néanmoins, les stratégies existantes sont souvent limitées par :

- la synthèse du dispositif de thioestérification qui est souvent complexe
- le couplage difficile de l'acide aminé C-terminal sur l'amine secondaire encombrée de ces dispositifs
- l'instabilité en milieu aqueux des peptides portant ces dispositifs, due à un transfert d'acyle  $N \rightarrow S$  précoce, pouvant nécessiter une étape post-SPPS
- la réactivité faible de ces dispositifs peut conduire à la nécessité d'une étape préalable de thioestérification en milieu acide, seules quelques stratégies ont décrit l'utilisation directe en NCL de peptides portant des dispositifs de thioestérification

## Chapitre 1

La haute technicité qui résulte des étapes nécessaires à la préparation et à l'utilisation de ces précurseurs de peptides thioester peut empêcher la mise en œuvre de ces stratégies par des personnes non-spécialistes.

C'est dans ce cadre que s'inscrivent les travaux décrits dans cette thèse. Nous avons cherché à développer une méthode de synthèse de précurseurs de peptides thioester entièrement automatisable à l'aide d'un synthétiseur de peptides à partir de réactifs commerciaux à bas coûts dont les peptides obtenus seraient directement utilisables dans une réaction de ligation chimique native sans transformations post-SPPS préalables nécessaires. Une méthode réunissant tous ces critères serait facilement mise en place par des personnes hors du domaine de la synthèse chimique et représenterait de ce fait une avancée significative pour la démocratisation de la synthèse de peptides et protéines par NCL.

## Partie 2 : Les peptides et protéines riches en ponts disulfure : des cibles idéales pour la NCL

Comme nous l'avons vu, la NCL est une réaction très puissante pour lier des fragments peptidiques dans le but d'atteindre une séquence de taille supérieure à celle possiblement atteinte par l'approche de SPPS classique. De plus, cette stratégie conduit à la formation d'une liaison amide avec une cystéine au point de jonction; elle est donc particulièrement adaptée à la synthèse de peptides contenant cet acide aminé. Les peptides riches en cystéines appariées par formation de ponts disulfure (DRP pour *disulfide-rich peptides*), des produits naturels de taille modérée (14 à une centaine d'acides aminés dont plus de 10 % de cystéine), représentent donc des cibles idéales pour la NCL.<sup>137</sup>

Ces peptides, conformationnellement contraints, possèdent un fort potentiel en tant qu'agents thérapeutiques ou diagnostiques; cette classe de peptides émerge comme une nouvelle classe de médicaments très prometteurs.<sup>138,139</sup> En effet, ces molécules, de taille intermédiaire à celle des molécules appartenant aux deux classes majeures actuelles de molécules-médicaments (les petites molécules et les *biologics*, les anticorps thérapeutique par exemple), combinent les avantages de ces deux classes : une biodisponibilité correcte, une faible immunogénéicité ainsi qu'une haute sélectivité et affinité; de plus, leur structure très compacte leur confère une stabilité importante face à la dégradation métabolique due aux protéases (figure 70).<sup>140</sup>



**Figure 70 :** Espace chimique des molécules thérapeutiques <sup>140</sup>

Les DRPs apparaissent dans tous les règnes du vivant, des bactéries aux animaux, plantes et champignons. La très grande biodiversité de ces molécules, illustrée par des millions de séquences distinctes (il a été estimé, par exemple, que seulement 0,01 % des toxines animales

<sup>137</sup> Okumura M.; Shimamoto S.; Hidaka Y., *FEBS Journal*, **2012**, 279, 2283-2295.

<sup>138</sup> Craik D. J.; Fairlie D. P.; Liras S.; Price D., *Chem. Biol. Drug. Des.*, **2013**, 81, 136-147.

<sup>139</sup> Terrett N., *Med. Chem. Commun.*, **2013**, 4, 474-475.

<sup>140</sup> Góngora-Benítez M.; Tulla-Puche J.; Albericio F., *Chem. Rev.*, **2014**, 114, 901-926.

riches en ponts disulfure sont répertoriées dans les bases de données), représente une source exceptionnelle de modèles pour le développement de médicaments basés sur les peptides.

Depuis le début des années 2000, une dizaine de peptides-médicaments contenant des ponts disulfure ont été approuvés pour commercialisation, traitant notamment des maladies comme le diabète, les douleurs chroniques ou les problèmes cardiaques.<sup>140</sup> De plus, une vingtaine de peptides de ce type sont actuellement sous essais cliniques ou précliniques pour des traitements divers contre, par exemple, le virus de l'immunodéficience humaine (VIH), des cancers, des scléroses ou en tant qu'antibactériens, antifongiques et antimicrobiens.<sup>141</sup> Par ailleurs, plusieurs DRPs ont été utilisés en tant que gabarit (ou *scaffold*) pour introduire des séquences peptidiques actives, formant ainsi des chimères stables, pour le développement de médicaments.<sup>142,143</sup>

Les DRPs ont été classés dans de très nombreux groupes et familles en fonction de leur structure imposée par les ponts disulfure (repliement ou *fold*) ou leurs similarités de séquence.<sup>144</sup> Cette partie a pour vocation de donner un simple aperçu, sans être exhaustif, des types de repliement généralement rencontrés ainsi que de leurs sources et leurs activités.

### ▪ Les peptides à nœud de cystine

Les DRPs à nœud de cystine (CK pour *cystine knot* ou *knottins*) sont rencontrés chez de nombreuses espèces, dont les plantes, les insectes, les champignons ou encore dans le venin d'araignées et de cônes marins.<sup>145</sup> Il a été estimé que ce repliement représente 40 % des DRPs naturels.<sup>144</sup> Il est caractérisé par une séquence de 30 à 50 acides aminés comprenant, au minimum, 3 ponts disulfure dont la connectivité caractéristique est C1-C4, C2-C5 et C3-C6; le pont disulfure C3-C6 passe à l'intérieur du macrocycle formé par les deux autres formant une architecture entrelacée de type rotaxane (figure 71).<sup>145,146</sup>

Le squelette peptidique de certains peptides CK est cyclique (figure 71); ces peptides à nœud de cystine cyclique (CCK pour *cyclic cystine knot*), également appelés « cyclotides », constituent la famille principale de peptides cycliques rencontrés chez les plantes. Ils sont constitués d'une trentaine d'acides aminés et démontrent généralement une forte stabilité vis-à-

---

<sup>141</sup> King G. F., *Expert Opin. Biol. Ther.*, **2011**, *11*, 1469-1484.

<sup>142</sup> Henriques S. T.; Craik D. J., *Drug Discovery Today*, **2010**, *15*, 57-64.

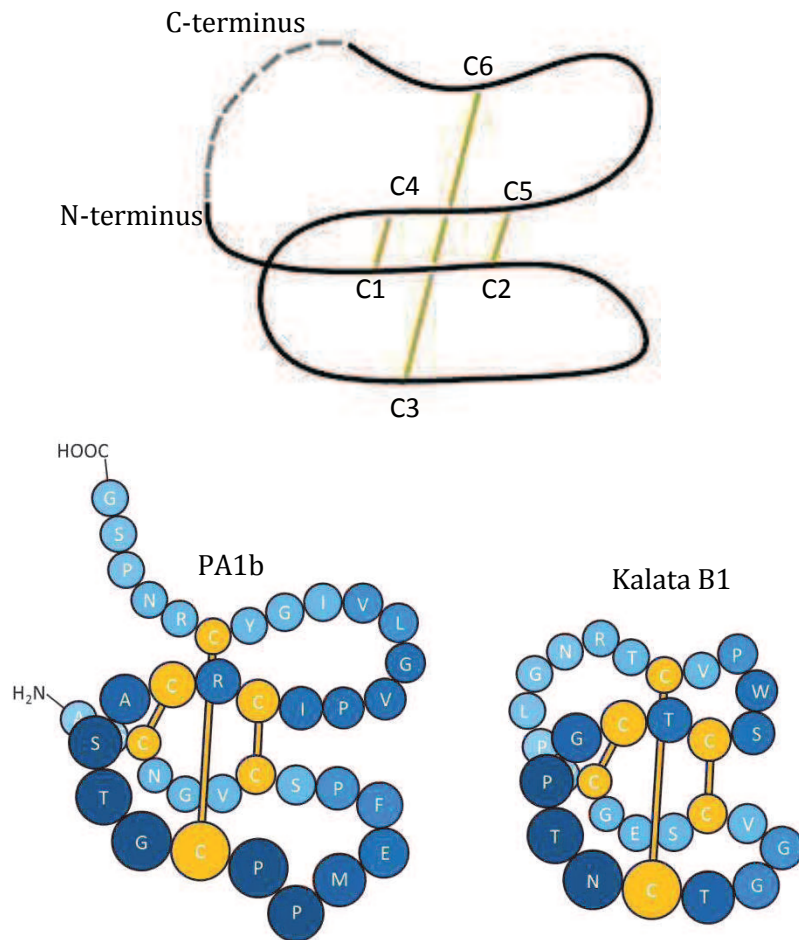
<sup>143</sup> Moore S. J.; Cochran J. R., *Methods Enzymol.*, **2012**, *503*, 223-251.

<sup>144</sup> Cheek S.; Krishna S. S.; Grishin N. V., *J. Mol. Biol.*, **2006**, *359*, 215-237.

<sup>145</sup> Craik D. J.; Daly N. L.; Wayne C., *Toxicon*, **2001**, *39*, 43-60.

<sup>146</sup> Ackerman S. E.; Currier N. V.; Bergen J. M.; Cochran J. R., *Expert Rev. Proteomics*, **2014**, *11*, 561-572.

vis de la dégradation enzymatique et une bonne biodisponibilité grâce à leur squelette cyclique.<sup>147</sup>



**Figure 71 :** Représentation schématique d'un peptide CK (schéma extrait de la ref 146) et exemples d'un peptide CK issu du pois (PA1b)<sup>148</sup> et d'un cyclotide (Kalata B1)<sup>149</sup>

Les peptides CK couvrent une large gamme d'activités biologiques dont des actions analgésiques, antiparasitaires, insecticides ou encore antimicrobiennes.<sup>150</sup>

<sup>147</sup> Craik D. J.; Daly N. L.; Bond T.; Waine C., *J. Mol. Biol.*, **1999**, 294, 1327-1336.

<sup>148</sup> Gressent F.; Da Silva P.; Eyraud V.; Karaki L.; Royer C., *Toxins*, **2011**, 3, 1502-1517.

<sup>149</sup> Daly N. L.; Love S.; Alewood P. F.; Craik D. J., *Biochemistry*, **1999**, 38, 10606-10614.

<sup>150</sup> The KNOTTIN database, disponible en ligne à : <http://knottin.cbs.cnrs.fr>

### ▪ Les défensines

Les défensines constituent une autre famille de DRPs naturels largement étudiée. Ces petits peptides (une cinquantaine d'acides aminés) cationiques et dotés de propriétés antimicrobiennes, sont majoritairement rencontrés chez les plantes,<sup>151</sup> les mammifères<sup>152</sup> et les insectes.<sup>153</sup>

Les défensines de plantes sont stabilisées par 4 ponts disulfure alors que les défensines de mammifères n'en contiennent que 3. Les défensines de plantes présentent une structure tertiaire hautement conservée constituée d'hélices  $\alpha$  et de feuillets  $\beta$  (motif  $CS\alpha\beta$ ).<sup>154</sup>

Chez les mammifères, les défensines sont divisées en trois classes, suivant la connectivité de leurs ponts disulfure : les  $\alpha$ -,  $\beta$ - et  $\theta$ -défensines, les deux premières familles étant observées chez l'humain.<sup>152</sup> La connectivité des ponts disulfure des  $\alpha$ -défensines est C1-C6, C2-C4 et C3-C5 alors que la connectivité des ponts dans les  $\beta$ -défensines est C1-C5, C2-C4 et C3-C6. Les  $\theta$ -défensines sont de petits peptides (18 acides aminés) rencontrés chez certains primates et caractérisés par un squelette cyclique et 3 ponts disulfure formant une structure en forme d'échelle.<sup>155</sup>

Les défensines d'insectes sont constituées de 40 à 65 acides aminés et sont contraintes par 3 à 4 ponts disulfure, adoptant un repliement de type  $CS\alpha\beta$ .<sup>153</sup>

En complément de leurs propriétés antimicrobiennes, les défensines possèdent une gamme étendue d'activités biologiques connexes : des propriétés insecticides, d'inhibition de synthèse de protéines, d'inhibition de digestion enzymatique ou encore de tolérance au zinc.<sup>151,153,154</sup>

### ▪ Les DRPs issus de venins

De nombreux peptides riches en ponts disulfure (d'une taille de 10 à 60 acides aminés environ) sont également rencontrés dans le venin d'animaux, comme les serpents, araignées, scorpions ou les cônes marins, qui peuvent contenir plusieurs centaines de ces peptides. Ces molécules, jouant un rôle de défense voire de prédation dans la nature, possèdent un potentiel thérapeutique très importants.<sup>141,156</sup> Leur très grande diversité est illustrée par un nombre

---

<sup>151</sup> Thomma B. P.; Cammue B. P.; Thevissen K., *Planta*, **2002**, 216, 196-202.

<sup>152</sup> Yang D.; Biragyn A.; Kwak L. W.; Oppenheim J. J., *Trends Immunol.*, **2002**, 23, 291-296.

<sup>153</sup> Bulet P.; Stocklin R., *Protein Pept. Lett.*, **2005**, 12, 3-11.

<sup>154</sup> Lacerda A. F.; Vasconcelos E. A. R.; Barbosa Pelegrini P.; Grossi de Sa M. F., *Front. Microbiol.*, **2014**, 5, 1-10.

<sup>155</sup> Conibear A. C.; Craik D. J., *Angew. Chem. Int. Ed.*, **2014**, 53, 10612-10623.

<sup>156</sup> Lewis R. J.; Garcia M. L., *Nat. Rev. Drug Discovery*, **2003**, 2, 790-802.

important de séquences distinctes et de connectivités différentes des ponts disulfure, en découlent de très nombreux repliements caractérisés par leur composition en hélices  $\alpha$  et/ou feuillets  $\beta$ .<sup>157,158</sup> A titre d'exemple, les toxines à trois doigts (3FT pour *three-finger toxins*), sont retrouvées chez toutes les familles de serpents; elles sont constituées de trois boucles en feuillet  $\beta$  partant d'un cœur, hautement conservé, composé de 4 ponts disulfure.<sup>159</sup>

La grande diversité de ces DRPs issus de venins résulte en une large variété de fonctions biologiques dont, par exemple, une action sur les canaux ioniques ou encore des actions de régulation de l'activité de différents récepteurs dont certains récepteurs couplés aux protéines G (GPCR).<sup>156,157,160</sup>

Il est à noter que certains DRPs naturels peuvent également approcher la taille d'une centaine d'acides aminés et sont parfois composés de deux domaines intervenant chacun dans des fonctions biologiques différentes dont la synergie reste encore à explorer.<sup>161-163</sup>

Les propriétés pharmacologiques intéressantes ont stimulé la recherche sur la production de DRPs qui doivent être facilement synthétisables, et ce, à large échelle et à moindre coût, pour explorer leurs applications thérapeutiques. La synthèse de DRP par voie recombinante à large échelle implique des coûts de production très importants.<sup>164</sup> Les développements de la synthèse peptidique par voie chimique ont permis de diminuer les coûts de production des peptides par cette voie et cette approche est privilégiée pour la synthèse à large échelle.<sup>164</sup> Néanmoins, si la séquence peptidique de DRPs de moins de 50 acides aminés peut être assemblée par SPPS classique, les séquences dépassant les 50 acides aminés nécessitent l'utilisation de méthodes de ligation chimique. La SPPS en stratégie Boc a récemment été appliquée pour la préparation du partenaire thioester nécessaire à la NCL pour la synthèse de DRPs.<sup>165,166</sup> Néanmoins, comme cela a été souligné dans la partie 1 de cette étude bibliographique, il n'existe pas pour l'instant de méthode totalement générale, robuste, simple à

<sup>157</sup> Lavergne V.; Taft R. J.; Alewood P. F., *Curr. Top. Med. Chem.*, **2012**, *12*, 1514-1533.

<sup>158</sup> Mouhat S.; Jouirou B.; Mosbah A.; De Waard M.; Sabatier J-M., *Biochem. J.*, **2004**, *378*, 717-726.

<sup>159</sup> Manjunatha Kini R.; Doley R., *Toxicon*, **2010**, *56*, 855-867.

<sup>160</sup> Rouget C.; Quinton L.; Maïga A.; Gales C.; Masuyer G.; Malosse C.; Chamot-Rooke J.; Thai R.; Mourier G.; de Pauw E.; Gilles N.; Servent D., *Br. J. Pharmacol.*, **2010**, *161*, 1361-1374.

<sup>161</sup> Saito T.; Kawabata S-I.; Shigenaga T.; Takayenoki Y.; Cho J.; Nakajima H.; Hirata M.; Iwanaga S., *J. Biochem.*, **1995**, *117*, 1131-1137.

<sup>162</sup> Hervé-Grépinet V.; Réhault-Godbert S.; Labas V.; Magallon T.; Derache C.; Lavergne M.; Gautron J.; Lalmanach A-C.; Nys Y., *Antimicrob. Agents Chemother.*, **2010**, *54*, 4401-4408.

<sup>163</sup> Bae C.; Kalia J.; Song I.; Yu J.; Kim H. H.; Swartz K. J.; Kim J. I., *PLoS One*, **2012**, *7*, e51516.

<sup>164</sup> Vlieghe P.; Lisowski V.; Martinez J.; Khrestchatisky M., *Drug Discovery Today*, **2010**, *15*, 40-56.

<sup>165</sup> Medini K.; Harris P. W. R.; Hards K.; Dingley A. J.; Cook G. M.; Brimble M. A., *ChemBioChem*, **2015**, *16*, 328-336.

<sup>166</sup> Tsuda S.; Yoshiya T.; Mochizuki M.; Nishiuchi Y., *Org. Lett.*, **2015**, *17*, 1806-1809.

## Chapitre 1

mettre en place et peu coûteuse pour la synthèse de peptides thioester par Fmoc-SPPS qui est plus utilisée dans les laboratoires. Comme cela a déjà été dit, le développement d'une telle méthode représente encore un challenge pour les chimistes des peptides.

Pour étudier le potentiel de la nouvelle méthode de synthèse de peptides cryptothioester développée dans cette thèse, nous avons choisi d'appliquer cette approche à la synthèse par NCL de la forme réduite de plusieurs DRPs représentatifs de ces produits naturels :

- de deux longs DRPs linéaires, un dont la synthèse est déjà connue, constituant un modèle pour notre approche, l'autre étant un cas réel dont la synthèse et le repliement oxydatif n'étaient pas connus au début de mes travaux de thèse
- des cas particuliers de DRPs linéaires présentant une cystéine C-terminale liée à des réactions secondaires lors de leur synthèse par les méthodes classiques
- plusieurs DRPs cycliques représentant des cas idéaux pour étudier la portée de notre stratégie.



## Références bibliographiques

### Partie 1

1. International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature*, **2001**, *409*, 860-921.
2. International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature*, **2004**, *431*, 931-945.
3. Unwin R. D.; Gaskell S. J.; Evans C. A.; Whetton A. D., The potential for proteomic definition of stem cell populations. *Exp. Hematol.*, **2003**, *31*, 1147-1159.
4. Dumas A.; Lercher L.; Spicer C. D.; Davis B. G., Designing logical codon reassignment – Expanding the chemistry in biology. *Chem. Sci.*, **2015**, *6*, 50-69.
5. Nilsson B. L.; Soellner M. B.; Raines R. T., Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **2005**, *34*, 91-118.
6. Fischer E.; Fourneau E., Ueber einige derivate des glykocolls. *Ber. Deutsch. Chem. Ges.*, **1901**, *34*, 2868-2877.
7. Bergmann M.; Zervas L., Über ein allgemeines verfahren der peptid-synthese. *Ber. Deutsch. Chem. Ges.*, **1932**, *65*, 1192-1201.
8. Du Vigneaud V.; Ressler C.; Swan J. M.; Roberts C. W.; Katsoyannis P. G.; Gordon S., The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.*, **1953**, *75*, 4879-4880.
9. El-Faham A.; Albericio F., Peptide coupling reagents, more than a letter soup. *Chem. Rev.*, **2011**, *111*, 6557–6602.
10. Sakakibara S., Chemical synthesis of proteins in solution. *Biopolymers*, **1999**, *51*, 279-296.
11. Kent S. B. H., Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.*, **1988**, *57*, 957-989.
12. Merrifield R. B., Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.*, **1963**, *85*, 2149-2154.
13. Merrifield R. B.; Stewart J. M.; Jernberg N., Instrument for automated synthesis of peptides. *Anal. Chem.*, **1966**, *38*, 1905-1914.
14. Sheppard R., The fluorenylmethoxycarbonyl group in solid phase synthesis. *J. Pept. Sci.*, **2003**, *9*, 545-552.
15. Garcia Martin F.; Albericio F., Solid supports for the synthesis of peptides. From the first resin used to the most sophisticated in the market. *Chimica Oggi*, **2009**, *26*, 29-34.

16. Góngora-Benítez M.; Tulla-Puche J.; Albericio F., Handles for Fmoc solid-phase synthesis of protected peptides. *ACS Comb. Sci.*, **2013**, *15*, 217–228.
17. Isidro-Llobet A.; Alvarez M.; Albericio F., Amino acid-protecting groups. *Chem. Rev.*, **2009**, *109*, 2455-2504.
18. Albericio F., Developments in peptide and amide synthesis. *Curr. Opin. Chem. Biol.*, **2004**, *8*, 211-221.
19. Nishiuchi Y.; Inui T.; Nishio H.; Bodi J.; Kimura T.; Tsuji F. I.; Sakakibara S., Chemical synthesis of the precursor molecule of the *Aequorea* green fluorescent protein, subsequent folding, and development of fluorescence. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 13549-13554.
20. Blake J.; Li C. H., New segment-coupling method for peptide synthesis in aqueous solution: application to synthesis of human [Gly<sup>17</sup>]- $\beta$ -endorphin. *Proc. Natl. Acad. Sci. USA*, **1981**, *78*, 4055-4058.
21. Hojo H.; Aimoto S., Polypeptide synthesis using the *S*-Alkyl thioester of a partially protected peptide segment. synthesis of the DNA-binding domain of *c*-Myb Protein (142–193)-NH<sub>2</sub>. *Bull. Chem. Soc. Jpn*, **1991**, *64*, 111-117.
22. Sato T.; Aimoto S., Use of thiosulfonate for the protection of thiol groups in peptide ligation by the thioester method. *Tetrahedron Lett.*, **2003**, *44*, 8085-8087.
23. Aimoto S., Polypeptide synthesis by the thioester method. *Biopolymers*, **1999**, *51*, 247-265.
24. Roberts A. G.; Johnston E. V.; Shieh J. H.; Sondey J. P.; Hendrickson R. C.; Moore M. A. S.; Danishefsky S. J., Fully synthetic granulocyte colony-stimulating factor enabled by isonitrile-mediated coupling of large, side-chain-unprotected peptides. *J. Am. Chem. Soc.*, **2015**, *137*, 13167-13175.
25. Schnolzer M.; Kent S. B. H., Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science*, **1992**, *256*, 221-225.
26. 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*, 1821-1829.
27. 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*, 1829-1838.
28. Dawson P. E.; Muir T. W.; Clark-Lewis I.; Kent S. B. H., Synthesis of proteins by native chemical ligation. *Science*, **1994**, *266*, 776-779.
29. 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. *Liebigs Ann. Chem.*, **1953**, *583*, 129-149.

30. Tam J. P.; Lu Y. A.; Liu C. F.; Shao J., Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. Usa*, **1995**, *92*, 12485-12489.
31. Villain M.; Gaertner H.; Botti P., Native chemical ligation with aspartic and glutamic acids as C-terminal residues: scope and limitations. *Eur. J. Org. Chem.*, **2003**, 3267-3272.
32. Hackeng T. M.; Griffin J. H.; Dawson P. E., Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. Usa*, **1999**, *96*, 10068-10073.
33. Pollock S. B.; Kent S. B. H., An investigation into the origin of the dramatically reduced reactivity of peptide-prolyl-thioesters in native chemical ligation. *Chem. Commun.*, **2011**, *47*, 2342-2344.
34. Raibaut L.; Ollivier N.; Melnyk O., Sequential native peptide ligation strategies for total chemical protein synthesis. *Chem. Soc. Rev.*, **2012**, *41*, 7001-7015.
35. Muir T. W.; Dawson P. E.; Kent S. B. H., Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.*, **1997**, *289*, 266-298.
36. Ueda S.; Fujita M.; Tamamura H.; Fujii N.; Otaka A., Photolabile protection for one-pot sequential native chemical ligation. *ChemBioChem*, **2005**, *6*, 1983-1986.
37. Bang D.; Chopra N.; Kent S. B. H., Total chemical synthesis of crambin. *J. Am. Chem. Soc.*, **2004**, *126*, 1377-1383.
38. Villain M.; Vizzanova J.; Rose K., Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chem. Biol.*, **2001**, *8*, 673-679.
39. Bang D.; Kent S. B. H., A one-pot total synthesis of crambin. *Angew. Chem. Int. Ed.*, **2004**, *43*, 2534-2538.
40. Dawson P. E.; Churchill M. J.; Ghadiri M. R.; Kent S. B. H., Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.*, **1997**, *119*, 4325-4329.
41. Johnson E. C. B.; Kent S. B. H., Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.*, **2006**, *128*, 6640-6646.
42. 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*, 3985-3988.
43. 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*, 11679-11684.
44. 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*, 8720-8727.
45. Brik A.; Keinan E.; Dawson P. E., Protein synthesis by solid-phase chemical ligation using a safety catch linker. *J. Org. Chem.*, **2000**, *65*, 3829-3835.

46. 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*, 253-261.
47. Bang D.; Kent S. B. H., His<sub>6</sub> tag-assisted chemical protein synthesis. *Proc. Natl. Acad. Sci. Usa*, **2005**, *102*, 5014-5019.
48. 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*, 4061-4066.
49. 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*, 12534-12537.
50. Trinquier G.; Sanejouand Y-H., Which effective property of amino acids is best preserved by the genetic code? *Protein Eng.*, **1998**, *11*, 153-159.
51. Yan L. Z.; Dawson P. E., Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.*, **2001**, *123*, 526-533.
52. Ma J.; Zeng J.; Wan Q., Postligation-desulfurization: a general approach for chemical protein synthesis. *Top. Curr. Chem.*, **2015**, *363*, 57-101.
53. Malins L. R.; Payne R. J., Modern extensions of native chemical ligation for chemical protein synthesis. *Top. Curr. Chem.*, **2015**, *362*, 27-87.
54. 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.*, **2007**, *46*, 9248-9252.
55. Burns J. A.; Butler J. C.; Moran J.; Whitesides G. M., Selective reduction of disulfides by tris(2-carboxyethyl)phosphine. *J. Org. Chem.*, **1991**, *56*, 2648-2650.
56. Offer J., Native chemical ligation with N<sup>α</sup> acyl transfer auxiliaries. *Biopolymers*, **2010**, *94*, 530-541.
57. Loibl S. F.; Harpaz Z.; Seitz O., A type of auxiliary for native chemical peptide ligation beyond cysteine and glycine junctions. *Angew. Chem. Int. Ed.*, **2015**, DOI: 10.1002/anie.201505274.
58. Brik A.; Yang Y. Y.; Ficht S.; Wong C. H., Sugar-assisted glycopeptide ligation. *J. Am. Chem. Soc.*, **2006**, *128*, 5626-5627.
59. Ficht S.; Payne R. J.; Brik A.; Wong C. H., Second-generation sugar-assisted ligation: a method for the synthesis of cysteine-containing glycopeptides. *Angew. Chem. Int. Ed.*, **2007**, *46*, 5975-5979.
60. Hojo H.; Ozawa C.; Katayama H.; Ueki A.; Nakahara Y.; Nakahara Y., The mercaptomethyl group facilitates an efficient one-pot ligation at Xaa-Ser/Thr for (glyco)peptide synthesis. *Angew. Chem. Int. Ed.*, **2010**, *49*, 5318-5321.
61. Zhang L.; Tam J. P., Orthogonal coupling of unprotected peptide segments through histidyl amino terminus. *Tetrahedron Lett.*, **1997**, *38*, 3-6.

62. Tam J. P.; Yu Q., Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones. *Biopolymers*, **1998**, *46*, 319-327.
63. Gieselman M. D.; Xie L.; Van Der Donk W. A., Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.*, **2001**, *3*, 1331-1334.
64. Malins L. R.; Mitchell N. J.; Payne R. J., Peptide ligation chemistry at selenol amino acids. *J. Pept. Sci.*, **2014**, *20*, 64-77.
65. Metanis N.; Keinan E.; Dawson P. E., Traceless ligation of cysteine peptides using selective deselenization. *Angew. Chem. Int. Ed.*, **2010**, *49*, 7049-7053.
66. Mitchell N. J.; Malins L. R.; Liu X.; Thompson R. E.; Chan B.; Radom L.; Payne R. J., Rapid additive-free selenocystine-selenoester peptide ligation. *J. Am. Chem. Soc.*, **2015**, DOI: 10.1021/jacs.5b07237.
67. Thomas G. L.; Payne R. J., Phosphate-assisted peptide ligation. *Chem. Commun.*, **2009**, *45*, 4260-4262.
68. Saxon E.; Bertozzi C., Cell surface engineering by a modified Staudinger reaction. *Science*, **2000**, *287*, 2007-2010.
69. Nilsson B. L.; Kiessling L. L.; Raines R. T., Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.*, **2000**, *2*, 1939-1941.
70. Soellner M. B.; Nilsson B. L.; Raines R. T., Reaction mechanism and kinetics of the traceless Staudinger ligation. *J. Am. Chem. Soc.*, **2006**, *128*, 8820-8828.
71. Mhidia R.; Boll E.; Fécourt F.; Ermolenko M.; Ollivier N.; Sasaki K.; Crich D.; Delpech B.; Melnyk O., Exploration of an imide capture/*N,N*-acyl shift sequence for asparagine native peptide bond formation. *Bioorg. Med. Chem.*, **2013**, *21*, 3479-3485.
72. Bode J. W.; Fox R. M.; Baucom K. D., Chemoselective amide ligations by decarboxylative condensations of *N*-alkylhydroxylamines and  $\alpha$ -ketoacids. *Angew. Chem. Int. Ed.*, **2006**, *45*, 1248-1252.
73. Pusterla I.; Bode J. W., The mechanism of the  $\alpha$ -ketoacid-hydroxylamine amide-forming ligation. *Angew. Chem. Int. Ed.*, **2012**, *51*, 513-516.
74. Pattabiraman V. R.; Ogunkoya A. O.; Bode J. W., Chemical protein synthesis by chemoselective  $\alpha$ -ketoacid-hydroxylamine (KAHA) ligations with 5-oxaproline. *Angew. Chem. Int. Ed.*, **2012**, *51*, 5114-5118.
75. Wucherpfennig T. G.; Rohrbacher F.; Pattabiraman V. R.; Bode J. W., Formation and rearrangement of homoserine depsipeptides and depsiproteins in the  $\alpha$ -ketoacid-hydroxylamine ligation with 5-oxaproline. *Angew. Chem. Int. Ed.*, **2014**, *53*, 1-5.
76. Pusterla I.; Bode J. W., An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations. *Nat. Chem.*, **2015**, *7*, 668-672.

77. Liu C. F.; Tam J. P., Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. Usa*, **1994**, *91*, 6584-6588.
78. Li X.; Lam H. Y.; Zhang Y.; Chan C. K., Salicylaldehyde ester-induced chemoselective peptide ligations: enabling generation of natural peptidic linkages at the serine/threonine sites. *Org. Lett.*, **2010**, *12*, 1724-1727.
79. Bu X. Z.; Xie G. Y.; Law C. W.; Guo Z. H., An improved deblocking agent for direct Fmoc solid-phase synthesis of peptide thioesters. *Tetrahedron Lett.*, **2002**, *43*, 2419-2422.
80. Futaki S.; Sogawa K.; Maruyama J.; Asahara T.; Niwa M.; Hojo H., Preparation of peptide thioesters using Fmoc-solid-phase peptide synthesis and its application to the construction of a template-assembled synthetic protein (TASP). *Tetrahedron Lett.*, **1997**, *38*, 6237-6240.
81. Yamamoto N.; Tanabe Y.; Okamoto R.; Dawson P. E.; Kajihara Y., Chemical synthesis of a glycoprotein having an intact human complex-type sialyloligosaccharide under the Boc and Fmoc synthetic strategies. *J. Am. Chem. Soc.*, **2008**, *130*, 501-510.
82. Nakamura K.; Hanai N.; Kanno M.; Kobayashi A.; Ohnishi Y.; Ito Y.; Nakahara Y., Design and synthesis of silyl ether-based linker for solid-phase synthesis of glycopeptides. *Tetrahedron Lett.*, **1999**, *40*, 515-518.
83. Li L.; Wang P., Transfer allyl esters to thioesters in solid phase condition: synthesis of peptide thioesters by Fmoc chemistry. *Tetrahedron Lett.*, **2007**, *48*, 29-32.
84. Lelièvre D.; Barta P.; Aucagne V.; Delmas A. F., Preparation of peptide thioesters using Fmoc strategy through hydroxyl side chain anchoring. *Tetrahedron Lett.*, **2008**, *49*, 4016-4019.
85. Ficht S.; Payne R. J.; Guy R. T.; Wong C. H., Solid-phase synthesis of peptide and glycopeptide thioesters through side-chain-anchoring strategies. *Chem. Eur. J.*, **2008**, *14*, 3620-3629.
86. Alsina J.; Yokum T. S.; Albericio F.; Barany G., Backbone amide linker (BAL) strategy for  $N^{\alpha}$ -9-fluorenylmethoxycarbonyl (Fmoc) solid-phase synthesis of unprotected peptide *p* nitroanilides and thioesters. *J. Org. Chem.*, **1999**, *64*, 8761-8769.
87. Gross C. M.; Lelièvre D.; Woodward C. K.; Barany G., Preparation of protected peptidyl thioester intermediates for native chemical ligation by  $N^{\alpha}$ -9-fluorenylmethoxycarbonyl (Fmoc) chemistry: considerations of side-chain and backbone anchoring strategies, and compatible protection for N-terminal cysteine. *J. Peptide Res.*, **2005**, *65*, 395-410.
88. Kenner G. W.; McDermott J. R.; Sheppard R. C., The safety catch principle in solid phase peptide synthesis. *J. Chem. Soc. D*, **1971**, 636-637.
89. Backes B. J.; Ellman J. A., An alkanesulfonamide "safety-catch" linker for solid-phase synthesis. *J. Org. Chem.*, **1999**, *64*, 2322-2330.
90. 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*, 11369-11374.

91. Shin Y.; Winans K. A.; Backes B. J.; Kent S. B. H.; Ellman J. A.; Bertozzi C. R., Fmoc-based synthesis of peptide- $\alpha$ thioesters: application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.*, **1999**, *121*, 11684-11689.
92. Camarero J. A.; Hackel B. J.; de Yoreo J. J.; Mitchell A. R., Fmoc-based synthesis of peptide  $\alpha$ -thioesters using an aryl hydrazine support. *J. Org. Chem.*, **2004**, *69*, 4145-4151.
93. 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*, 7645-7649.
94. Fang G-M.; Wang J-X.; Liu L., Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angew. Chem. Int. Ed.*, **2012**, *51*, 10347-10350.
95. Mong S. K.; Vinogradov A. A.; Simon M. D.; Pentelute B. L., Rapid total synthesis of DARPin pE59 and Barnase. *ChemBioChem*, **2014**, *15*, 721 – 733.
96. Reif A.; Siebenhaar S.; Tröster A.; Schmälzlein M.; Lechener C.; Velisetty P.; Gottwald K.; Pöhner C.; Boos I.; Schubert V.; Rose-John S.; Unverzagt C., Semisynthesis of biologically active glycoforms of the human cytokine interleukin 6. *Angew. Chem. Int. Ed.*, **2014**, *53*, 12125-12131.
97. Seenaiiah M.; Jbara M.; Mali S. M.; Brik A., Convergent versus sequential protein synthesis: the case of ubiquitinated and glycosylated H2B. *Angew. Chem. Int. Ed.*, **2015**, *54*, 12374-12378.
98. Blanco-Canosa J. B.; Dawson P. E., An Efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem. Int. Ed.*, **2008**, *47*, 6851-6855.
99. Lindgren J.; Ekblad C.; Abrahmsen L.; Karlstrom A. E., A native chemical ligation approach for combinatorial assembly of affibody molecules. *ChemBioChem*, **2012**, *13*, 1024-1031.
100. Fauvet B.; Butterfield S. M.; Fuks J.; Brik A.; Lashuel H. A., One-pot total chemical synthesis of human  $\alpha$ -synuclein *Chem. Commun.*, **2013**, *49*, 9254-9256.
101. Okamoto R.; Mandal K.; Ling M.; Luster A. D.; Kajihara Y.; Kent S. B. H., Total chemical synthesis and biological activities of glycosylated and non-glycosylated forms of the chemokines CCL1 and Ser-CCL1. *Angew. Chem. Int. Ed.*, **2014**, *53*, 5188-5193.
102. Wang J. X.; Fang G. M.; He Y.; Qu D. L.; Yu M.; Hong Z. Y.; Liu L., Peptide *o*-aminoanilides as crypto-thioesters for protein chemical synthesis. *Angew. Chem. Int. Ed.*, **2015**, *54*, 2194-2198.
103. 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*, 7197-7209.
104. Tofteng A. P.; Sorensen K. K.; Conde-Frieboes K. W.; Hoeg-Jensen T.; Jensen K. J., Fmoc solid-phase synthesis of C-terminal peptide thioesters by formation of a backbone pyroglutamyl imide moiety. *Angew. Chem. Int. Ed.*, **2009**, *48*, 7411-7414.
105. Mende F.; Seitz O., 9-Fluorenylmethoxycarbonyl-based solid-phase synthesis of peptide  $\alpha$ -thioesters. *Angew. Chem. Int. Ed.*, **2011**, *50*, 1232-1240.

106. Kawakami T., Peptide thioester formation via an intramolecular *N* to *S* acyl shift for peptide ligation. *Top. Curr. Chem.*, **2015**, *362*, 102-135.
107. Warren J. D.; Miller J. S.; Keding S. J.; Danishesfky S. J., Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc.*, **2004**, *126*, 6576-6578.
108. Botti P.; Villain M.; Manganiello S.; Gaertner H., Native chemical ligation through in situ *O* to *S* acyl shift. *Org. Lett.*, **2004**, *6*, 4861-4864.
109. Zheng J. S.; Cui H. K.; Fang G. M.; Xi W. X.; Liu L., Chemical protein synthesis by kinetically controlled ligation of peptide *O*-Esters. *ChemBioChem*, **2010**, *11*, 511-515.
110. Eom K. D.; Tam J. P., Acid-catalyzed tandem thiol switch for preparing peptide thioesters from mercaptoethyl esters. *Org. Lett.*, **2011**, *13*, 2610-2613.
111. Mills K. V.; Johnson M. A.; Perler F. B., Protein splicing: how inteins escape from precursor proteins. *J. Biol. Chem.*, **2014**, *289*, 14498-14505.
112. Shah N. H.; Muir T. W., Inteins: nature's gift to protein chemists. *Chem. Sci.*, **2014**, *5*, 446-461.
113. Muir T. W.; Sondhi D.; Cole P. A., Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. Usa*, **1998**, *95*, 6705-6710.
114. Ollivier N.; Behr J. B.; El-Mahdi O.; Blanpain A.; Melnyk O., Fmoc solid-phase synthesis of peptide thioesters using an intramolecular *N,S*-acyl shift. *Org. Lett.*, **2005**, *7*, 2647-2650.
115. Kawakami T.; Sumida M.; Nakamura K.; Vorherr T.; Aimoto S., Peptide thioester preparation based on an *N-S* acyl shift reaction mediated by a thiol ligation auxiliary. *Tetrahedron Lett.*, **2005**, *46*, 8805-8807.
116. Nagaike F.; Onuma Y.; Kanazawa C.; Hojo H.; Ueki A.; Nakahara Y.; Nakahara Y., Efficient microwave-assisted tandem *N*- to *S*-acyl transfer and thioester exchange for the preparation of a glycosylated peptide thioester. *Org. Lett.*, **2006**, *8*, 4465-4468.
117. Hojo H.; Onuma Y.; Akimoto Y.; Nakahara Y.; Nakahara Y., *N*-Alkyl cysteine-assisted thioesterification of peptides. *Tetrahedron Lett.*, **2007**, *48*, 25-28.
118. Kang J.; Richardson J.P.; Macmillan D., 3-Mercaptopropionic acid-mediated synthesis of peptide and protein thioesters. *Chem. Commun.*, **2009**, 407-409.
119. Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., Expanding the scope of *N* → *S* acyl transfer in native peptide sequences. *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.
120. Zheng J. S.; Chang H. N.; Wang F. L.; Liu L., Fmoc synthesis of peptide thioesters without post-chain-assembly manipulation. *J. Am. Chem. Soc.*, **2011**, *133*, 11080-11083.
121. Zheng J. S.; Chen X.; Tang S.; Chang H. N.; Wang F. L.; Zuo C., A new method for synthesis of peptide thioesters via irreversible *N*-to-*S* acyl transfer. *Org. Lett.*, **2014**, *16*, 4908-4911.



122. Taichi M.; Hemu X.; Qiu Y.; Tam J. P., A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift. *Org. Lett.*, **2013**, *15*, 2620-2623.
123. Kawakami T.; Aimoto S., The use of a cysteinyl prolyl ester (CPE) autoactivating unit in peptide ligation reactions. *Tetrahedron*, **2009**, *65*, 3871-3877.
124. Kawakami T.; Shimizu S.; Aimoto S., Peptide thioester formation by an *N* to *S* acyl shift reaction at the cysteinyl prolyl cysteine position. *Bull. Chem. Soc. Jpn*, **2010**, *83*, 570-574.
125. Ollivier N.; Dheur J.; Mhidia R.; Blanpain A.; Melnyk O., Bis(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.*, **2010**, *12*, 5238-5241.
126. 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*, 386-389.
127. Ollivier N.; Vicogne J.; Vallin A.; Drobecq H.; Desmet R.; El Mahdi O.; Leclercq B.; Goormachtigh G.; Fajeur V.; Melnyk O., A one-pot three-segment ligation strategy for protein chemical synthesis. *Angew. Chem. Int. Ed.*, **2012**, *51*, 209-213.
128. 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*, 374-377.
129. Dheur J.; Ollivier N.; Melnyk O., Synthesis of thiazolidine thioester peptides and acceleration of native chemical ligation. *Org. Lett.*, **2011**, *13*, 1560-1563.
130. Raibaut L.; Seeberger P.; Melnyk O., Bis(2-sulfanylethyl)amido peptides enable native chemical ligation at proline and minimize deletion side-product formation. *Org. Lett.*, **2013**, *15*, 5516-5519.
131. Tsuda S.; Shigenaga A.; Bando K.; Otaka A., *N*→*S* acyl-transfer-mediated synthesis of peptide thioesters using anilide derivatives. *Org. Lett.*, **2009**, *11*, 823-826.
132. Sato K.; Shigenaga A.; Tsuji K.; Tsuda S.; Sumikawa Y.; Sakamoto K.; Otaka A., *N*-sulfanylethylanilide peptide as a crypto-thioester peptide. *ChemBioChem*, **2011**, *12*, 1840-1844.
133. 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*, 7855-7859.
134. Yang R.; Qi L.; Liu Y.; Ding Y.; Sheng Yi Kwek M.; Liu C. F., Chemical synthesis of *N*-peptidyl 2-pyrrolidinemethanethiol for peptide ligation. *Tetrahedron Lett.*, **2013**, *54*, 3777-3780.
135. Burlina F.; Papageorgiou G.; Morris C.; White P. D.; Offer J., *In situ* thioester formation for protein ligation using  $\alpha$ -methylcysteine. *Chem. Sci.*, **2014**, *5*, 766-770.
136. Asahuna Y.; Nabeshima K.; Hojo H., Peptidyl *N*-alkylcysteine as a peptide thioester surrogate in the native chemical ligation. *Tetrahedron Lett.*, **2015**, *56*, 1370-1373.

## Partie 2

137. Okumura M.; Shimamoto S.; Hidaka Y., A chemical method for investigating disulfide-coupled peptide and protein folding. *FEBS Journal*, **2012**, *279*, 2283-2295.
138. Craik D. J.; Fairlie D. P.; Liras S.; Price D., The future of peptide-based drugs. *Chem. Biol. Drug. Des.*, **2013**, *81*, 136-147.
139. Terrett N., Drugs in middle space. *Med. Chem. Commun.*, **2013**, *4*, 474-475.
140. Góngora-Benítez M.; Tulla-Puche J.; Albericio F., Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chem. Rev.*, **2014**, *114*, 901-926.
141. King G. F., Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther.*, **2011**, *11*, 1469-1484
142. Henriques S. T.; Craik D. J., Cyclotides as templates for drug design. *Drug Discovery Today*, **2010**, *15*, 57-64.
143. Moore S. J.; Cochran J. R., Engineering knottins as novel binding agents. *Methods Enzymol.*, **2012**, *503*, 223-251.
144. Cheek S.; Krishna S. S.; Grishin N. V., Structural classification of small, disulfide-rich protein domains. *J. Mol. Biol.*, **2006**, *359*, 215-237.
145. Craik D. J.; Daly N. L.; Wayne C., The cystine knot motif in toxins and implications for drug design. *Toxicon*, **2001**, *39*, 43-60.
146. 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*, 561-572.
147. Craik D. J.; Daly N. L.; Bond T.; Wayne C., Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.*, **1999**, *294*, 1327-1336.
148. Gressent F.; Da Silva P.; Eyraud V.; Karaki L.; Royer C., Pea albumin 1 subunit b (PA1b), a promising bioinsecticide of plant origin. *Toxins*, **2011**, *3*, 1502-1517.
149. Daly N. L.; Love S.; Alewood P. F.; Craik D. J., Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide Kalata B1. *Biochemistry*, **1999**, *38*, 10606-10614.
150. The KNOTTIN database, Knottins are intriguing miniproteins with strong potential in drug design. Disponible en ligne à : <http://knottin.cbs.cnrs.fr>
151. Thomma B. P.; Cammue B. P.; Thevissen K., Plant defensins. *Planta*, **2002**, *216*, 196-202.
152. Yang D.; Biragyn A.; Kwak L. W.; Oppenheim J. J., Mammalian defensins in immunity: more than just microbicidal. *Trends Immunol.*, **2002**, *23*, 291-296.
153. Bulet P.; Stocklin R., Insect antimicrobial peptides: structures, properties and gene regulation. *Protein Pept. Lett.*, **2005**, *12*, 3-11.

154. Lacerda A. F.; Vasconcelos E. A. R.; Barbosa Pelegrini P.; Grossi de Sa M. F., Antifungal defensins and their role in plant defense. *Front. Microbiol.*, **2014**, *5*, 1-10.
155. Conibear A. C.; Craik D. J., The chemistry and biology of theta defensins. *Angew. Chem. Int. Ed.*, **2014**, *53*, 10612-10623.
156. Lewis R. J.; Garcia M. L., Therapeutic potential of venom peptides. *Nat. Rev. Drug Discovery*, **2003**, *2*, 790-802.
157. Lavergne V.; Taft R. J.; Alewood P. F., Cysteine-rich mini proteins in human biology. *Curr. Top. Med. Chem.*, **2012**, *12*, 1514-1533.
158. Mouhat S.; Jouirou B.; Mosbah A.; De Waard M.; Sabatier J-M., Diversity of folds in animal toxins acting on ion channels. *Biochem. J.*, **2004**, *378*, 717-726.
159. Manjunatha Kini R.; Doley R., Structure, function and evolution of three-finger toxins: Mini proteins with multiple targets. *Toxicon*, **2010**, *56*, 855-867.
160. Rouget C.; Quinton L.; Maïga A.; Gales C.; Masuyer G.; Malosse C.; Chamot-Rooke J.; Thai R.; Mourier G.; de Pauw E.; Gilles N.; Servent D., Identification of a novel snake peptide toxin displaying high affinity and antagonist behaviour for the  $\alpha_2$ -adrenoceptors. *Br. J. Pharmacol.*, **2010**, *161*, 1361-1374.
161. Saito T.; Kawabata S-I.; Shigenaga T.; Takayenoki Y.; Cho J.; Nakajima H.; Hirata M.; Iwanaga S., A novel big defensin identified in horseshoe crab hemocytes: isolation, amino acid sequence, and antibacterial activity. *J. Biochem.*, **1995**, *117*, 1131-1137.
162. Hervé-Grépinet V.; Réhault-Godbert S.; Labas V.; Magallon T.; Derache C.; Lavergne M.; Gautron J.; Lalmanach A-C.; Nys Y., Purification and characterization of avian  $\beta$ -defensin 11, an antimicrobial peptide of the hen egg. *Antimicrob. Agents Chemother.*, **2010**, *54*, 4401-4408.
163. Bae C.; Kalia J.; Song I.; Yu J.; Kim H. H.; Swartz K. J.; Kim J. I., High yield production and refolding of the double-knot toxin, an activator of TRPV1 channels. *PLoS One*, **2012**, *7*, e51516.
164. Vlieghe P.; Lisowski V.; Martinez J.; Khrestchatsky M., Synthetic therapeutic peptides: science and market. *Drug Discovery Today*, **2010**, *15*, 40-56.
165. Medini K.; Harris P. W. R.; Hards K.; Dingley A. J.; Cook G. M.; Brimble M. A., Chemical synthesis of a pore-forming antimicrobial protein, caenopore-5, by using native chemical ligation at a glu-cys site. *ChemBioChem*, **2015**, *16*, 328-336.
166. 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*, 1806-1809.



# CHAPITRE 2



## Chapitre 2 : Développement d'une méthode simple et bio-inspirée de synthèse de peptides crypto-thioester par Fmoc-SPPS

La partie 1 du chapitre précédent a souligné le rôle central que jouent les peptides C $\alpha$ -thioester dans la synthèse de peptides et protéines par voie chimique. Ils sont notamment des partenaires clés dans la ligation chimique native, réaction qui a révolutionné ce domaine. Comme nous l'avons vu précédemment, la synthèse de ces partenaires thioester s'avère délicate par la stratégie Fmoc largement utilisée en SPPS. La section 1.5 du chapitre 1 a mis en évidence le nombre important de méthodes proposées par les scientifiques du domaine pour répondre à ce problème; ces méthodes reposent principalement sur la synthèse de précurseurs stables à la pipéridine pouvant être converti en thioester après la SPPS.

<sup>1</sup> Néanmoins, comme nous l'avons également vu, aucune des méthodes développées ne s'est pas vraiment encore imposée universellement dans les laboratoires. Les méthodes proposées par l'équipe de L. Liu (peptides hydrazide)<sup>2</sup> et par l'équipe de Dawson (peptides Nbz et MeNbz)<sup>3,4</sup> sont aujourd'hui très utilisées mais elles souffrent néanmoins de limitations, exposées lors de leur description.

De nombreuses stratégies exploitant un réarrangement par transfert d'acyle intramoléculaire  $N \rightarrow S$  par utilisation d'un dispositif de thioestérification  $\beta$ - ou  $\gamma$ -mercapto amide C-terminal ont été développées,<sup>5</sup> démontrant ainsi le potentiel de cette approche très prometteuse. Néanmoins, lors de l'étude bibliographique des méthodes existantes, nous avons identifié les points suivants comme inconvénients souvent récurrents des stratégies existantes : la synthèse complexe du dispositif, le couplage difficile de l'acide aminé C-terminal sur l'amine secondaire encombrée de ces dispositifs. De plus, la plupart de ces dispositifs montrent une faible réactivité à pH neutre, nécessitant une étape de thioestérification en milieu acide avant la NCL, seules quelques stratégies ont montré l'utilisation directe en NCL de peptides portant ce type de dispositif (peptides crypto-thioester). Enfin, dans ces derniers cas, il peut être nécessaire

---

<sup>1</sup> Mende F.; Seitz O., *Angew. Chem. Int. Ed.*, **2011**, *50*, 1232-1240.

<sup>2</sup> Fang G. M.; Li Y. M.; Shen F.; Huang Y. C.; Li J. B.; Lin Y.; Cui H. K.; Liu L., *Angew. Chem. Int. Ed.*, **2011**, *50*, 7645-7649.

<sup>3</sup> Blanco-Canosa J. B.; Dawson P. E., *Angew. Chem. Int. Ed.*, **2008**, *47*, 6851-6855.

<sup>4</sup> Blanco-Canosa J. B.; Nardone B.; Albericio F.; Dawson P. E., *J. Am. Chem. Soc.*, **2015**, *137*, 7197-7209.

<sup>5</sup> Kawakami T., *Top. Curr. Chem.*, **2015**, *362*, 102-135.

d'appliquer une étape post-SPPS pour stabiliser les peptides portant ces dispositifs afin d'éviter une activation précoce du transfert d'acyle  $N \rightarrow S$  lors de leur manipulation en milieux aqueux.

Etant convaincus qu'une méthode générale, peu coûteuse et facile à mettre en œuvre aurait un fort impact au sein de la communauté des sciences des peptides et protéines, nous avons cherché à développer une stratégie répondant, point par point, aux limitations habituellement rencontrées dans les stratégies existantes de thioestérification par transfert d'acyle  $N \rightarrow S$ . Ces résultats ont été acceptés pour publication dans le journal *Chemical Science* (article joint ci-après : 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-thioester. *Chemical Science*, **2016**, 7, 339-345).

Lors du *design* du dispositif validé et exploité au cours de cette thèse, nous avons choisi de nous focaliser sur la stratégie utilisant la *N*-alkyl-cystéine étant donné que ce dispositif peut être facilement installé sur support solide à partir de réactifs commerciaux peu coûteux.<sup>6,7</sup> Pour tenter d'accélérer les cinétiques lentes observées à pH neutre lors de l'utilisation des *N*-alkyl-cystéines existantes directement dans des réactions de NCL, nous nous sommes inspirés de la modification post-traductionnelle opérée *in vivo* par les intéines (d'où le terme de stratégie « bio-inspirée »).<sup>8</sup> La première étape de l'épissage protéique réalisé par les intéines consiste en un transfert d'acyle  $N \rightarrow S$  intramoléculaire auto-catalysé, faisant notamment intervenir un résidu histidine hautement conservé, pour former une liaison thioester.<sup>9</sup> L'histidine *N*-protonée (pKa d'environ 6) à pH physiologique est supposée accélérer ce transfert d'acyle par une catalyse acido-basique intramoléculaire (figure 1a). Nous avons postulé que la fonctionnalisation adéquate d'un dispositif de type *N*-alkyl-cystéine par un groupement donneur de proton, mimant l'histidine conservée chez les intéines, résulterait en une accélération du transfert d'acyle  $N \rightarrow S$  à pH neutre. Nous avons donc conçu un dispositif de type *N*-(2-hydroxybenzyl)cystéine dont le groupement phénol clé, avec un pKa proche de la neutralité, pourrait participer à la catalyse intramoléculaire par transfert de proton (figure 1b). Ce type de catalyse rappelle le mécanisme proposé par Otaka et ses collaborateurs pour la thioestérification de peptides SEALide assistée par une catalyse pseudo-intramoléculaire par les sels de phosphate,<sup>10</sup> même si dans notre cas, le dispositif participerait directement dans la catalyse de façon intramoléculaire.

<sup>6</sup> Hojo H.; Onuma Y.; Akimoto Y.; Nakahara Y.; Nakahara Y., *Tetrahedron Lett.*, **2007**, 48, 25-28.

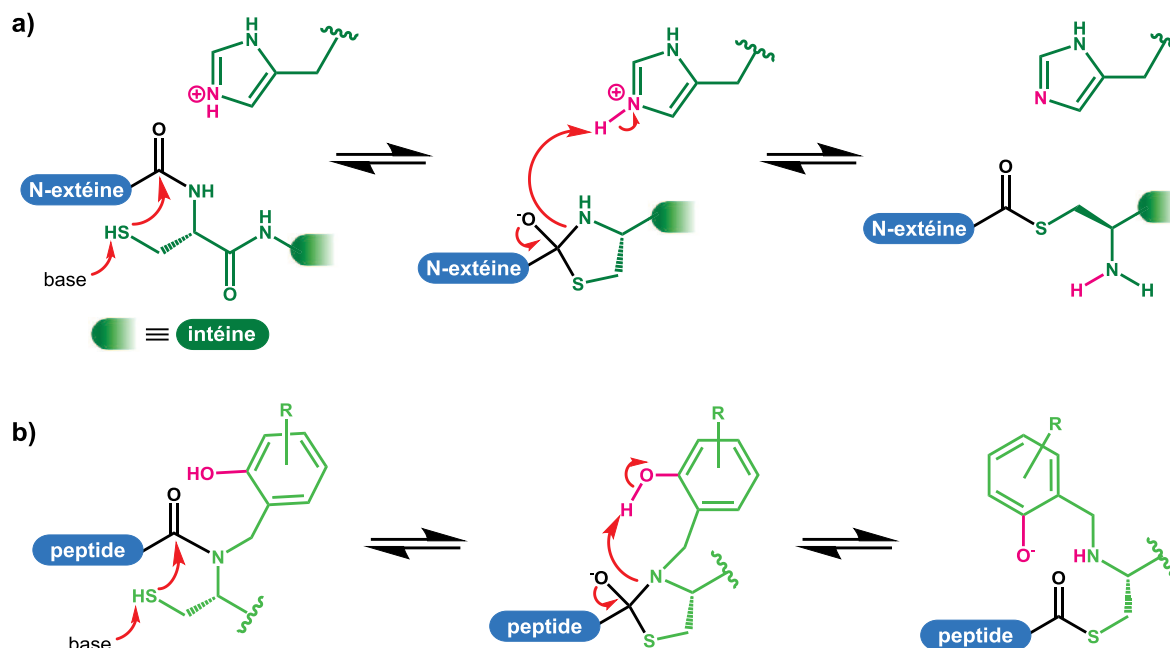
<sup>7</sup> Asahuna Y.; Nabeshima K.; Hojo H., *Tetrahedron Lett.*, **2015**, 56, 1370-1373.

<sup>8</sup> Mills K. V.; Johnson M. A.; Perler F. B., *J. Biol. Chem.*, **2014**, 289, 14498-14505.

<sup>9</sup> Binschik J.; Mootz H. D., *Angew. Chem. Int. Ed.*, **2013**, 52, 4260-4264.

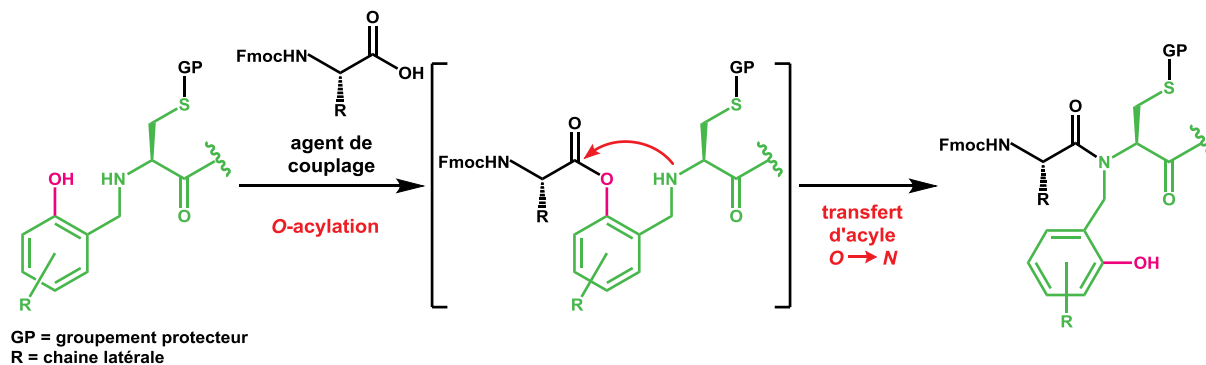
<sup>10</sup> Sato K.; Shigenaga A.; Tsuji K.; Tsuda S.; Sumikawa Y.; Sakamoto K.; Otaka A., *ChemBioChem*, **2011**, 12, 1840-1844.





**Figure 1 :** a) Catalyse acido-basique intramoléculaire lors de l'excision des intéines par une histidine hautement conservée; b) mécanisme bio-inspiré hypothétique pour la formation d'une fonction thioester à l'aide d'un dispositif de type *N*-(2-hydroxybenzyl)cystéine

Un avantage additionnel de cette approche biomimétique réside dans l'assistance potentielle du groupement hydroxyle de la *N*-acylation de l'amine secondaire du dispositif lors l'introduction de l'acide aminé C-terminal de la séquence peptidique. Le mécanisme anticipé pourrait reposer sur la participation du groupement hydroxyle qui serait d'abord *O*-acylé, puis la liaison amide serait formée par un transfert d'acyle *O*→*N* intramoléculaire (figure 2). Ce mécanisme est rencontré lors de la *N*-acylation des amines protégées par un groupement 2-hydroxy-4-méthoxybenzyle (Hmb).<sup>11</sup> Une telle assistance permettrait de résoudre le problème du couplage souvent difficile de l'acide aminé C-terminal sur ce type de dispositif portant une amine encombrée.



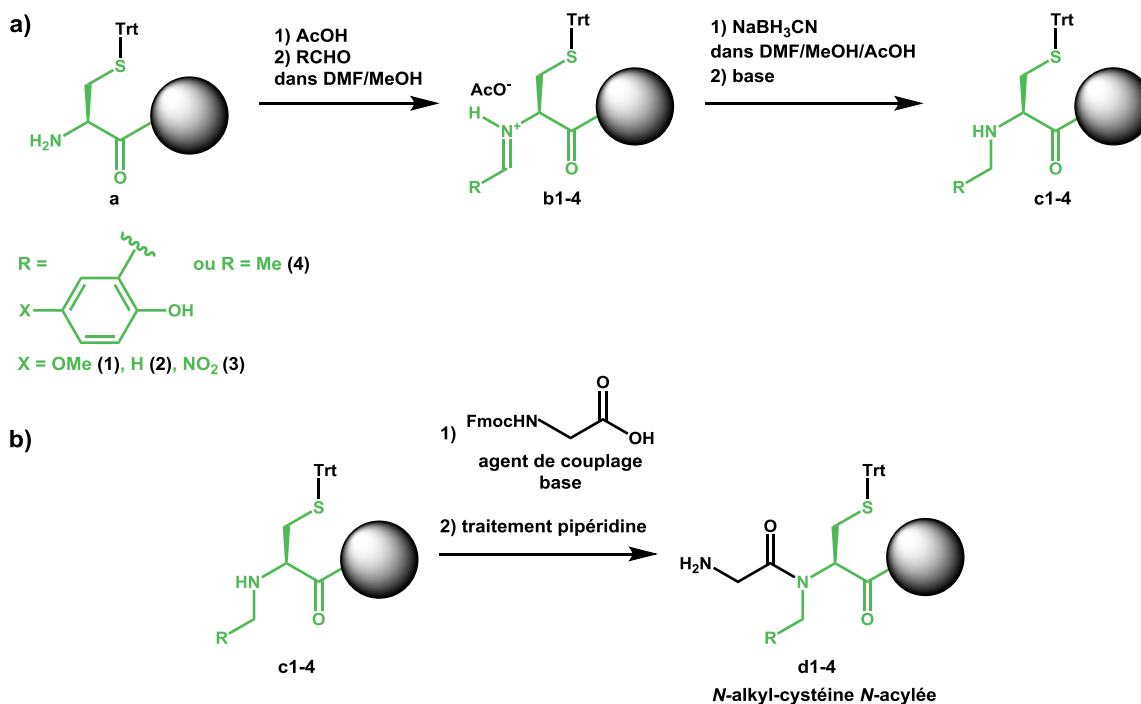
**Figure 2 :** Mécanisme de la *N*-acylation d'un dispositif de type *N*-(2-hydroxybenzyl)cystéine

<sup>11</sup> Johnson T.; Quibell M.; Owen D.; Sheppard R. C., *J. Chem. Soc., Chem. Commun.*, **1993**, 137, 369-372.

L'étude préliminaire de cette nouvelle stratégie avait été initiée par le Dr. Hélène Adihou, la doctorante m'ayant précédée au sein de notre équipe. Il a d'abord fallu d'abord optimiser la structure du groupement 2-hydroxybenzyle du dispositif, dans le but d'obtenir une assistance maximale du groupement hydroxyle pour la *N*-acylation de la *N*-benzyl-cystéine. Différents groupements, ayant différentes propriétés électroniques, ont été installés sur une cystéine *S*-protégée greffée sur support solide *via* une réaction d'amination réductrice à l'aide de l'aldéhyde approprié et d'un réducteur (composés **c1** à **c3**, figure 3a). La cystéine est d'abord traitée par une solution d'acide acétique pour former le sel d'acétate, permettant la formation de l'imine par réaction avec un aldéhyde (composés **b**, figure 3a), un traitement basique est réalisé après la réduction pour éliminer ce sel d'acétate et permettre ensuite le couplage de l'acide aminé C-terminal sur la *N*-alkyl-cystéine. Il est à noter que les conditions d'amination réductrice ont été entièrement automatisées, permettant la réalisation complète de la synthèse de notre dispositif à l'aide d'un synthétiseur de peptides. Ces dispositifs ont été couplés avec une Fmoc-Gly-OH dans des conditions de couplage standard pour obtenir les composés *N*-acylés **d1** à **d3** (figure 3b). Les différences d'effet électronique des différents groupements hydroxybenzyle testés résultent en des rendements de *N*-acylation variables. Le dispositif *N*-(2-hydroxy-5-nitrobenzyl)cystéine (*N*-Hnb-Cys, composé **c3**) a été retenu car il a démontré les meilleurs rendements de *N*-acylation, en comparaison d'un contrôle négatif représenté par le couplage de la Fmoc-Gly-OH sur une *N*-éthyl-cystéine (composé **c4**) qui n'est pas réactive. Cette assistance de la *N*-acylation avait déjà été observée dans la littérature pour des groupements 2,5- et 2,6-Hnb.<sup>12</sup> De plus, le pKa du groupement phénol du dispositif *N*-Hnb-Cys (~ 6,5) est proche de la neutralité, ce qui répond à un critère souligné comme important lors du *design* du dispositif pour une catalyse biomimétique de la thioestérification efficace.

---

<sup>12</sup> Miranda L. P.; Meutermans W. D.; Smythe M. L.; Alewood P. F., *J. Org. Chem.*, **2000**, 65, 5460-5468.



**Figure 3 :** a) Amination réductrice d'une cystéine sur support solide à l'aide d'un aldéhyde et d'un réducteur; b) couplage d'une Fmoc-Gly-OH sur les N-alkyl-cystéines supportées **c1-4**

Lors de la synthèse de peptides comportant notre dispositif avec une cystéine, dont le groupement protecteur (*S*-trityle) est éliminé dans des conditions de coupure TFA, nous avons observé l'instabilité des peptides obtenus lors de leur purification, compliquant ainsi leur manipulation. La cause en est un transfert d'acyle *N*→*S* précoce en milieu acide. Ce problème est fréquent pour les peptides portant de tels dispositifs de thioestérification, c'est le cas par exemple pour les peptides TEA décrits par Tam et ses collaborateurs<sup>13</sup> ou encore les peptides SEA, Melnyk et ses collaborateurs ont montré qu'il était possible de les stabiliser par une étape d'oxydation post-SPPS sous forme de disulfure.<sup>14</sup> Dans le cas du dispositif *N*-Hnb-Cys, il suffit d'utiliser simplement un groupement résistant aux conditions de SPPS et de coupure TFA, nous avons choisi le groupement *tert*-butylsulfanyle (*StBu*) protégeant la cystéine sous forme d'un pont disulfure.<sup>15</sup> Ce groupement est complètement stable lors de la Fmoc-SPPS et nous avons vérifié qu'il était rapidement éliminé, comme attendu,<sup>14</sup> dans les conditions réductrices de la NCL.

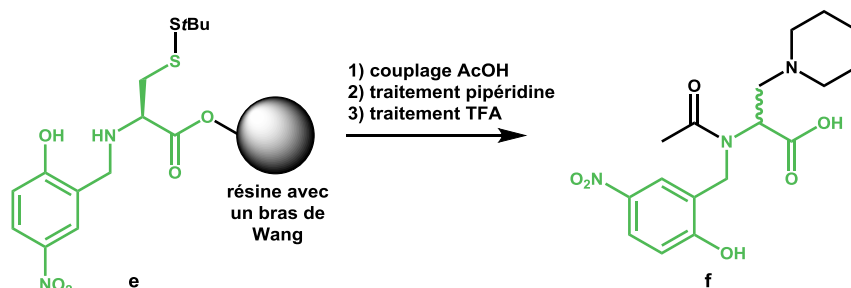
Le choix du lien entre la *N*-Hnb-cystéine et la résine s'est également révélé être important. Dans le cas d'une résine fonctionnalisée par un bras de Wang, l'épimérisation de la

<sup>13</sup> Taichi M.; Hemu X.; Qiu Y.; Tam J. P., *Org. Lett.*, **2013**, *15*, 2620-2623.

<sup>14</sup> Ollivier N.; Dheur J.; Mhidia R.; Blanpain A.; Melnyk O., *Org. Lett.*, **2010**, *12*, 5238-5241.

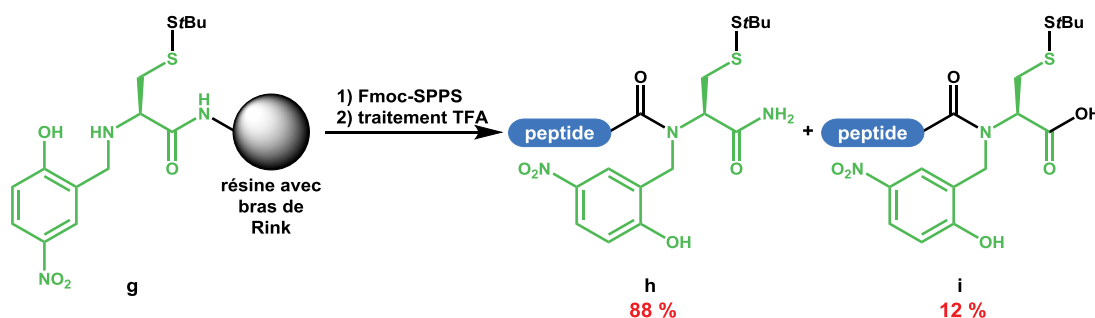
<sup>15</sup> Weber U.; Hartter P., *Hoppe-Seyler's Z. Physiol. Chem.*, **1970**, *351*, 1384-1388.

cystéine C-terminale<sup>16</sup> et la  $\beta$ -élimination du groupement thiol, suivies de l'addition de la pipéridine<sup>17</sup> conduisent à l'obtention exclusive d'un produit *N*-Hnb-3-(*N*-pipéridinyl)-alanine C-terminal (composé **f**, figure 4). L'utilisation d'une résine fonctionnalisée par un bras de type chloro-trityle a également été infructueuse.



**Figure 4 :** Utilisation d'une résine-acide pour la synthèse de peptides *N*-Hnb-Cys

Dans le cas d'une résine fonctionnalisée par un bras de Rink, après traitement TFA, un mélange de produit est obtenu (88:12) contenant le produit désiré (composé **h**, figure 5) en proportion majoritaire et un produit provenant de l'hydrolyse de l'amide C-terminal en acide carboxylique (composé **i**, figure 5) lors du traitement TFA.



**Figure 5 :** Utilisation d'une résine-amide pour la synthèse de peptides *N*-Hnb-Cys

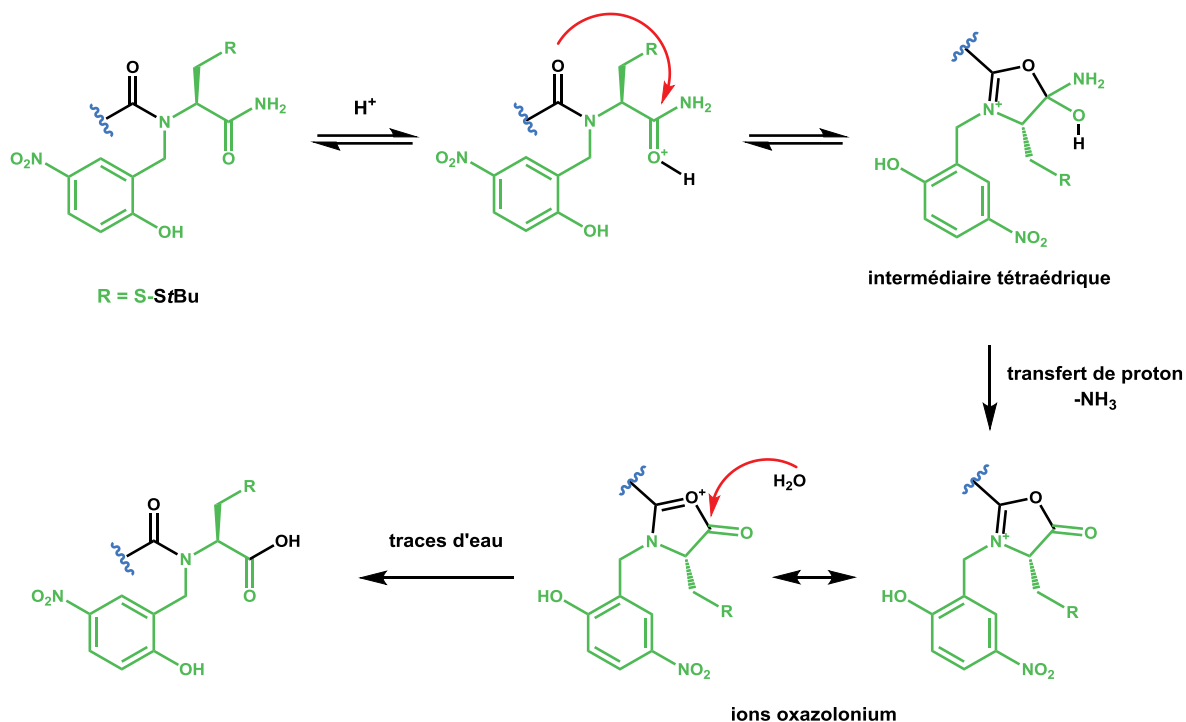
Une réaction secondaire apparentée a été décrite pour des acides aminés *N*-alkylés.<sup>18,19</sup> Le mécanisme postulé consiste en une attaque nucléophile du carbonyle de l'acide aminé *N*-alkylé protoné par le TFA, formant un intermédiaire tétraédrique. Cet intermédiaire peut se réarranger par transfert de proton et perte d'une molécule d'ammoniaque pour conduire à la formation d'ions oxazolonium dont l'ouverture par attaque nucléophile de l'eau présente à l'état de trace conduit à l'acide carboxylique C-terminal (figure 6).

<sup>16</sup> Tawaragi Y.; Fuchimara K.; Tanaka S.; Minamino N.; Kangawa K.; Matsuo H., *Biochem. Biophys. Res. Commun.*, **1991**, 175, 645-651.

<sup>17</sup> Lukszo J.; Patterson D.; Albericio F.; Kates S. A., *Lett. Pept. Sci.*, **1996**, 3, 157-166.

<sup>18</sup> Urban J.; Vaisar T.; Shen R.; Lee M. S., *Int. J. Pept. Prot. Res.*, **1996**, 47, 182-189.

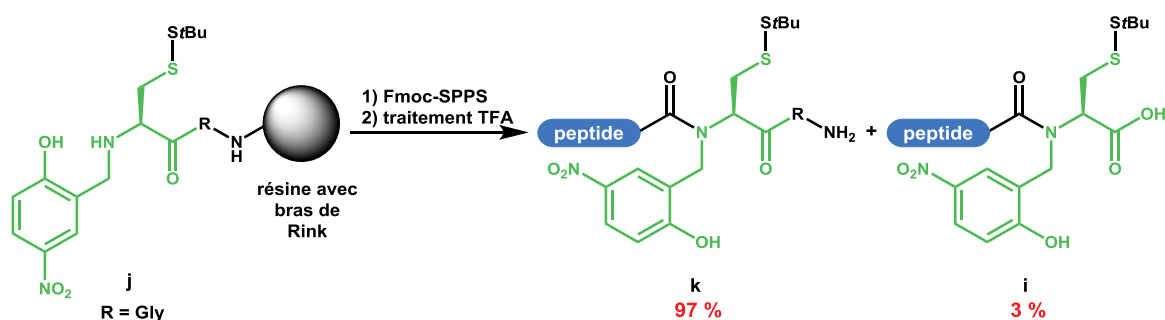
<sup>19</sup> Teixidó M.; Albericio F.; Giralt E., *J. Pept. Res.*, **2005**, 65, 153-166.



**Figure 6 :** Mécanisme proposé pour l'hydrolyse acide de la fonction amide C-terminale d'un peptide *N-Hnb-Cys*<sup>18,19</sup>

De façon intéressante, nous avons observé que la quantité de produit hydrolysé augmentait quand la durée du traitement TFA était rallongée, indiquant que ce phénomène serait lent. Nous avons également observé qu'en diminuant la proportion d'eau, il était possible d'abaisser la quantité de produit hydrolysé formé, ce qui va dans le sens du mécanisme proposé.

Néanmoins pour éviter ce problème d'hydrolyse, une solution simple consiste à insérer une glycine entre la *N-Hnb-cystéine* et la résine fonctionnalisée par un bras de Rink. La réaction d'hydrolyse de l'amide de la cystéine *N*-alkylée est alors limitée à moins de 3 % (figure 7).

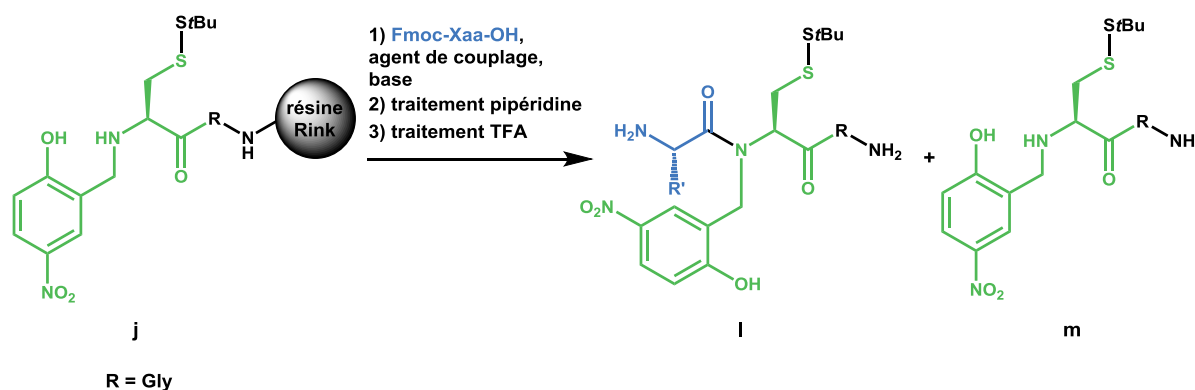


**Figure 7 :** Insertion d'une glycine entre la résine et le dispositif *N-Hnb-Cys*

De plus, lors d'études préliminaires de la NCL avec les peptides **h**, **i** et **k** et un cystéinyl-peptide modèle, nous n'avons pas observé de différence significative dans les cinétiques de réaction. Alors que Macmillan et ses collaborateurs ont décrit des NCLs plus rapides avec

l'utilisation d'un peptide comportant une cystéine C-terminale acide plutôt qu'amide, postulant une catalyse intramoléculaire de la thioestérification par la fonction C-terminale,<sup>20</sup> la faible différence de réactivité des peptides *N*-Hnb-Cys-OH, *N*-Hnb-Cys-NH<sub>2</sub> et *N*-Hnb-Cys-Gly-NH<sub>2</sub> exclut possiblement une catalyse de ce type dans le cas de l'utilisation de notre dispositif.

Nous avons étudié la *N*-acylation du dispositif optimisé par un couplage simple avec les vingt *N*-Fmoc-acides aminés canoniques dans des conditions standards (figure 8). Dans la majorité des cas, les rendements de *N*-acylation obtenus sont bons voire excellents, appuyant le phénomène d'assistance du groupement hydroxyle dans la *N*-acylation de ce dispositif. Néanmoins, comme attendu, la *N*-acylation avec des acides aminés encombrés stériquement (Thr, Pro, Val et Ile) s'est avérée difficile. Dans un souci de conserver la possibilité de synthèse entièrement automatisée de peptides portant le motif *N*-Hnb-Cys à partir de réactifs peu coûteux, une solution que nous avons proposée réside dans la réalisation de couplages multiples pour augmenter le rendement de *N*-acylation. Une autre solution consiste à réaliser le couplage dans des conditions standards sous activation thermique, par exemple, à l'aide de micro-ondes, ce qui permet d'obtenir de très bons rendements avec les acides aminés les plus encombrés.

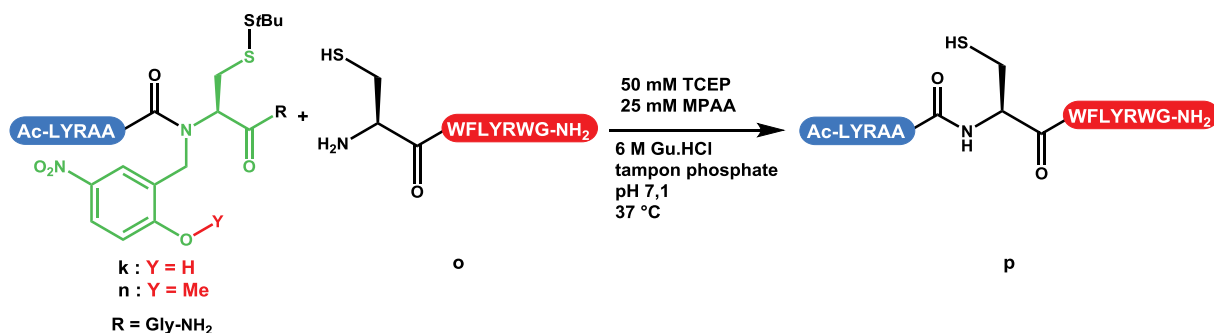


**Figure 8 :** *N*-acylation du dispositif *N*-Hnb-Cys

Nous avons ensuite voulu vérifier notre hypothèse de catalyse biomimétique de la thioestérification de notre dispositif à pH neutre par protonation intramoléculaire (voir figure 1). Pour cela, nous avons comparé les vitesses de NCLs utilisant un peptide portant notre dispositif *N*-Hnb-Cys (composé **k**) ou un peptide dont le groupement hydroxyle du dispositif serait masqué sous forme d'éther méthylique de phénol (composé **n**) ne changeant pas l'encombrement stérique du dispositif de façon prononcée, avec un cystéinyl-peptide modèle (composé **o**, figure 9). La NCL faisant intervenir le composé *O*-méthylé **n** est au moins cinquante fois plus lente que celle faisant intervenir le composé **k**. Ce résultat très important est une

<sup>20</sup> Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.

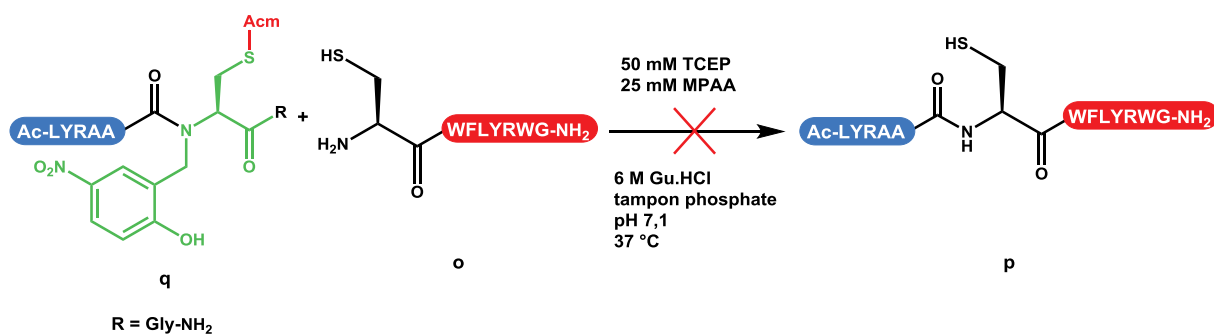
preuve de l'importance cruciale que joue le groupement hydroxyle de notre dispositif dans la thioestérification à pH neutre.



**Figure 9 :** NCL avec un peptide *N*-Hnb-Cys **k** ou un peptide avec un dispositif *O*-méthylé **n**

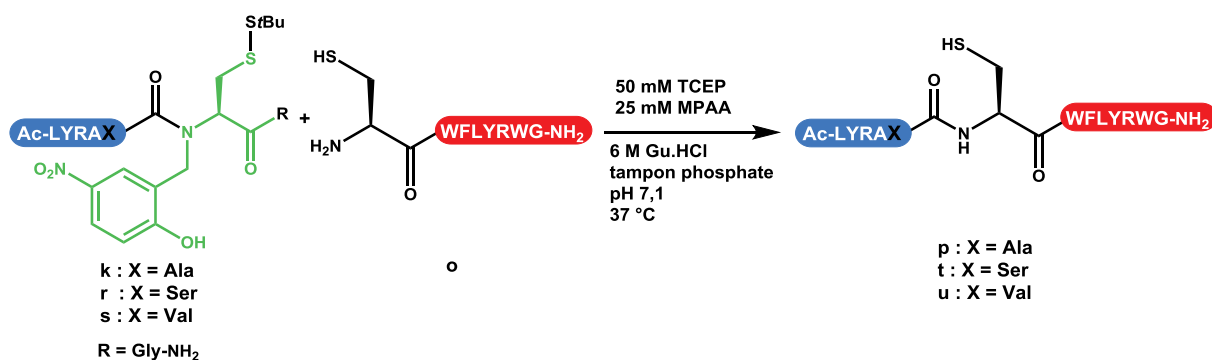
Des études poussées de NCLs utilisant un peptide *N*-Hnb-Cys **k** à différents pHs ont montré que le pH optimal de cette réaction se situait vers pH 6,5 alors que la cinétique de la réaction à un pH d'environ 5,5 est plus lente. Ce résultat, en accord avec le pKa du groupement phénol de notre dispositif, suggère la nécessité d'un juste milieu entre la protonation de ce groupement phénol et la déprotonation du groupement thiol du dispositif. La NCL avec les peptides **k** et **o** en l'absence de tampon phosphate (tampon HEPES) n'a pas montré de différence significative de cinétique par rapport à une NCL utilisant le tampon phosphate, discriminant la possibilité d'une catalyse pseudo-intramoléculaire des sels de phosphate comme postulée par Otaka et ses collaborateurs.<sup>10</sup>

Pour écarter la possibilité d'un mécanisme de transfert d'acyle intramoléculaire *N*→*O* lors de la NCL utilisant notre dispositif, nous avons bloqué son groupement thiol par utilisation d'un groupement acétamidométhyle (Acm) stable dans les conditions de NCL (figure 10). Nous n'avons pas observé de réaction significative, ce qui suggère également l'application potentielle de notre méthode à des NCLs successives de *N* vers *C* en désactivant temporairement le dispositif de thioestérification.



**Figure 10 :** NCL avec un peptide *N*-Hnb-Cys(Acm) **q**

Pour étudier le caractère général de notre stratégie, nous avons évalué les cinétiques de NCLs impliquant des peptides *N*-Hnb-Cys comportant différents acides aminés C-terminaux sur le dispositif (figure 11). Les cinétiques des réactions impliquant une alanine ou une sérine C-terminale sont par exemples similaires. Comme attendu, la réaction impliquant une valine C-terminale (compose **s**) est bien plus lente; à pH 6,6, une augmentation de la température et de la concentration en thiol exogène comme catalyseur permettent d'accélérer la NCL à un niveau proche des NCLs impliquant des acides aminés C-terminaux beaucoup moins encombrés stériquement.



**Figure 11** : NCL avec différents peptides *N*-Hnb-Cys

Il faut noter que nous avons étudié l'épimérisation pouvant avoir lieu lors du transfert d'acyle *N*→*S* ou la NCL à pH neutre. Pour cela nous avons synthétisé le produit de ligation contenant une D-Ala comme jonction en tant que référence HPLC du produit d'épimérisation. Après 24 h d'une NCL utilisant le peptide *N*-Hnb-Cys **k**, le produit d'épimérisation a été observé uniquement à l'état de trace (< 0,4 %) dans le mélange brut de NCL; notre stratégie n'est donc pas épimérisante.

La stratégie a été appliquée avec succès aux formes réduites de deux longs peptides naturels riches en ponts disulfure complexes à synthétiser, la toxine muscarinique 7 (MT7, 65 acides aminés, 8 cystéines)<sup>21</sup> issue du venin du mamba vert et la défensine antimicrobienne *Cg*-BigDef1 (93 acides aminés, 6 cystéines)<sup>22</sup> issue de l'huitre *Crassostrea gigas*. Les peptides *N*-Hnb-Cys utilisés pour ces applications, de 41 (MT7) et 56 acides aminés (*Cg*-BigDef1), représentent des exemples particulièrement longs de précurseurs de peptides thioester.

<sup>21</sup> Mourier G.; Dutertre S.; Fruchart-Gaillard C.; Ménez A.; Servent D., *Mol. Pharmacol.*, **2003**, 63, 26-35.

<sup>22</sup> Rosa R. D.; Santini A.; Fievet J.; Bulet P; Destoumieux-Garzón D.; Bachère E., *PLoS One*, **2011**, 6, e25594.



## Chapitre 2

En conclusion, nous avons donc développé un nouveau dispositif de thioestérification opérant à pH neutre dans des conditions de NCL par un transfert d'acyle  $N \rightarrow S$  grâce à une catalyse biomimétique intramoléculaire. La synthèse du dispositif est réalisable de façon entièrement automatisée à partir de réactifs commerciaux peu coûteux à l'aide d'un synthétiseur de peptide, ce qui représente une avancée significative dans ce domaine et rend possible l'utilisation de cette stratégie par des personnes non-spécialistes.



CrossMark  
click for updatesCite this: *Chem. Sci.*, 2016, 7, 339

## A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters†

Victor P. Terrier, H el ene Adihou, Mathieu Arnould, Agn es F. Delmas and Vincent Aucagne\*

Despite recent advances, the direct Fmoc-based solid phase synthesis of peptide  $\alpha$ -thioesters for the convergent synthesis of proteins *via* native chemical ligation (NCL) remains a challenge in the field. We herein report a simple and general methodology, enabling access to peptide thioester surrogates. A novel C-terminal *N*-(2-hydroxybenzyl)cysteine thioesterification device based on an amide-to-thioester rearrangement was developed, and the resulting peptide crypto-thioesters can be directly used in NCL reactions with fast  $N \rightarrow S$  shift kinetics at neutral pH. These fast kinetics arise from our bio-inspired design, *via* intein-like intramolecular catalysis. Due to a well-positioned phenol moiety, an impressive >50 fold increase in the kinetic rate is observed compared to an *O*-methylated derivative. Importantly, the synthesis of this new device can be fully automated using inexpensive commercially available materials and does not require any post-synthetic steps prior to NCL. We successfully applied this new method to the synthesis of two long naturally-occurring cysteine-rich peptide sequences.

Received 20th July 2015  
Accepted 22nd September 2015

DOI: 10.1039/c5sc02630j

www.rsc.org/chemicalscience

### Introduction

Peptide  $\alpha$ -thioesters are key intermediates for the convergent synthesis of proteins through native chemical ligation (NCL), a reaction that has revolutionized the field.<sup>1</sup> While peptide thioesters can be directly synthesized through Boc-based solid-phase peptide synthesis (SPPS), their access *via* the more widely used Fmoc-based strategy is not straightforward due to the instability of the thioester moiety to repeated piperidine treatments used for Fmoc deprotection. Most Fmoc-based methodologies developed to date rely on a post-SPPS conversion of a piperidine-stable precursor into a thioester, but despite considerable efforts, no universal strategy has emerged.<sup>2</sup> Thanks to their simple implementation, methods such as Dawson's *N*-acylurea (Nbz)<sup>3</sup> and Liu's hydrazides<sup>4</sup> for the generation of thioester precursors have recently become popular, but are still associated with several limitations.<sup>5</sup> Thus, a reliable and straightforward route to peptide  $\alpha$ -thioesters *via* Fmoc-SPPS is highly desirable.

A number of strategies based on  $\beta$ - or  $\gamma$ -mercapto amide thioesterification devices have recently appeared that exploit an amide-to-thioester rearrangement<sup>6,7</sup> (Fig. 1A), which is reminiscent of the first step of intein-promoted *in vivo* protein splicing.<sup>8</sup> In most cases, the  $N \rightarrow S$  acyl shift is reversible and

requires acidic conditions to proceed efficiently, meaning that the peptides have to be first converted into thioesters prior to NCL, which is optimal at pH  $\sim$  7. A few systems, referred to as crypto-thioesters, are able to operate under NCL conditions, thus enabling one-pot reactions (thioester formation and NCL). Cysteine-prolyl esters (CPEs) developed by Aimoto<sup>7d</sup> (Fig. 1B-b) rely on an elegant intramolecular  $O \rightarrow N$  shift to displace the amide-thioester equilibrium under NCL conditions, but with slow ligation kinetics. Although simple to implement, the  $\alpha$ -methyl cysteine crypto-thioesters recently introduced by Offer<sup>7j</sup> (Fig. 1B-e) are impaired by similarly slow kinetics. In both cases, the limiting factor is likely the disfavored *trans*-to-*cis* isomerization of the amide bond to generate the geometrically feasible intramolecular attack of the thiolate. Use of *N,N*-disubstituted  $\beta$ -mercapto amides can overpass this limitation by disfavoring the unproductive *trans* conformer. The symmetrical bis-sulfanylethylamides (SEAs, Fig. 1B-c) introduced by Melnyk<sup>7g</sup> and Liu<sup>7h</sup> can be used in NCL at pH 6–7 with relatively fast kinetics, but work better at lower pH. Otaka's *N*-sulfanylethylamides<sup>7e</sup> (Sealides, Fig. 1B-d) exploit the electronic effect of the anilide to enhance the electrophilicity of the amide bond. They are efficient NCL substrates at neutral pH only when using a phosphate-based buffer, probably arising from pseudo-intramolecular acid–base catalysis, but the exact mechanism remains elusive.<sup>9</sup> Interestingly, while we were preparing this paper, a method appeared based on an *N*-ethyl cysteine device featuring a free carboxylic acid; peptides equipped with this device can act as crypto-thioesters. The  $\alpha$ -acid group was shown to accelerate the reaction compared to an  $\alpha$ -amide-containing

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

† Electronic supplementary information (ESI) available: Detailed synthetic procedures, characterization and kinetics studies. See DOI: 10.1039/c5sc02630j

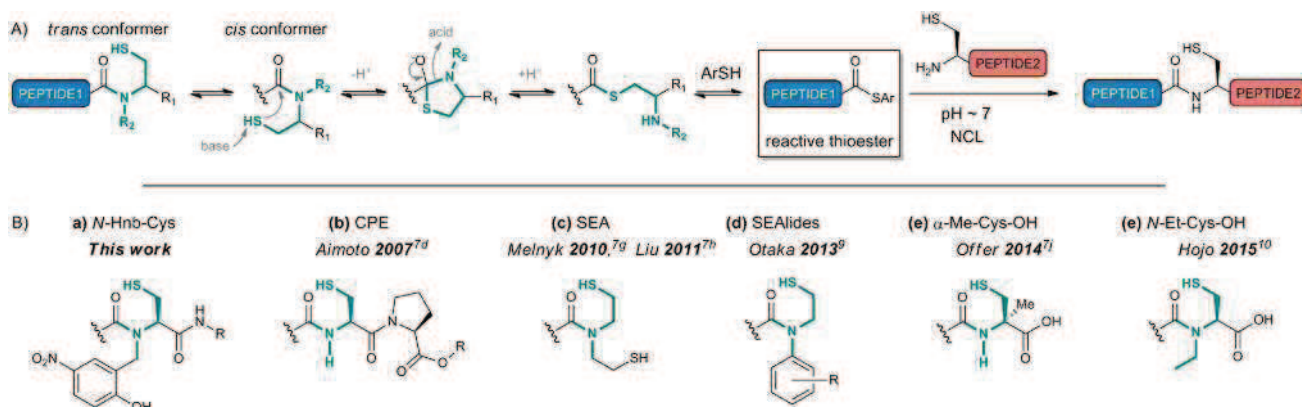


Fig. 1  $N \rightarrow S$  shift-based *in situ* synthesis of peptide thioesters for NCL. (A) Mechanism. (B) Crypto-thioesters bearing thioesterification devices operating under NCL conditions.

analogue, with optimum ligation kinetics at pH 5.5, but with very slow kinetics at pH 7.<sup>10</sup>

We herein report the design and optimization of a simple and general methodology based on an  $N$ -(2-hydroxybenzyl) cysteine device, which can be automatically assembled in the solid-phase using inexpensive commercially available materials. Peptides bearing this device show fast  $N \rightarrow S$  shift kinetics under NCL conditions around neutral pH, likely arising from a biomimetic intein-like intramolecular catalysis mechanism. The synthesis of these novel crypto-thioesters is straightforward and does not require an additional step prior to NCL. We believe that this accessible and robust Fmoc-based thioesterification device provides a significant advance to chemical protein synthesis.

## Results and discussion

We based our design on  $N$ -alkyl cysteine<sup>7c</sup> as a scaffold for the design of our thioesterification device, as it can be readily assembled in the solid-phase from inexpensive and commercially available materials.<sup>11</sup> However, its  $N \rightarrow S$  shift is usually very slow at neutral pH,<sup>12</sup> preventing its direct use in NCL. We focused our efforts on addressing this challenge by taking inspiration from the mechanism of *in vivo* protein splicing promoted by class 1 inteins. The first step of this process consists of a self-catalyzed  $N \rightarrow S$  acyl shift of an amine into a thioester,<sup>8</sup> where a highly conserved histidine residue plays a pivotal role. This  $N$ -protonated His is believed to catalyze C–N bond scission through polarization and then protonation of the nitrogen leaving group (Fig. 2A).<sup>13</sup> We reasoned that the  $N$ -alkyl group of cysteine could be functionalized by an appropriate intramolecular proton donor able to mimic the catalytic histidine side chain, resulting in a much faster  $N \rightarrow S$  shift at neutral pH. Accordingly, we designed an  $N$ -(2-hydroxybenzyl) cysteine device (Fig. 2B) leading to a hypothetical thermodynamically favored six-membered transition state for intramolecular  $N$ -protonation. The key to this design is a phenol group with a  $pK_a$  close to neutral, as this would result in a low energy cost for the intramolecular proton transfer at pH 7.

We expected an additional advantage from the 2-hydroxybenzyl group: acceleration of  $N$ -acylation *via*  $O$ -acylation followed by the intramolecular  $O \rightarrow N$  shift (Fig. 2C), as in the  $N$ -acylation of 2-hydroxy-5-methoxybenzyl (Hmb) protected amines.<sup>14</sup> This could overcome a common drawback encountered with most  $N \rightarrow S$  acyl shift-based thioesterification devices developed so far, the difficult  $N$ -acylation of

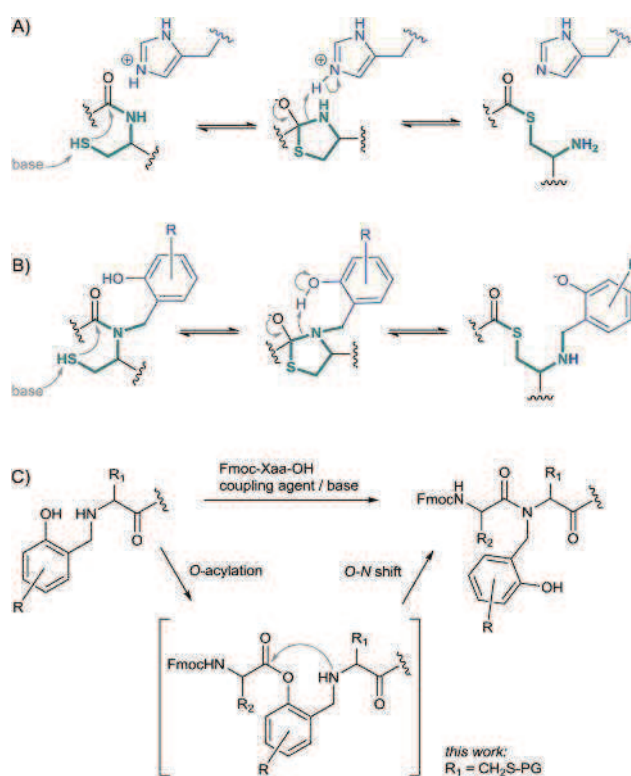


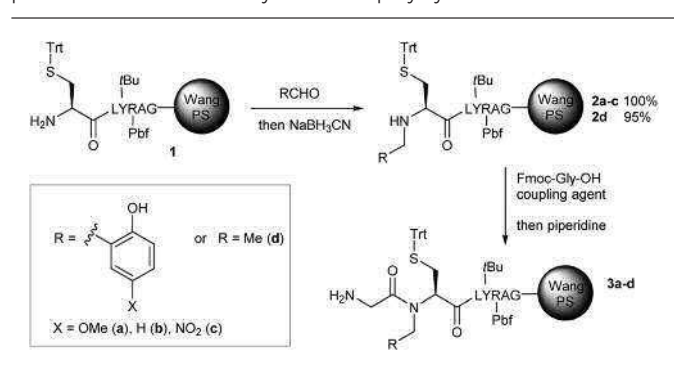
Fig. 2 Rational design of  $N$ -(2-hydroxybenzyl)cysteine thioesterification devices. (A) Mechanism of the intein self-catalyzed  $N \rightarrow S$  shift in water. (B) Bio-inspired putative mechanism for the self-catalysis of the  $N \rightarrow S$  shift in  $N$ -acyl- $N$ -(2-hydroxybenzyl)cysteine peptides in water. (C) Mechanism of the intramolecular  $O \rightarrow N$  shift in organic solvents shared by  $N$ -Hmb peptides and the new  $N$ -(2-hydroxybenzyl) cysteine thioesterification devices. PG: protecting group.

a secondary amine,<sup>7g,h,15</sup> to introduce the first residue of the peptide sequence.

### Optimization of the hydroxybenzyl group

Three different *N*-2-hydroxybenzyl groups were introduced on a solid supported cysteine derivative through an automated reductive amination protocol (compounds **2a–c**) and were compared with a reference *N*-ethyl cysteine device (**2d**). For this preliminary study, a reporter peptide (LYRAG-NH<sub>2</sub>) was introduced C-terminal to cysteine to facilitate the analysis using LC-MS after TFA-mediated deprotection and cleavage from the resin. Excellent results were obtained with the 2-hydroxy-5-nitrobenzyl (Hnb) group,<sup>16</sup> giving quantitative *N*-acylation yields for a Gly residue under mild amide coupling conditions (HBTU/HOBt) (Table 1, entry 5), while the ethyl-substituted compound (Table 1, entry 7) was unreactive. Hmb or unsubstituted 2-hydroxybenzyl groups (Table 1, entries 1 and 3) showed only moderate assistance. These results are in line with the literature:<sup>16</sup> Alewood and collaborators already observed the much superior acyl transfer efficiencies of 2,6- and 2,5-Hnb compared to Hmb. Moreover, the pK<sub>a</sub> of the hydroxyl group of Hnb is close to 7 (~6.5, see ESI, Fig. S11†), in line with our requirements discussed above for efficient biomimetic catalysis. We accordingly selected *N*-Hnb-Cys as our preferred scaffold.

**Table 1** Screening of different *N*-(2-hydroxybenzyl) groups for their ability to enhance the *N*-acylation yield of a solid-supported *S*-trityl-protected *N*-substituted cysteine. PS: polystyrene



Entry	Resin	Coupling agent <sup>a</sup>	Product	Yield <sup>b</sup> [%]
1	<b>2a</b>	HBTU/HOBt	<b>3a</b>	<5 <sup>c</sup>
2	<b>2a</b>	HATU	<b>3a</b>	45 <sup>c</sup>
3	<b>2b</b>	HBTU/HOBt	<b>3b</b>	11
4	<b>2b</b>	HATU	<b>3b</b>	12
5	<b>2c</b>	HBTU/HOBt	<b>3c</b>	>99
6	<b>2c</b>	HATU	<b>3c</b>	>99
7	<b>2d</b>	HBTU/HOBt	<b>3d</b>	<1
8	<b>2d</b>	HATU	<b>3d</b>	12
9	<b>2d</b>	PyBrop	<b>3d</b>	1
10	<b>2d</b>	HATU/DMAP <sup>d</sup>	<b>3d</b>	13

<sup>a</sup> All reactions were conducted for 2 h, in DMF, using 10 equiv. Fmoc-Xaa-OH, 9.5 equiv. coupling agent and 20 equiv. iPr<sub>2</sub>NEt. <sup>b</sup> Yields determined after TFA-mediated cleavage of the peptide resin using HPLC analysis of the amine-to-amide conversion. <sup>c</sup> Estimated yields due to closely eluting HPLC peaks. <sup>d</sup> 0.1 equiv. DMAP.

### Optimization of the thiol protecting group of cysteine and linkage to the resin

When exploring the scope of peptides bearing our *N*-Hnb-Cys thioesterification device, we observed an instability upon purification and storage due to a premature spontaneous *N* → *S* shift. Fortunately, the protection of the cysteine thiol as an *S*-*S*tBu disulfide, stable to Fmoc-SPPS and readily removed under NCL conditions,<sup>17</sup> resolved this problem (see ESI, pS24–S29†).<sup>18</sup> As the use of a reporter peptide was no longer needed at this point, we looked at introducing Cys(*S*tBu) as either the  $\alpha$ -acid or the amide, using a Wang-type or a Rink's amide linker, respectively. Disappointingly, in the former case we were confronted with Cys epimerization<sup>19</sup> and  $\beta$ -elimination, followed by the addition of piperidine.<sup>20</sup> In the latter case, we observed a TFA-catalyzed hydrolysis<sup>21</sup> reaction, leading to a mixture of amide-(7) and acid-terminated (8) compounds (see ESI pS27–S29† for details). Simple incorporation of a Gly spacer between the Rink's linker and the device solved these problems. Preliminary evaluation of NCL reactions with the model cysteinyl peptide **6** showed that this additional residue (compound **9**) did not affect the NCL kinetics compared to cysteine  $\alpha$ -amide **7**, the reaction being even slightly faster than for the  $\alpha$ -acid **8** (ESI, Fig. S57†). As expected, cysteine deprotection was very fast, showing the complete removal of the *S*tBu group within minutes.

### Study of the *N*-acylation of the device

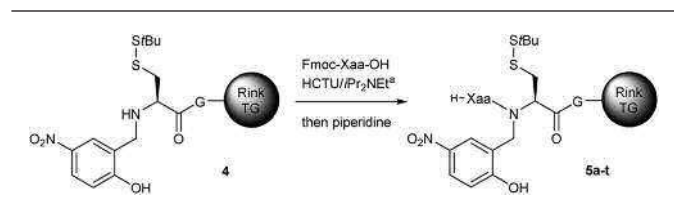
*N*-Acylation of the optimized device was then examined for the twenty proteogenic amino acids, using a standard SPPS protocol (Table 2). In most cases the coupling yields were good to excellent. No significant epimerization was observed, except for Fmoc-Cys(Trt) (7%), which is well known for its propensity for racemization.<sup>22</sup> Six residues showed modest *N*-acylation yields (<50%). In these difficult cases, we looked at further optimizing the coupling, avoiding the use of sensitive or costly reagents, and keeping in mind an easy-to-automate procedure. Three successive couplings of Fmoc-Ser(*t*Bu) resulted in an excellent yield (85%, entry 16). The five-fold coupling of the more demanding Fmoc-Val led to a good 65% yield (entry 21). Alternatively, single coupling under microwave heating at 70 °C led to an excellent 92% yield for Fmoc-Val (entry 22) and 78% for Fmoc-Ile (entry 24). Importantly, the unreacted secondary amine could be quantitatively capped by acetylation, giving a byproduct that can be readily removed during the peptide precipitation step following TFA-based cleavage from the resin.<sup>23</sup>

### Mechanistic insights

Next, we wanted to assess the validity of our intein-inspired design hypothesis that should result in the self-catalysis of the *N* → *S* shift at neutral pH through intramolecular protonation. Here, we synthesized a model peptide equipped with a 2-methoxy-5-nitrobenzylcysteine device (**12**), designed to inhibit the acid-base properties of the phenol group while not being expected to lead to major changes in terms of steric hindrance



**Table 2** Acylation of the resin **4** with the 20 different protected proteogenic amino acids under automated coupling conditions. TG: Tentagel



Entry	Xaa	Product	Nb of couplings <sup>b</sup>	Yield <sup>c</sup> [%]
1	Gly	<b>5a</b>	1	97
2	Asp(OtBu)	<b>5b</b>	1	94
3	Ala	<b>5c</b>	1	90
4	Glu(OtBu)	<b>5d</b>	1	89 <sup>d</sup>
5	Cys(Trt)	<b>5e</b>	1	86 <sup>e</sup>
6	Met	<b>5f</b>	1	85
7	Asn(Trt)	<b>5g</b>	1	80 <sup>f</sup>
8	His(Trt)	<b>5h</b>	1	77
9	Phe	<b>5i</b>	1	77
10	Gln(Trt)	<b>5j</b>	1	72
11	Leu	<b>5k</b>	1	71
12	Trp(Boc)	<b>5l</b>	1	69
13	Arg(Pbf)	<b>5m</b>	1	67
14	Tyr(tBu)	<b>5n</b>	1	58
15	Ser(tBu)	<b>5o</b>	1	49
16	Ser(tBu)	<b>5o</b>	3	85
17	Lys(Boc)	<b>5p</b>	1	45
18	Thr(tBu)	<b>5q</b>	1	19
19	Pro	<b>5r</b>	1	19
20	Val	<b>5s</b>	1	16
21	Val	<b>5s</b>	5	65
22	Val	<b>5s</b>	1	92 <sup>g</sup>
23	Ile	<b>5t</b>	1	12
24	Ile	<b>5t</b>	1	78 <sup>g</sup>

<sup>a</sup> Reactions were conducted at a 0.025 mmol scale using a Prelude synthesizer at RT for 30 min in NMP/DMF 5 : 1, using 10 equiv. Fmoc-Xaa-OH, 9.5 equiv. HCTU and 20 equiv. *i*Pr<sub>2</sub>NEt. <sup>b</sup> When indicated, the coupling step was repeated under the same conditions. <sup>c</sup> Yields determined after TFA-mediated cleavage of the peptidyl resin and subsequent HPLC analysis of the amine-to-amide conversion. <sup>d</sup> Concomitant formation of pyroglutamate during TFA cleavage. <sup>e</sup> 93 : 7 mixture of *L*-Cys/*D*-Cys. <sup>f</sup> Incomplete deprotection of the Trt group. <sup>g</sup> Coupling at 70 °C.

compared to Hnb (peptide **9**). We were delighted to see that the hydroxyl group very efficiently catalyzes the *N* → *S* acyl shift: >50-fold rate enhancement of NCL was measured for **9** compared to **12** under the NCL conditions (Fig. 3), supporting our hypothesis.<sup>24</sup>

Moreover, the ligation of **9** at different pH values showed the *N* → *S* acyl shift/NCL process to be fastest at pH ~ 6.5, and slightly slower at lower pH (ESI, Fig. S67†). This fits with the p*K*<sub>a</sub> of the Hnb group (6.5), suggesting the need for a subtle balance between the phenol group protonation and the thiol group deprotonation (cysteine p*K*<sub>a</sub> ~ 8.5) in influencing the *N* → *S* shift. To eliminate a possible *N* → *O* shift mechanism, the thiol was protected with a stable acetamidomethyl (Acm) group (**13**). This NCL-unreactive compound confirms the crucial role of the SH group and provides further data to suggest an intramolecular *N* → *S* shift mechanism. Critically, this latter finding

also highlights the potential of the *N*-Hnb-Cys thioester precursors for successive ligations in the *N*-terminus-to-*C*-terminus (*N*-to-*C*) direction. We are currently evaluating latent crypto-thioesters that could be activated by deprotection of a NCL-stable cysteine protecting group, as has been exploited for related *N*-methyl Cys,<sup>11,25</sup> CPE,<sup>26</sup> SEA<sup>17</sup> and Sealide<sup>27</sup> thioester precursors.

### Study of the NCL with model peptides

Next, we turned our attention to the effectiveness of our new thioesterification device in NCL reactions for a range of thioesters bearing different amino acids *C*-terminal to the *N*-Hnb-Cys device (Fig. 3). Note that we used relatively dilute conditions for the peptide reactants (~1 mM) and low concentrations of the 4-mercaptophenylacetic acid catalyst (MPAA, 25 mM).<sup>28</sup> Relatively fast NCL kinetics were observed using Ala (peptide **9**) and Ser (**10**) crypto-thioesters, as the reactions were complete after 24 h at 37 °C. For comparison with the NCL kinetics data available in the literature, we determined the apparent second order kinetic constants, with results in the 0.03–0.06 M<sup>-1</sup> s<sup>-1</sup> range, in line with the reported values for standard NCL with preformed thioesters.<sup>29</sup> These data validate the rapid *N* → *S* shift promoted by our device. Ligation at valine (crypto-thioester **11**) was 10 times slower, as we had anticipated from the steric hindrance of its side chain.<sup>30</sup> However, increasing the MPAA concentration to 300 mM, lowering the pH close to the optimum (6.6) and heating to 50 °C led to kinetics comparable to Ala and Ser crypto-thioesters. The latter result suggests that the method could be applied in routine NCL even for peptide sequences known to be kinetically demanding.

Potential epimerization during the *N* → *S*-shift/NCL process was rigorously examined by synthesizing a *D*-Ala-containing product (**15**) as an HPLC standard. Only trace epimerization (<0.4%) was detected in the crude NCL mixture starting from the *L*-Ala crypto-thioester **9**. We also examined the hydrolysis of the crypto-thioesters into the corresponding acids. All reactions were clean and did not show significant amounts of hydrolysis (<4% for **9** after 24 h and using excess crypto-thioester).

### Application of the methodology to two cysteine-rich peptides

Importantly, our new method was applied to the synthesis of two long naturally-occurring cysteine-rich peptide sequences, MT7 and Cg-BigDef1. Muscarinic toxin 7 (MT7), 65 residues, was isolated from green mamba venom.<sup>31</sup> MT7 naturally forms a three-finger-fold stabilized by four disulfide bonds. Big defensin 1 (Cg-BigDef1) is a 93-residue host-defense peptide isolated from Japanese oysters and is composed of a hydrophobic *N*-terminal domain and a β-defensin-like cationic *C*-terminal domain containing six cysteine residues involved in three disulfide bridges.<sup>32</sup>

These two peptide sequences were chosen as challenging examples to demonstrate the applicability of our methodology, as their syntheses *via* the NCL of two segments require long thioesters, 41 and 56 amino acids, respectively. In addition, the *N*-terminal Cg-BigDef1 domain is a “difficult sequence”,<sup>33</sup> *i.e.* hard to synthesize using Fmoc-SPPS.<sup>34</sup> Both of the crypto-

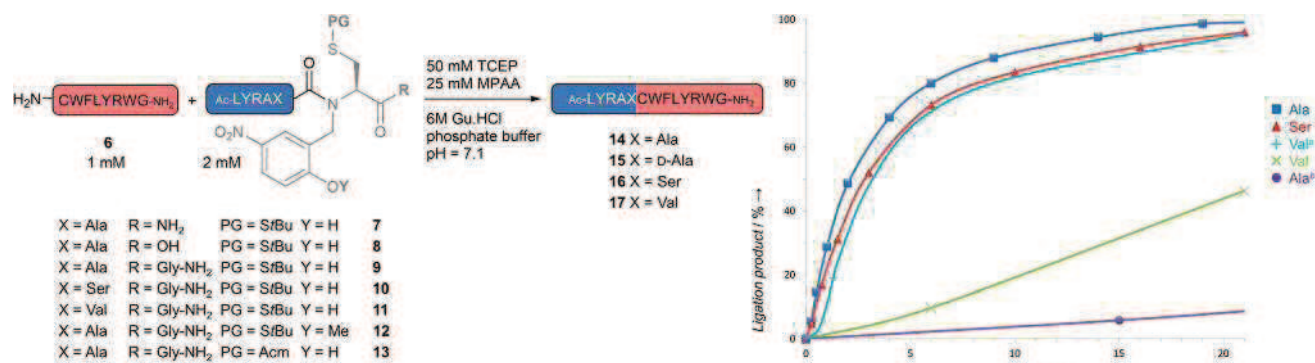


Fig. 3 NCL of the model crypto-thioester peptides 7–11 bearing an *N*-Hnb-cysteine device, and an analog (12) with the hydroxyl group masked as a methyl ether. <sup>a</sup> 300 mM MPAA, pH 6.6, 50 °C. <sup>b</sup> Hnb methyl ether (12).

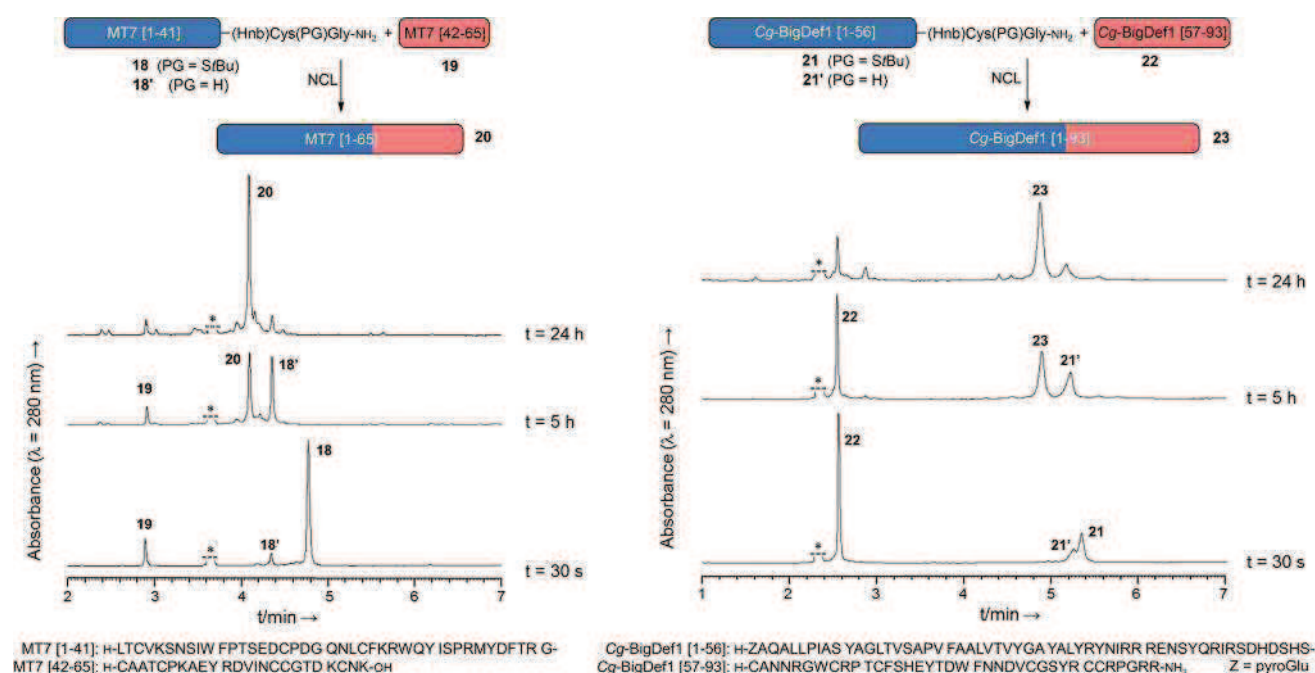


Fig. 4 Application of the *N*-(2-hydroxy-5-nitrobenzyl)cysteine thioesterification device to the NCL-based syntheses of two long and demanding peptide sequences.

thioesters (18 and 21) were synthesized and purified in good yields (13% and 7%, see ESI, pS86 and S93†). The two ligations were complete overnight at 1 and 2 mM concentrations, respectively, and showed clean HPLC profiles (Fig. 4). Finally, pure reduced forms of MT7 (20) and Cg-BigDef1 (23) were isolated after HPLC purification in 14% and 18% yields, respectively<sup>35</sup> (ESI, Fig. S85 and S93†), demonstrating the utility of the *N*-Hnb-Cys method.

## Conclusions

In conclusion, we introduced a novel C-terminal thioesterification device, *N*-Hnb-Cys(*S**t*Bu), that can rapidly rearrange into a thioester at neutral pH, allowing its use in NCL even under demanding conditions. Fast kinetics likely result from an

internal catalysis mechanism provided by a well-positioned phenol moiety arising from a bio-inspired intein-like design. Synthesis from inexpensive commercially available materials and further peptide elongation are straightforward, and can be fully automated on a peptide synthesizer without requiring a post-synthetic step, thus constituting a major advance in the field. Applicability of the strategy was demonstrated by the synthesis of two long naturally occurring cysteine-rich peptide sequences. Further work is ongoing in our laboratory for its application to other synthetically demanding targets.

## Acknowledgements

We thank Dr Guillaume Gabant, Dr Cyril Colas and the mass spectrometry platforms of CBM and FR2708 federation for the

MS and LC-MS analyses, as well as Hervé Meudal and the CBM NMR platform for the NMR spectra. The Pierre Potier foundation for the chemistry of natural products is gratefully acknowledged for a PhD fellowship for VPT, as well as ANR (JCJC08-0138, SynProt project) and Vaincre la Mucovisidose for financial support. Dr Denis Servent, Dr Gilles Mourier and Dr Pascal Kessler are greatly acknowledged for their expertise and support for the application of the method to MT7, as well as Dr Delphine Destoumieux-Garzón for Cg-BigDef1. Thanks also to Dr Michael Jacobsen and Dr Michael Kay for proof-reading the manuscript.

## Notes and references

- 1 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776–779.
- 2 F. Mende and O. Seitz, *Angew. Chem., Int. Ed.*, 2011, **50**, 1232–1240.
- 3 J. B. Blanco-Canosa and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2008, **47**, 6851–6855.
- 4 G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui and L. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 7645–7649.
- 5 For side reactions observed with the Nbz approach see: S. K. Mahto, C. J. Howard, J. C. Shimko and J. J. Ottesen, *ChemBioChem*, 2011, **12**, 2488–2494. Note that an article about a second generation Nbz linker was published during the preparation of this manuscript: J. B. Blanco-Canosa, B. Nardone, F. Albericio and P. E. Dawson, *J. Am. Chem. Soc.*, 2015, **137**, 7197–7209. For side reactions observed with the hydrazide approach, see: S. K. Mong, A. A. Vinogradov, M. D. Simon and B. L. Pentelute, *ChemBioChem*, 2014, **15**, 721–733.
- 6 For reviews on  $N \rightarrow S$  acyl shift-based thioesterification devices, see: (a) T. Kawakami, *Top. Curr. Chem.*, 2015, **362**, 107–135; (b) D. Macmillan, A. Adams and B. Premdjee, *Isr. J. Chem.*, 2011, **51**, 885–889.
- 7 (a) T. Kawakami, M. Sumida, K. Nakamura, T. Vorherr and S. Aimoto, *Tetrahedron Lett.*, 2005, **46**, 8805–8807; (b) F. Nagaike, Y. Onuma, C. Kanazawa, H. Hojo, A. Ueki, Y. Nakahara and Y. Nakahara, *Org. Lett.*, 2006, **8**, 4465–4468; (c) H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara and Y. Nakahara, *Tetrahedron Lett.*, 2007, **48**, 25–28; (d) T. Kawakami and S. Aimoto, *Chem. Lett.*, 2007, **36**, 76–77; (e) S. Tsuda, A. Shigenaga, K. Bando and A. Otaka, *Org. Lett.*, 2009, **11**, 823–826; (f) J. Kang, J. P. Richardson and D. Macmillan, *Chem. Commun.*, 2009, 407–409; (g) N. Ollivier, J. Dheur, R. Mhida, A. Blanpain and O. Melnyk, *Org. Lett.*, 2010, **12**, 5238–5241; (h) W. Hou, X. Zhang, F. Li and C.-F. Liu, *Org. Lett.*, 2011, **13**, 386–389; (i) J.-S. Zheng, H.-N. Chang, F.-L. Wang and L. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 11080–11083; (j) F. Burlina, G. Papageorgiou, C. Morris, P. D. White and J. Offer, *Chem. Sci.*, 2014, **5**, 766–770.
- 8 For recent reviews on inteins, see: (a) G. Volkmann and H. D. Mootz, *Cell. Mol. Life Sci.*, 2013, **70**, 1185–1206; (b) K. V. Mills, M. A. Johnson and F. B. Perler, *J. Biol. Chem.*, 2014, **289**, 14498–14505; (c) H. H. Shah and T. W. Muir, *Chem. Sci.*, 2014, **5**, 446–461.
- 9 K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto and A. Otaka, *ChemBioChem*, 2011, **12**, 1840–1844.
- 10 (a) Y. Asahina, K. Nabeshima and H. Hojo, *Tetrahedron Lett.*, 2015, **56**, 1370–1373; (b) Y. Asahina, S. Komiya, A. Ohagi, R. Fujimoto, H. Tamagaki, K. Nakagawa, T. Sato, S. Akira, T. Takao, A. Ishii, Y. Nakahara and H. Hojo, *Angew. Chem., Int. Ed.*, 2015, **54**, 8226–8230.
- 11 L. A. Erlich, K. S. Kumar, M. Haj-Yahya, P. E. Dawson and A. Brik, *Org. Biomol. Chem.*, 2010, **8**, 2392–2396.
- 12 (a) K. S. A. Kumar, L. Spasser, L. A. Erlich, S. N. Bavikar and A. Brik, *Angew. Chem., Int. Ed.*, 2010, **49**, 9126–9131; (b) M. Haj-Yahya, B. Fauvet, Y. Herman-Bachinsky, M. Hejjaoui, S. N. Bavikar, S. V. Karthikeyan, A. Ciechanover, H. A. Lashuel and A. Brik, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 17726–17731; (c) Y. Ruff, V. Garavini and N. Giuseppone, *J. Am. Chem. Soc.*, 2014, **136**, 6333–6339.
- 13 J. Binschik and H. D. Mootz, *Angew. Chem., Int. Ed.*, 2013, **52**, 4260–4264.
- 14 T. Johnson, M. Quibell, D. Owen and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1993, 369–372.
- 15 K. Sakamoto, K. Sato, A. Shigenaga, K. Tsuji, H. Hibino, Y. Nishiuchi and A. Otaka, *J. Org. Chem.*, 2012, **77**, 6948–6958.
- 16 L. P. Miranda, W. D. Meutermans, M. L. Smythe and P. F. Alewood, *J. Org. Chem.*, 2000, **65**, 5460–5468.
- 17 See for example: L. Raibaut, H. Adihou, R. Desmet, A. F. Delmas, V. Aucagne and O. Melnyk, *Chem. Sci.*, 2013, **4**, 4061–4066.
- 18 Instability of thioesterification devices upon handling and purification is a known problem that can be solved by thiol protection to inhibit a premature  $N \rightarrow S$  shift. For example, SEA crypto-thioesters<sup>7g</sup> can be post-synthetically oxidized to give stable compounds. See also ref. 10a.
- 19 Y. Tawaragi, K. Fuchimura, S. Tanaka, N. Minamino, K. Kangawa and H. Matsuo, *Biochem. Biophys. Res. Commun.*, 1991, **175**, 645–651.
- 20 J. Lukszo, D. Patterson, F. Albericio and S. A. Kates, *Let. Pept. Sci.*, 1996, **3**, 157–166.
- 21 (a) C. J. Creighton, T. T. Romoff, J. H. Bu and M. Goodman, *J. Am. Chem. Soc.*, 1999, **121**, 6786–6791; (b) M. Teixidó, F. Albericio and E. Giralt, *J. Pept. Res.*, 2005, **65**, 153–166.
- 22 Y. Han, F. Albericio and G. Barany, *J. Org. Chem.*, 1997, **62**, 4307–4312.
- 23 Note that a simple adjustment of the initial amount of resin allows the SPPS of the crypto-thioester sequence to be conducted at the desired scale.
- 24 The hydroxyl group could assist the reaction *via* a different mechanism to the one we postulated, such as stabilization of the oxyanion moiety of the tetrahedral intermediate. Note that the activation of the amide by hydrogen bonding between the phenol group and the carbonyl oxygen seems unlikely: NMR analysis of a small model peptide showed no significant differences in the *cis/trans* ratio between the



- hydroxy and methoxy compounds (see ESI pS14–S23†). A hydrogen bond would be expected to stabilize the *cis* form and thus modify the ratio. DFT studies are planned to help better understand the mechanism.
- 25 K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon and A. Brik, *Angew. Chem., Int. Ed.*, 2011, **50**, 6137–6141.
- 26 T. Kawakami and S. Aimoto, *Tetrahedron Lett.*, 2007, **48**, 1903–1905.
- 27 K. Sato, A. Shigenaga, K. Kitakaze, K. Sakamoto, D. Tsuji, K. Itoh and A. Otaka, *Angew. Chem., Int. Ed.*, 2013, **52**, 7855–7859.
- 28 E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640–6646.
- 29 (a) S. B. Pollock and S. B. H. Kent, *Chem. Commun.*, 2011, **47**, 2342–2344; (b) F. Saito, H. Noda and J. W. Bode, *ACS Chem. Biol.*, 2015, **10**, 1026–1033.
- 30 T. M. Hackeng, J. H. Griffin and P. E. Dawson, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 10068–10073.
- 31 G. Mourier, S. Dutertre, C. Fruchart-Gaillard, A. Ménez and D. Servent, *Mol. Pharmacol.*, 2003, **63**, 26–35.
- 32 R. D. Rosa, A. Santini, J. Fievet, P. Bulet, D. Destoumieux-Garzón and E. Bachère, *PLoS One*, 2011, **6**, e25594.
- 33 I. Coin, M. Beyermann and M. Bienert, *Nat. Protoc.*, 2007, **2**, 3247–3256.
- 34 Standard Fmoc-SPPS of Cg-BigDef [1–42] C-terminal  $\alpha$ -amide gave extremely poor results and low elongation yields, probably arising from its difficult hydrophobic sequence. Considering these results, we optimized the synthesis of the cryptothioester **21** by incorporating multiple coupling steps.
- 35 These isolated yields can appear rather low considering the excellent conversion yield, as determined from the HPLC chromatograms of the crude NCL mixtures, see Fig. 4. We attribute this loss of material to the non-optimized workup and HPLC purification protocols.



## Références bibliographiques

1. Mende F.; Seitz O., 9-Fluorenylmethoxycarbonyl-based solid-phase synthesis of peptide  $\alpha$ -thioesters. *Angew. Chem. Int. Ed.*, **2011**, *50*, 1232-1240.
2. 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*, 7645-7649.
3. Blanco-Canosa J. B.; Dawson P. E., An Efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem. Int. Ed.*, **2008**, *47*, 6851-6855.
4. 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*, 7197-7209.
5. Kawakami T., Peptide thioester formation via an intramolecular *N* to *S* acyl shift for peptide ligation. *Top. Curr. Chem.*, **2015**, *362*, 102-135.
6. Hojo H.; Onuma Y.; Akimoto Y.; Nakahara Y.; Nakahara Y., *N*-Alkyl cysteine-assisted thioesterification of peptides. *Tetrahedron Lett.*, **2007**, *48*, 25-28.
7. Asahuna Y.; Nabeshima K.; Hojo H., Peptidyl *N*-alkylcysteine as a peptide thioester surrogate in the native chemical ligation. *Tetrahedron Lett.*, **2015**, *56*, 1370-1373.
8. Mills K. V.; Johnson M. A.; Perler F. B., Protein splicing: how inteins escape from precursor proteins. *J. Biol. Chem.*, **2014**, *289*, 14498-14505.
9. Binschik J.; Mootz H. D., Chemical bypass of intein-catalyzed *N*-*S* acyl shift in protein splicing. *Angew. Chem. Int. Ed.*, **2013**, *52*, 4260-4264.
10. Sato K.; Shigenaga A.; Tsuji K.; Tsuda S.; Sumikawa Y.; Sakamoto K.; Otaka A., *N*-sulfanylethylanilide peptide as a crypto-thioester peptide. *ChemBioChem*, **2011**, *12*, 1840-1844.
11. Johnson T.; Quibell M.; Owen D.; Sheppard R. C., A reversible protecting group for the amide bond in peptides. Use in the synthesis of 'difficult sequences'. *J. Chem. Soc., Chem. Commun.*, **1993**, *137*, 369-372.
12. Miranda L. P.; Meutermans W. D.; Smythe M. L.; Alewood P. F., An activated *O*→*N* acyl transfer auxiliary: efficient amide-backbone substitution of hindered "difficult" peptides. *J. Org. Chem.*, **2000**, *65*, 5460-5468.
13. Taichi M.; Hemu X.; Qiu Y.; Tam J. P., A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift. *Org. Lett.*, **2013**, *15*, 2620-2623.
14. Ollivier N.; Dheur J.; Mhidia R.; Blanpain A.; Melnyk O., *Bis*(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.*, **2010**, *12*, 5238-5241.
15. Weber U.; Hartter P., *S*-Alkylmercapto groups for protection of the SH-function of cysteine. I. Synthesis and stability of some *S*-(alkylmercapto) cysteines. *Hoppe-Seyler's Z. Physiol. Chem.*, **1970**, *351*, 1384-1388.

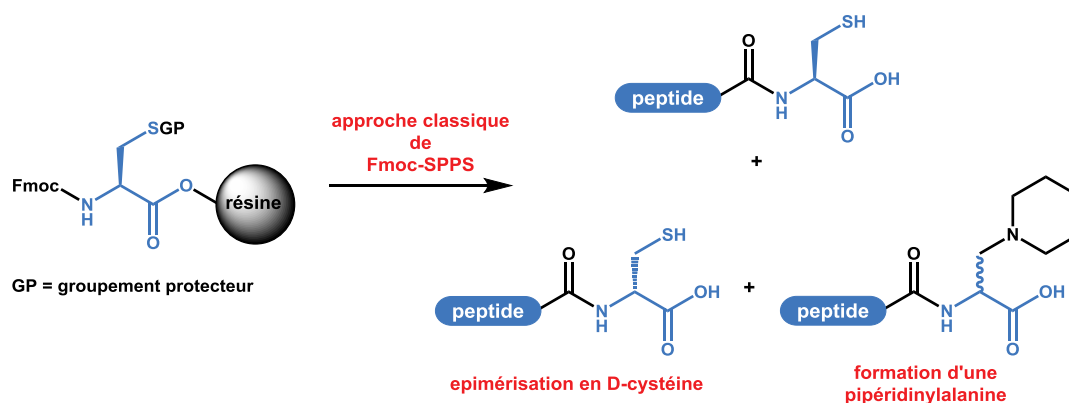
16. Tawaragi Y.; Fuchimara K.; Tanaka S.; Minamino N.; Kangawa K.; Matsuo H., Gene and precursor structures of human C-type natriuretic peptide. *Biochem. Biophys. Res. Commun.*, **1991**, *175*, 645-651.
17. Lukszo J.; Patterson D.; Albericio F.; Kates S. A., 3-(1-Piperidiny)alanine formation during the preparation of C-terminal cysteine peptides with the Fmoc/tBu strategy. *Lett. Pept. Sci.*, **1996**, *3*, 157-166.
18. Urban J.; Vaisar T.; Shen R.; Lee M. S., Lability of *N*-alkylated peptides towards TFA cleavage. *Int. J. Pept. Prot. Res.*, **1996**, *47*, 182-189.
19. Teixidó M.; Albericio F.; Giralt E., Solid-phase synthesis and characterization of *N*-methyl-rich peptides. *J. Pept. Res.*, **2005**, *65*, 153-166.
20. Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., Expanding the scope of *N* → *S* acyl transfer in native peptide sequences. *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.
21. Mourier G.; Dutertre S.; Fruchart-Gaillard C.; Ménez A.; Servent D., Chemical synthesis of MT1 and MT7 muscarinic toxins: critical role of Arg-34 in their interaction with M1 muscarinic receptor. *Mol. Pharmacol.*, **2003**, *63*, 26-35.
22. Rosa R. D.; Santini A.; Fievet J.; Bulet P.; Destoumieux-Garzón D.; Bachère E., Big defensins, a diverse family of antimicrobial peptides that follows different patterns of expression in hemocytes of the oyster *Crassostrea gigas*. *PLoS One*, **2011**, *6*, e25594.

# CHAPITRE 3



## Chapitre 3 : Stratégie pour la synthèse de peptides comportant une cystéine C-terminale par ligation chimique native

Dans la continuité du chapitre 2, nous nous sommes intéressés à l'application de notre méthode simple de synthèse de peptides crypto-thioester à la résolution d'une limitation de la synthèse peptidique par voie chimique : la synthèse de peptides contenant une cystéine C-terminale. Les cystéines C-terminales sont fréquentes dans les séquences de peptides naturels riches en ponts disulfure démontrant des propriétés biologiques intéressantes (par exemple, dans la somatostatine,<sup>1</sup> les conotoxines,<sup>2</sup> ou encore la crotalphine<sup>3</sup>). Néanmoins, la synthèse de telles cibles par Fmoc-SPPS classique est très limitée en raison de réactions secondaires dues à l'acidité prononcée du proton  $\alpha$  de la cystéine greffée au support solide par un lien ester. D'une part, la cystéine C-terminale peut être épimérisée lors des traitements répétés à l'aide de pipéridine utilisée pour éliminer les groupements Fmoc;<sup>4</sup> d'autre part, la  $\beta$ -élimination du groupement thiol catalysée par la pipéridine suivie de l'addition de cette amine sur la déhydroalanine formée conduit à l'obtention d'une 3-(N-pipéridinyl)alanine C-terminale (figure 1).<sup>5</sup>



**Figure 1 :** Synthèse d'un peptide contenant une cystéine C-terminale par Fmoc-SPPS

Plusieurs stratégies ont été proposées pour résoudre ce problème dont, par exemple, l'utilisation d'une résine fonctionnalisée par un bras de type 2-chloro-trityle,<sup>4</sup> ou encore

<sup>1</sup> Brazeau P.; Vale N.; Burgus R.; Ling N.; Butcher M.; Rivier, J.; Guillemin R., *Science*, **1973**, 179, 77–79.

<sup>2</sup> Halai, R.; Craik, D. J., *Nat. Prod. Rep.*, **2009**, 26, 526–536.

<sup>3</sup> Konno K.; Picolo G.; Gutierrez V. P.; Brigatte P.; Zambelli V. O.; Camargo A. C. M.; Cury Y., *Peptides*, **2008**, 29, 1293–1304.

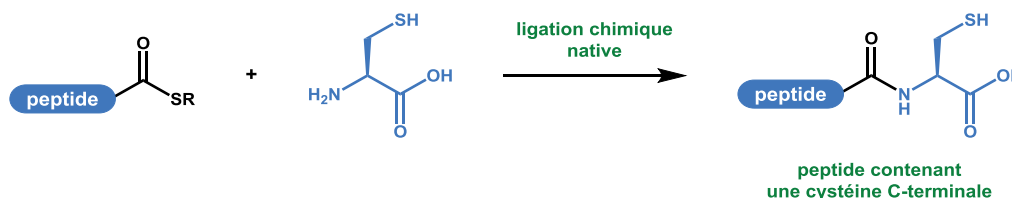
<sup>4</sup> Fujiwara Y.; Akaji K.; Kiso Y., *Chem. Pharm. Bull.*, **1994**, 42, 724-726.

<sup>5</sup> Lukszo J.; Patterson D.; Albericio F.; Kates S. A., *Lett. Pept. Sci.*, **1996**, 3, 157-166.

l'utilisation d'un dérivé de la cystéine de type orthoester.<sup>6</sup> Dans ce dernier cas, la synthèse laborieuse du *building block* cystéine orthoester limite l'utilisation de cette stratégie.

L'équipe dont je fais partie a récemment rencontré ce problème lors de la synthèse de la mini-protéine naturelle AhPDF1.1b, une défensine de 51 acides aminés issue de la plante *Arabidopsis halleri*.<sup>7</sup> Malgré l'utilisation d'un bras de type trityle, plus de 30 % d'épimérisation ont été observés dans le mélange brut, compliquant ainsi la purification du peptide cible et conduisant à son obtention avec un rendement global faible de 7 %.

Pour répondre à ce problème, la stratégie générale que nous proposons consiste en l'introduction post-SPPS de la cystéine C-terminale par une NCL avec un peptide thioester (figure 2). En effet, dans des conditions de NCL, la fonction C-terminale de la cystéine est sous forme de carboxylate, ce qui diminue fortement la labilité du proton  $\alpha$ , évitant ainsi l'épimérisation de cet acide aminé.



**Figure 2 :** Synthèse d'un peptide contenant une cystéine C-terminale par NCL

Vue la simplicité de sa mise en œuvre, nous avons appliqué notre méthode de synthèse de peptides crypto-thioester par utilisation du dispositif de thioestérification *N*-Hnb-Cys à cette approche pour la synthèse de la forme réduite de AhPDF1.1b.

La séquence de AhPDF1.1b contient un motif Pro-Cys C-terminal. Comme nous l'avons vu, cette jonction est particulièrement difficile à former par NCL car l'utilisation de peptide prolyl-thioester conduit à des cinétiques réactionnelles très lentes.<sup>8,9</sup> De plus, une réaction secondaire provenant de la délétion du motif Xaa-Pro d'un tel peptide thioester par formation de dicétopipérazine a récemment été décrite.<sup>10,11</sup> L'application de notre stratégie à ce cas difficile représente un exemple de choix pour poursuivre l'étude du potentiel de la méthode des peptides *N*-Hnb-Cys en NCL.

<sup>6</sup> Huang Z.; Derksen D. J.; Vederas J. C., *Org. Lett.*, **2010**, *12*, 2282-2285.

<sup>7</sup> Meindre F.; Lelièvre D.; Loth K.; Mith O.; Aucagne V.; Berthomieu P.; Marquès L.; Delmas A. F.; Landon C.; Paquet F., *Biochemistry*, **2014**, *53*, 7745-7754.

<sup>8</sup> Hackeng T. M.; Griffin J. H.; Dawson P. E., *Proc. Natl. Acad. Sci. Usa*, **1999**, *96*, 10068-10073.

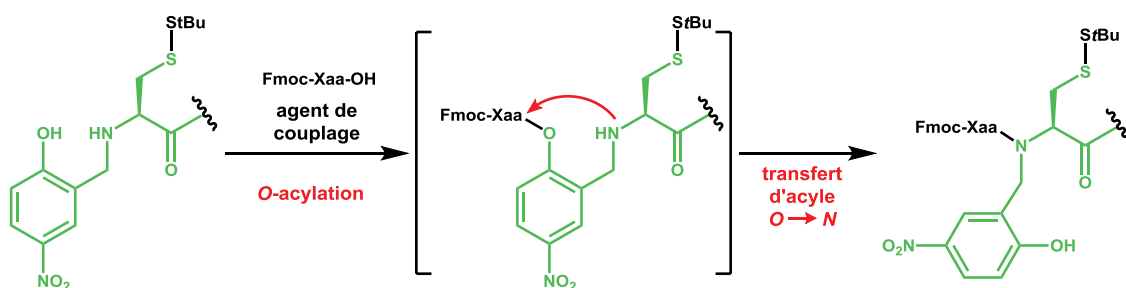
<sup>9</sup> Pollock S. B.; Kent S. B. H., *Chem. Commun.*, **2011**, *47*, 2342-2344.

<sup>10</sup> Raibaut L.; Seeberger P.; Melnyk O., *Org. Lett.*, **2013**, *15*, 5516-5519.

<sup>11</sup> Nakamura T.; Shigenaga A.; Sato K.; Tsuda Y.; Sakamoto K.; Otaka A., *Chem. Commun.*, **2014**, *50*, 58-60.



Par ailleurs, lors de la synthèse de notre dispositif de thioestérification *N*-Hnb-Cys (chapitre 2), nous avons mis en évidence qu'un couplage simple, dans des conditions standards, de la proline ou d'acides aminés encombrés stériquement (tels que Ile, Thr ou Val) sur ce dispositif conduisait à un rendement de *N*-acylation faible, représentant la limitation principale que nous avons identifiée. Nous avons proposé des solutions efficaces pour résoudre ce problème, consistant en des couplages multiples entièrement automatisés ou un couplage sous chauffage (par exemple aux micro-ondes) pour augmenter nettement ce rendement. Néanmoins, nous avons décidé de saisir l'occasion de l'utilisation d'un peptide proline-*N*-Hnb-Cys pour étudier plus en détails le mécanisme du couplage d'un acide aminé sur ce dispositif (rappel : figure 3).

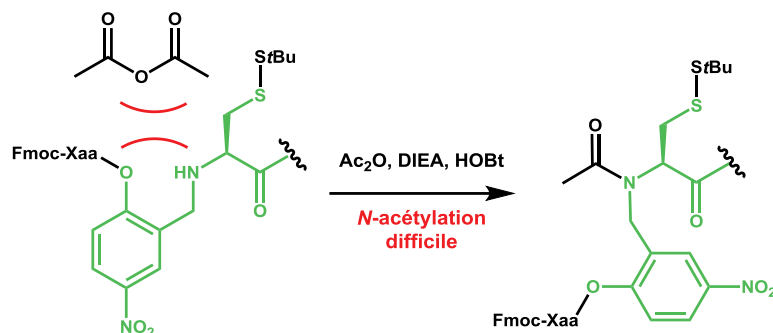


**Figure 3 :** Rappel du mécanisme supposé de la *N*-acylation du dispositif *N*-Hnb-Cys

Le déroulement de réaction sur support solide est difficile à suivre, il est compliqué d'identifier l'étape cinétiquement limitante du couplage de l'acide aminé (la *O*-acylation ou le transfert d'acyle intramoléculaire *O*→*N*). D'après les études existantes de l'acylation de groupement de type 2-hydroxy-5-nitrobenzyle ou 2-hydroxy-6-nitrobenzyle utilisés en tant que *N*-substituants d'acides aminés pour la synthèse de séquences difficiles d'un point de vue de l'encombrement stérique,<sup>12</sup> nous supposons que le transfert d'acyle *O*→*N* était rapide, le groupement Hnb étant un bon groupement partant à cause de l'effet électroattracteur du groupement nitro.

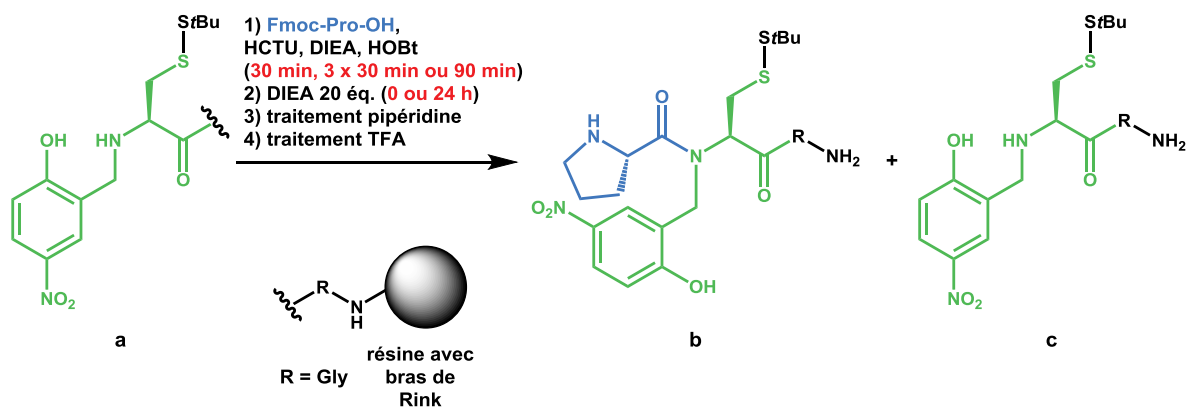
De façon surprenante, nous avons observé que l'étape d'acétylation, pour bloquer la réactivité de l'amine secondaire n'ayant pas réagi, était incomplète; seul un traitement prolongé (jusqu'à 1 h) ou la répétition de cette étape (jusqu'à 10 fois 7 min) permettent d'obtenir une réaction quantitative. Nous proposons une hypothèse pour expliquer ces résultats: le réarrangement par transfert d'acyle *O*→*N* pourrait être lent. L'acétylation de l'amine serait ainsi rendue difficile par l'encombrement stérique dû à la *O*-acylation du groupement Hnb par un Fmoc-acide aminé (figure 4).

<sup>12</sup> Miranda L. P.; Meutermaans W. D.; Smythe M. L.; Alewood P. F., *J. Org. Chem.*, **2000**, 65, 5460-5468.



**Figure 4 :** *N*-acétylation difficile réalisée directement après un couplage d'acide aminé sur le dispositif *N*-Hnb-Cys

A ce jour, nous n'avons jamais pu observer et caractériser cette forme mono-*O*-acylée du motif Hnb-Cys, cette forme ne semblant pas être stable lors du traitement TFA. Pour documenter notre hypothèse, nous avons mis en place des contrôles indirects simples. Nous avons réalisé différents couplages (30 min, 3 × 30 min ou 90 min) de Fmoc-Pro-OH sur le dispositif *N*-Hnb-Cys (composé **a**, figure 5). Après ces couplages, la résine est rincée abondamment pour éliminer l'excès d'acide aminé activé. La résine est ensuite rapidement traitée à la pipéridine puis au TFA en tant que contrôle. La résine est également incubée dans une solution de diisopropyléthylamine (DIEA) durant 24 h avant ces étapes de traitement pipéridine puis TFA. Enfin, le rendement de *N*-acylation est déterminé par intégration HPLC des composés **b** et **c**. Les résultats sont rassemblés dans le tableau 1.



**Figure 5 :** *Etude du couplage de Fmoc-Pro-OH sur le dispositif N-Hnb-Cys*

Couplage (h)	0,5	0,5 × 3	1,5	0,5
Traitement DIEA (h)	0	0	0	24
<i>N</i> -acylation (%)	16	32	33	61

**Tableau 1 :** *Etude du couplage de Fmoc-Pro-OH sur le dispositif N-Hnb-Cys a*

Il est important de noter qu'une simple incubation prolongée en milieu basique, en absence d'acide aminé additionnel ou de réactif de couplage, permet d'augmenter nettement le rendement de *N*-acylation final, ce qui met en évidence le transfert d'acyle *O*→*N* lent comme étape limitante du couplage d'un acide aminé sur le dispositif *N*-Hnb-Cys. De façon anecdotique, nous n'avons pas observé de différence significative de rendement après trois couplages successifs de 30 min ou un seul couplage d'une durée équivalente (1,5 h). Néanmoins, ces faibles rendements montrent que l'étape de *O*-acylation est également lente même si elle n'est pas limitante cinétiquement.

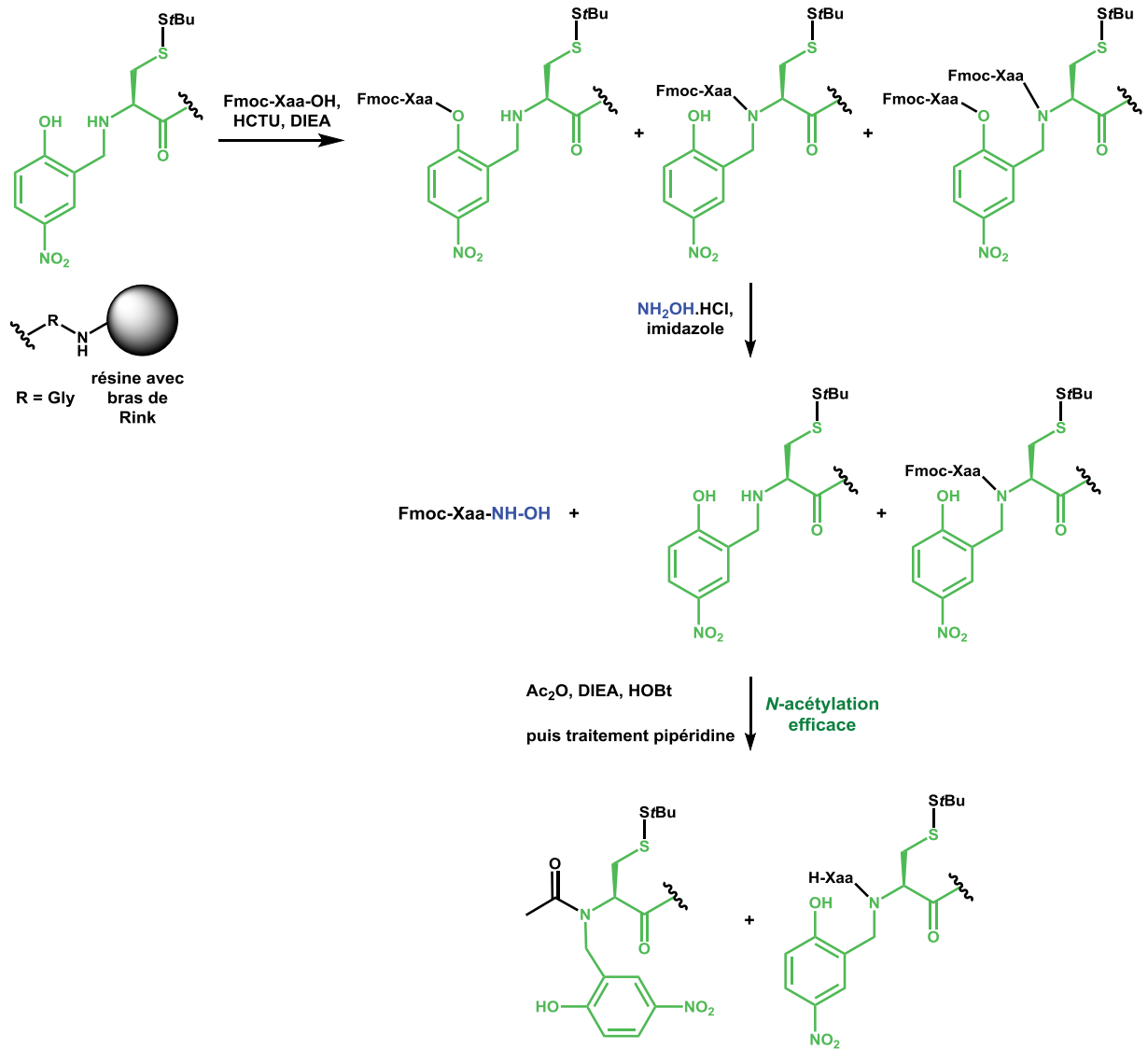
Enfin, dans le but de réaliser une étape d'acétylation quantitative en cas de couplage incomplet sur ce dispositif, il faut déprotéger sélectivement le groupement Hnb *O*-acylé sans déprotéger le groupement Fmoc préalablement à l'étape de *capping*. La solution que nous proposons consiste en un traitement automatisé à l'aide d'un nucléophile dans des conditions peu basiques (figure 6). Nous nous sommes inspirés de conditions développées par Bradley et ses collaborateurs pour éliminer sélectivement un groupement sensible à un nucléophile, le groupement 1-(4,4-diméthyl-2,6-dioxacyclohexylidène)éthyle (Dde), à l'aide d'hydroxylamine et d'imidazole sans réagir avec le groupement Fmoc.<sup>13</sup> Nous avons observé qu'un *capping* réalisé après ce traitement était quantitatif, ce qui va dans le sens de l'hypothèse d'une acétylation difficile du dispositif *N*-Hnb-Cys *O*-acylé (figure 4).

Ce traitement possède l'avantage secondaire de rendre possible une évaluation de l'efficacité de l'installation de l'acide aminé C-terminal ainsi qu'une quantification de l'élongation peptidique par dosage de l'adduit fluorénylméthylpipéridine, libéré après un traitement à base de pipéridine, par spectroscopie UV.

---

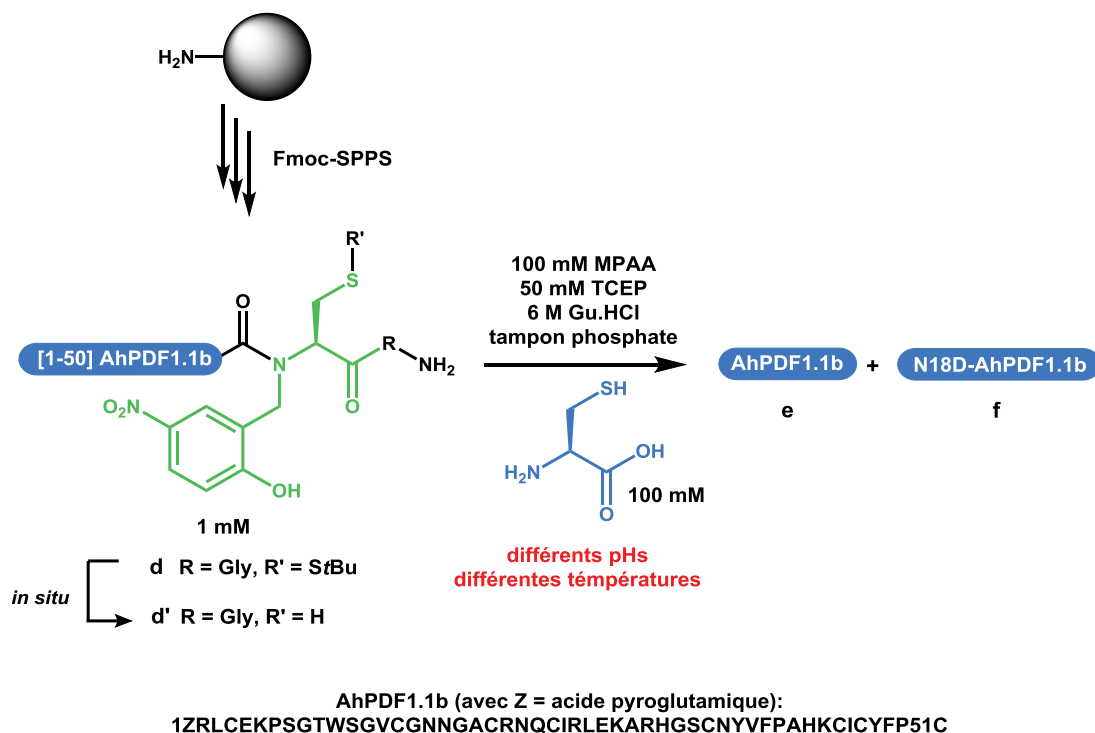
<sup>13</sup> Diaz-Mochón J. J.; Bialy L.; Bradley M., *Org. Lett.*, **2004**, 6, 1127-1129.

### Chapitre 3



**Figure 6 :** Traitement à l'aide d'hydroxylamine préalable à l'étape d'acétylation

Nous avons ensuite synthétisé le crypto-thioester *N*-Hnb-Cys de la séquence [1-50] de AhPDF1.1b (composé **d**) de façon totalement automatisée pour appliquer la stratégie d'installation de la cystéine C-terminale par NCL (figure 7). Notre idée initiale était de réaliser la NCL à partir du mélange brut obtenu après SPPS; néanmoins, pour optimiser les conditions réactionnelles, nous avons réalisé les essais de NCL à partir du composé **d** purifié par HPLC. Toutes les NCLs (optimisation des conditions et application au mélange brut) ont été réalisées par Dominique Lelièvre.



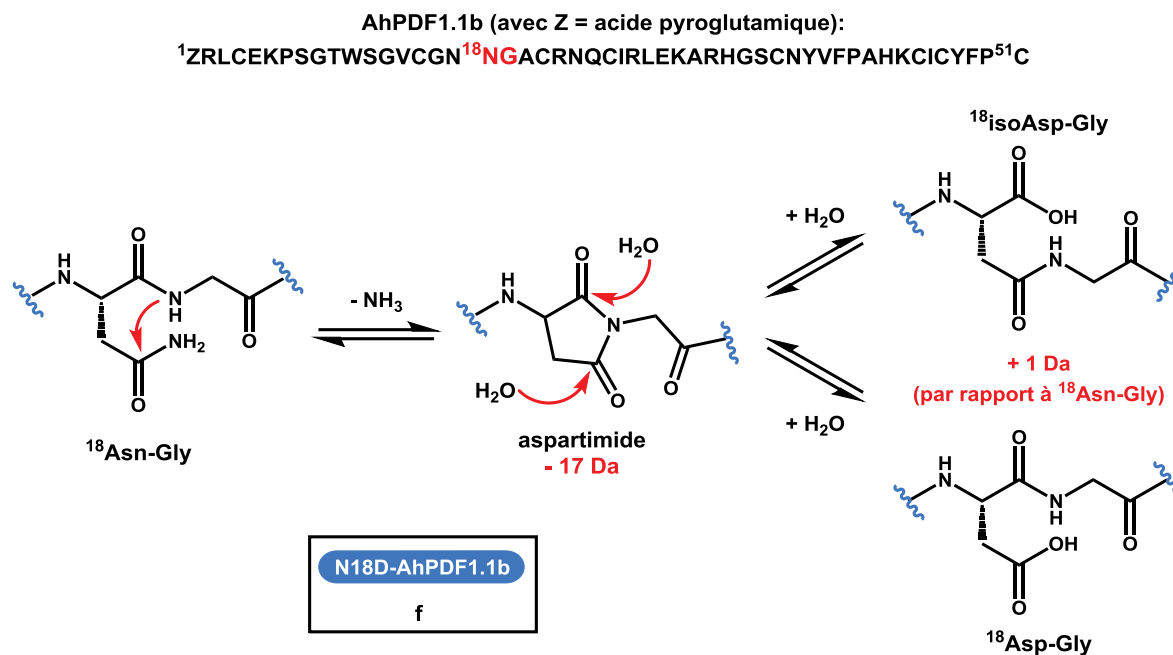
**Figure 7 :** Synthèse de la forme réduite d'AhPDF1.1b **e** à partir du peptide *N*-Hnb-Cys **d**

Les conditions initialement utilisées (20 °C et pH 7,2) conduisent à une NCL très lente, comme cela était attendu avec l'utilisation d'un peptide prolyl-thioester. De plus, nous avons observé la formation d'un produit secondaire (composé **f**, figure 7) lors des analyses de LC/HRMS, à hauteur de 5 % après 24 h de réaction (*ratio* quantifié par intégration du chromatogramme d'ion extrait pour les composés **e** et **f**), montrant une masse de + 1 Da, par rapport au composé **e** attendu. Ce produit a été attribué à la formation d'un motif aspartimide à partir du site <sup>18</sup>Asn-Gly, suivie de son hydrolyse, conduisant à l'obtention d'un motif Asp-Gly et/ou isoAsp-Gly (figure 8).<sup>14,15</sup> Il a été démontré que le site Asn-Gly est particulièrement enclin à

<sup>14</sup> Clarke S., *Int. J. Pept. Prot. Res.*, **1987**, 30, 808-821.

<sup>15</sup> Stephenson R. C.; Clarke S., *J. Biol. Chem.*, **1989**, 264, 6164-6170.

se dégrader par formation d'aspartimide.<sup>16,17</sup> De plus, la formation d'un tel motif à partir d'un site Asp-Gly, durant une NCL, a été décrite récemment.<sup>18</sup>



**Figure 8 :** Mécanisme de formation du coproduit **f**

Une augmentation de la température permet d'augmenter la vitesse de la NCL et donc d'obtenir de meilleurs taux de conversion,<sup>10</sup> mais aussi d'augmenter la vitesse de formation du coproduit **f**.<sup>14,19</sup> Des études ont montré que cette réaction secondaire, pH-dépendante, est catalysée en milieu basique,<sup>20-22</sup> ainsi pour résoudre ce problème, nous avons réalisé la NCL à pH 5,8, réduisant ainsi la formation du coproduit **f**.

<sup>16</sup> Robinson N. E., *Proc. Natl. Acad. Sci. USA*, **2002**, 99, 5283-5288.

<sup>17</sup> Tyler-Cross R.; Schirch V., *J. Biol. Chem.*, **1991**, 266, 22549-22556.

<sup>18</sup> Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., *Org. Biomol. Chem.*, **2015**, 13, 7469-7476.

<sup>19</sup> Geiger T.; Clarke S., *J. Biol. Chem.*, **1987**, 262, 785-794.

<sup>20</sup> Patel K.; Borchardt R. T., *Pharm. Res.*, **1990**, 7, 703-711.

<sup>21</sup> Xie M.; Vander Velde D.; Morton M.; Borchardt R. T.; Schowen R. L., *J. Am. Chem. Soc.*, **1996**, 118, 8955-8956.

<sup>22</sup> Song Y.; Schowen R. L.; Borchardt R. T.; Topp E. M., *J. Pharm. Sci.*, **2001**, 90, 141-156.

Les conditions de NCL optimisées ont été appliquées à une NCL à partir du mélange brut obtenu après SPPS du composé **d**. La forme réduite de AhPDF1.1b **e** a été obtenue, après purification HPLC, avec un très bon rendement global de 13,4 %, à comparer avec les 7 % obtenus par Fmoc-SPPS classique à partir d'une résine trityle, d'autant que notre approche rajoute une étape (la NCL). Il est à noter qu'aucune trace du produit de délétion du dipeptide <sup>49</sup>Tyr-Pro n'a été observée dans ce cas.<sup>10,11</sup>

Dans le but de prouver le côté général de cette approche, nous l'avons finalement appliqué à la synthèse de la forme réduite d'une autre défensine, AhPDF1.5, également issue de la plante *Arabidopsis halleri*.<sup>7</sup> La séquence de AhPDF1.5, présentant 41 % d'homologie avec AhPDF1.1b, est la suivante (avec Z = acide pyroglutamique) :



Dans ce cas, le site Ser-Cys C-terminal est bien plus facile à former que le site Pro-Cys du cas précédent. La ligation à partir du mélange brut contenant le précurseur Ser-N-Hnb-Cys correspondant, avec la cystéine dans les conditions optimisées décrites précédemment est quantitative en moins de 2 h; la forme réduite de AhPDF1.5 a été isolée par purification HPLC, avec un rendement global de 16 %.

En conclusion, nous avons tiré parti de la simplicité de la mise en œuvre de la synthèse des peptides crypto-thioester N-Hnb-Cys pour développer une approche originale de synthèse de peptides contenant une cystéine C-terminale. L'approche développée permet d'éviter les réactions secondaires rencontrées lors de la synthèse de ces cibles par une approche classique de Fmoc-SPPS et conduit à l'obtention des peptides cibles avec de meilleurs rendements. Le cas d'un peptide proline-thioester nous a permis d'explorer plus en détails le mécanisme de la N-acylation du dispositif N-Hnb-Cys en mettant en évidence un transfert d'acyle O→N lent dans le cas d'acides aminés décrits comme difficile à coupler sur ce type de dispositif.

Cette nouvelle approche peut nous permettre de constituer une librairie d'analogues de AhPDF1.1b dans le but de poursuivre l'étude des propriétés de cette cible.

Ces travaux sont détaillés dans l'article suivant, soumis à l'éditeur pour publication dans le journal *Organic Letters* : Lelièvre D.; Terrier V. P.; Delmas A. F.; Aucagne V., A native chemical ligation strategy to overcome side reactions during Fmoc-based synthesis of C-terminal cysteine-containing peptides. *Org. Lett.*, **2016**, article soumis.

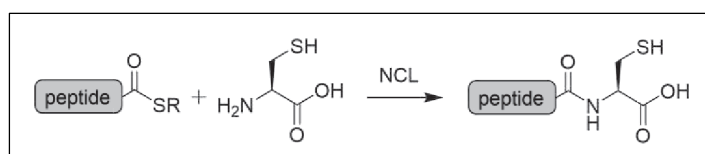




# A native chemical ligation strategy to overcome side reactions during Fmoc-based synthesis of C-terminal cysteine-containing peptides.

Dominique Lelièvre, Victor P. Terrier, Agnès F. Delmas, and Vincent Aucagne\*

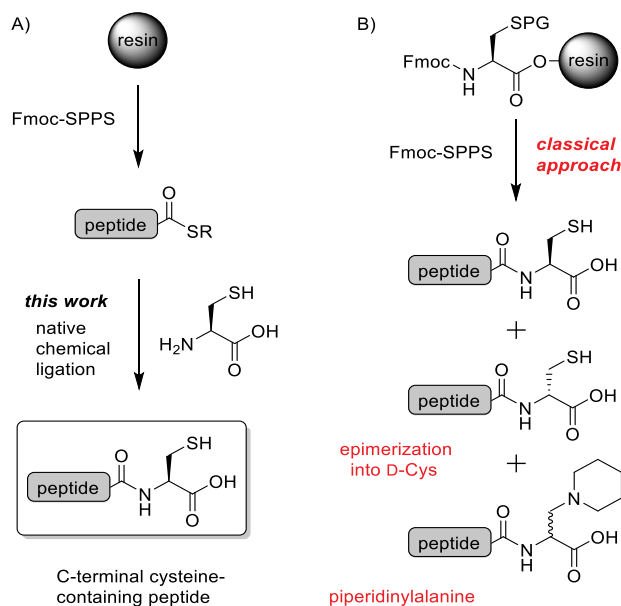
Centre de Biophysique Moléculaire, CNRS UPR 4301, Rue Charles Sadron, 45071 Orléans Cedex 2, France.



**ABSTRACT:** The Fmoc-based solid phase synthesis of C-terminal cysteine-containing peptides is problematic, due to side reactions provoked by the pronounced acidity of the C $\alpha$  proton of cysteine esters. Several methodologies have been developed to try to overcome this limitation, but no simple and efficient method emerged yet. We herein describe a general strategy consisting in the post-synthetic introduction of the C-terminal Cys through a key chemoselective native chemical ligation reaction with *N*-Hnb-Cys peptide crypto-thioesters. This method was successfully applied to the demanding peptide sequences of two natural products of biological interest, giving remarkably high overall yields compared to a state of the art strategy.

C-terminal cysteines are frequent in the sequences of disulfide-rich peptide natural products of biological or therapeutic interest. However, the efficacy of their Fmoc-based solid phase peptide synthesis (SPPS) is impaired by the pronounced acidity of the C $\alpha$  proton of Cys esters. Epimerization can occur not only when forming the ester bond to immobilize Cys on a solid support, but also during the repeated piperidine treatments used for Fmoc deprotection during SPPS elongation.<sup>1</sup> A second commonly encountered side-reaction is the formation of 3-(*N*-piperidinyl)alanine<sup>2</sup> arising from piperidine-mediated  $\beta$ -elimination followed by Michael-type addition to the resulting dehydroalanine (Figure 1B). Contamination by these side-products becomes particularly problematic when synthesizing long peptide sequences. Several methodologies have been developed in the past years to try to overcome these limitations: use of a 2-chloro-trityl linker,<sup>1</sup> side-chain anchoring strategies,<sup>3</sup> or *S*-protecting groups alternative to the classical *S*-trityl protection<sup>4</sup> were shown to dramatically slow down the kinetics of the two parasite reactions. Alternatively, introduction of the Cys as an orthoester derivative<sup>5</sup> can totally suppress the formation of the two side-products but the latter strategy is hampered by tedious synthesis of the building block and requires post-SPPS base-catalyzed hydrolysis of the orthoester.

We recently faced these problems during the SPPS of the reduced form (**1**) of the disulfide-rich miniprotein AhPDF1.1b,<sup>6</sup> an intriguing 51 amino acid plant defensin that, besides classical antifungal properties, is implied in zinc tolerance of *Arabidopsis halleri*.<sup>7</sup> Even when using a 2-chloro-trityl linker, more than 30% epimerization (relative to **1**) was observed in the crude target peptide,<sup>8</sup> that considerably complicated its purification and led to a disappointing overall 7% isolated yield.



**Figure 1.** A general NCL-based approach for the post-synthetic introduction of the C-terminal cysteine on synthetic peptides (A) versus classical Fmoc SPPS (B). PG: protective group.

To improve this yield, we thought to design a general strategy based on post-synthetic introduction of the C-terminal Cys through native chemical ligation (NCL)<sup>9,10</sup> with a peptide thioester (figure 1A). Due to its C-terminal Pro-Cys, the peptide sequence of AhPDF1.1b, peptide **1** constitutes a highly challenging example: indeed, NCL using proline thioesters is known to lead to the slowest kinetics amongst the 20 amino-

genic amino acids.<sup>11</sup> Moreover, a side-product arising from deletion of the C-terminal Xaa-Pro dipeptide through diketopiperazine formation has been recently reported.<sup>12</sup>

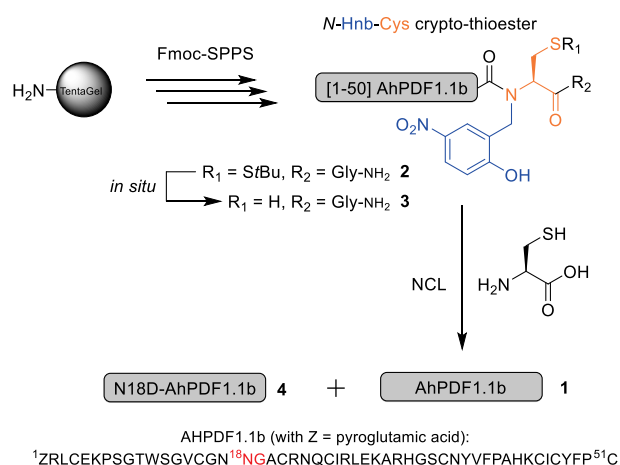
Key compounds for NCL-based synthetic approaches are peptide C $\alpha$  thioesters, but their Fmoc-based synthesis is not straightforward using available tools.<sup>13</sup> We recently introduced a simple, inexpensive and fully automated methodology for the synthesis of peptides crypto-thioesters based on a C-terminal *N*-(2-hydroxy-5-nitrobenzyl)cysteine (*N*-Hnb-Cys) thioesterification device: such peptides are efficient thioester surrogates for NCL.<sup>14</sup> The key phenol group of this device plays a dual role: accelerating the intramolecular *N*→*S* acyl shift-based conversion into a reactive thioester during NCL by a biomimetic acid-base catalysis mechanism, and also assisting the acylation of the secondary amine of the device to introduce the C-terminal residue of the sequence, through an *O*→*N* acyl shift.

The main limitation we identified so far for this otherwise quite general strategy relies in the low yields observed under standard SPPS conditions for the coupling of the *N*-Hnb-Cys device with Pro as well as the sterically-hindered Ile, Val and Thr. Although multiple couplings or microwave heating of the coupling reaction can overcome this problem,<sup>14</sup> the present application to a C-terminal Pro-containing crypto-thioester prompted us to further explore the underlying reasons for this slow coupling. We clearly showed that, at least in the case of Pro, the *N*-acylation process is limited by a slow *O*→*N* acyl shift (see detailed study in the supporting information p S9–15) as it has been previously speculated for a related *O*→*N* shift-assisted *N*-acylation process.<sup>15</sup> This finding is supported by a simple experiment: if a single 30 min coupling at room temperature of Fmoc-Pro-OH using HCTU leads to only 20% *N*-acylation, elimination of the reagents by thorough washing of the resin followed by a 24 h incubation shows a much greater 61% yield. A simple five-fold coupling without further incubation leads to a satisfying 50% yield.

The initial amount of aminomethyl resin can be conveniently increased to achieve the expected synthetic scale, knowing that the capped Ac-(Hnb)Cys(*St*Bu)-Gly-NH<sub>2</sub> co-product is readily eliminated during the ether precipitation of the peptide crypto-thioester after TFA-mediated cleavage. Therefore, robust conditions were optimized to ensure quantitative capping to avoid contamination with C-terminally truncated peptide crypto-thioesters.

Based on these results, automated SPPS of an *N*-Hnb-Cys crypto-thioester corresponding to the [1-50] AhPDF1.1b sequence (**2**) proceeded with a good 55% elongation yield, determined by UV titration of the fluorenylmethyl-piperidine byproduct from the last Fmoc group removal, (Arg<sup>2</sup>) as compared to the first one (Pro<sup>50</sup>) (AhPDF1.1b numbering, see scheme 1). The crude mixture was pure enough to be directly engaged in the next step without any chromatographic purification (see supporting information p S17).

For analytical purposes, the NCL conditions were first optimized using an HPLC-purified sampled of **2**: we performed all the optimization processes (Scheme 1) employing a rather low 1 mM concentration of **2**, representative of NCL with demanding peptides in terms of limited solubility. We started with using a large excess of cysteine (100 mM) in conjunction with



**Scheme 1.** Application to the reduced form of AhPDF1.1b **1** of an NCL-based approach for the post-synthetic introduction of the C-terminal cysteine.

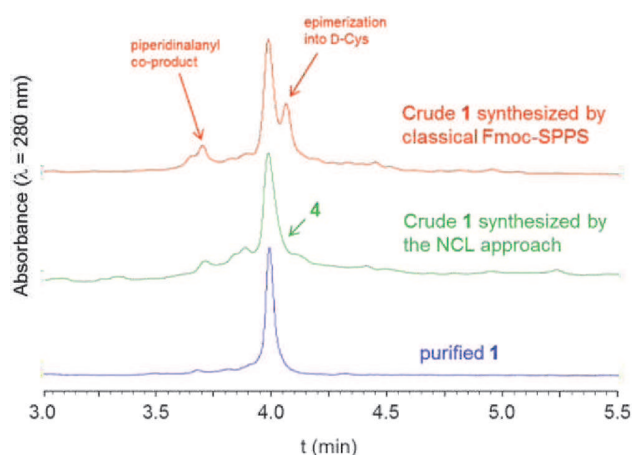
**Table 1.** Optimization of the NCL conditions.

T (°C)	pH	t (h)	NCL yield <sup>a,b</sup> (%)	Ratio <b>1</b> / <b>4</b>	Yield <sup>b</sup> in <b>1</b> (%)
20	7.2	6	3	<i>nd</i>	<i>nd</i>
20	7.2	24	12	95 : 5	11
37	7.2	24	40	81 : 19	32
50	7.2	24	58	41 : 59	24
50	5.8	24	47	93 : 7	44
50	5.8	48	92	87 : 13	80

<sup>a</sup>NCL yields correspond to the formations of both **1** and **4**, not isolated. <sup>b</sup>yields determined by integrations of the HPLC peaks corresponding to compounds **1**, **3** and **4**. *nd*: not determined.

the classical arylthiol catalyst 4-mercaptophenylacetic acid (MPAA, 100 mM) in a pH 7.2 phosphate buffer containing 6 M guanidinium chloride (Gu.HCl) and 50 mM of the classically used *tris*-(2-carboxyethyl)phosphine (TCEP) reducing agent. Under these conditions, *St*Bu-protected **2** was nearly instantaneously converted into latent thioester **3**, ready for *in situ* *N*→*S* acyl shift- based conversion into a NCL-reactive thioester. As anticipated from the C-terminal Pro residue, the reaction was sluggish at room temperature,<sup>11, 12</sup> leading to low 3% and 12% conversion after 6 and 24 h, respectively (Table 1). In the latter case, careful LC-HRMS analysis of the reaction mixture showed the formation of a minor co-product (5% relative to the expected product **1**) with a +1 Da increase in mass, that we attributed to aspartimide formation at the <sup>18</sup>Asn-Gly site, followed by hydrolysis into Asp-Gly and/or isoAsp-Gly (**4**).<sup>16</sup> Asn-Gly sequences are well known to be prone to aspartimide-mediated degradation,<sup>16, 17</sup> being a general problem for the stability of synthetic peptides or recombinant proteins.<sup>18</sup> Further identification of a -17 Da co-product (3.5% relative to **1**), as well as traces (< 0.1%) of [19-51]AhPDF1.1b attributed to the aspartimide-associated cleavage of the Asn-Gly bond,<sup>19</sup> are additional evidences for the <sup>18</sup>Asn-Gly aspartimide hypothesis. As expected, heating of the ligation mixture

2



**Figure 1.** Comparative RP-HPLC analyses of AhPDF1.1b **1**. Red trace: crude peptide obtained with a classical Fmoc-SPPS strategy using a trityl linker. Green trace: crude peptide from the NCL-based approach for the post-synthetic introduction of the C-terminal cysteine. Blue trace: HPLC-purified **1**.

at 37°C or 50°C<sup>12a</sup> led to improved NCL yields (40% and 58%, respectively), but also to a dramatic enhancement of the kinetics of formation of co-product **4** (19% and 59% relative to **1**, respectively). Such a marked dependence on temperature has previously been observed for aspartimide formation at Asn-Gly sites.<sup>16b, 19</sup> This unanticipated difficulty was solved by performing NCL at a lower pH, taking our inspiration from the observed dramatic pH dependence on the base-catalyzed aspartimide-mediated degradation process.<sup>12, 20</sup> Note that, if the optimal pH for NCL using a model *N*-Hnb-Cys peptide was previously evaluated to be around 6.5, the reaction was shown to also perform with fast kinetics at lower pH.<sup>14</sup>

Gratifyingly, conducting the ligation of **2** with cysteine at pH 5.8 for either 24 or 48 h led to a much cleaner reaction (co-product **4** limited to 7% and 13%, respectively), and a 80% yield in **1** in the latter case, an excellent yield when considering the specific limitations arising from the highly demanding AhPDF1.1b sequence. Further careful LC-MS analysis of the crude ligation mixture starting from crude **2** showed about 5% aspartimide formation (-17 Da) together with 6% hydrolysis of the crypto-thioester into [1-50]AhPDF1.1b.<sup>21</sup> The surprising relatively high amount of hydrolysis compared to our model studies<sup>14</sup> was attributed to the C-terminal proline-containing sequence of crypto-thioester **2**. No traces of deletion of the <sup>49</sup>Tyr-Pro dipeptide were observed in any tested conditions.<sup>12</sup>

Further highlighting the efficiency of our strategy, NCL on the crude SPPS mixture gratifyingly led to a much improved chromatographic profile as compared with classical Fmoc-SPPS using a trityl linker (Figure 1). HPLC purification furnished the pure reduced form of AhPDF1.1b **1** in a good 13.4 % overall yield, that is about the double of the yield initially obtained, in spite of the sequence-specific aspartimide problems we encountered. Despite a close elution in RP-HPLC of **4** with **1**, less than 2% contamination was observed in the purified material.

In order to illustrate the general scope of our strategy, we finally applied the protocol optimized for AhPDF1.1b to an-

other related defensin from *A. hallieri*, AhPDF1.5<sup>6, 22</sup> without trying to adapt it, even if the sequence implies a much more kinetically favorable NCL at serine. The SPPS of [1-50]AhPDF1.5 crypto-thioester proceeded in a good 56% elongation yield (see supporting information p S25). As expected, NCL was much faster (ligation completed in less than 2 h), and the overall process including NCL and purification starting from the crude peptide crypto-thioester led to isolation of the pure (> 95%) reduced form of AhPDF1.5 in an excellent 16% overall yield (see supporting information p S27). Note that, in this second case, less than 2% thioester hydrolysis was observed, supporting our hypothesis from a C-terminal Pro-specific –and somewhat limited– side reaction in the first example.

In conclusion, we herein introduced a simple and general NCL-based strategy for the high-yielding synthesis of C-terminal Cys-containing peptides, which allows bypassing the problematic side-reactions associated to these valuable peptide sequences. The work described here conveniently makes use of the facile synthesis of peptide crypto-thioesters through the *N*-Hnb-Cys technology, but could probably be generalized to other popular Fmoc-based strategies for thioester synthesis. Work is ongoing to generate a diverse library of naturally occurring C-terminal Cys-containing plant defensins and analogues, related to our primary target AhPDF1.1b, and thus to help deciphering the unique zinc-tolerance properties induced by this molecule through further structural, bio-inorganic and biological studies.

## ASSOCIATED CONTENT

### Supporting Information

Materials and methods, experimental section, compound characterization, and RP-HPLC chromatograms and their analyses. Detailed studies on the mechanism of the *N*-acylation of the device. This material is available free of charge on the ACS Publications website.

## AUTHOR INFORMATION

### Corresponding Author

\* E-mail: aucagne@cns-orleans.fr

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENT

We thank Dr Guillaume Gabant, Dr Cyril Colas and the mass spectrometry platforms of CBM and FR2708 federation for the MS and LC-MS analyses. The Pierre Potier foundation for the chemistry of natural products is gratefully acknowledged for a PhD fellowship for VPT.

## REFERENCES

- (1) Fujiwara, Y.; Akaji, K.; Kiso, Y. *Chem. Pharm. Bull.* **1994**, *42*, 724–726.
- (2) Lukszo, J.; Patterson, D.; Albericio, F.; Kates, S. A. *Lett. Pept. Sci.* **1996**, *3*, 157–166.
- (3) Barany, G.; Han, Y.; Hargittai, B.; Liu R.-Q.; Varkey J. T. *Biopolymers* **2003**, *71*, 652–666.

- (4) (a) Hibino, H.; Nishiuchi, Y. *Org. Lett.* **2012**, *14*, 1926–1929. (b) Hibino, H.; Miki, Y.; Nishiuchi, Y. *J. Pep. Sci.* **2014**, *20*, 30–35. (c) Ramos-Tornillero, Y.; Rodriguez, R.; Albericio, F. *Org. Lett.* **2015**, *17*, 1680–1683.
- (5) Huang, Z.; Derksen, D. J.; Vederas J. C. *Org. Lett.* **2010**, *12*, 2282–2285.
- (6) Meindre, F.; Lelièvre, D.; Loth, K.; Mith, O.; Aucagne, V.; Berthomieu, P.; Marquès, L.; Delmas, A. F.; Landon, F.; Paquet F. *Biochemistry* **2014**, *53*, 7745–7754.
- (7) (a) Mirouze, M.; Sels, J.; Richard, O.; Czernic, P.; Loubet, S.; Jacquier, A.; Francois, I. E.; Cammue, B. P.; Lebrun, M.; Berthomieu, P.; Marques, L. *Plant J.* **2006**, *47*, 329–342. (b) Mith, O.; Benhamdi, A.; Castillo, T.; Bergé, M.; MacDiarmid, C. W.; Steffen, J.; Eide, D. J.; Perrier, V.; Subileau, M.; Gosti, F.; Berthomieu, P.; Marques, L. *Microbiol. Open* **2015**, *4*, 409–422.
- (8) About 8% epimerization into D-Cys was observed after 10 coupling/capping/Fmoc deprotection cycles, and 27% after 32 cycles, see supporting information p S4–6.
- (9) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776–779.
- (10) Note that NCL using a single residue cysteinamide, but unrelated to the scope of the present work, has been described recently, see: Dang, B.; Kubota, T.; Correa, A. M.; Bezanilla, F.; Kent, S. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 8970–8974.
- (11) (a) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073. (b) Pollock, S. B.; Kent, S. B. H. *Chem. Commun.* **2011**, *47*, 2342–2344.
- (12) (a) Raibaut, L.; Seeberger, P., Melnyk O. *Org. Lett.* **2013**, *15*, 5516–5519. (b) Nakamura, T.; Shigenaga, A.; Sato, K.; Tsuda, Y.; Sakamoto, K.; Otaka, A. *Chem. Commun.* **2014**, *50*, 58–60.
- (13) Mende, F.; Seitz, O. *Angew. Chem., Int. Ed.* **2011**, *50*, 1232–1240.
- (14) Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas A. F., Aucagne, V. *Chem. Sci.* **2016**, DOI: 10.1039/C5SC02630J.
- (15) Miranda, L. P.; Meutermans, W. D.; Smythe, M. L.; Alewood, P. F. *J. Org. Chem.* **2000**, *65*, 5460–5468.
- (16) (a) Clarke, S. *Int. J. Pept. Prot. Res.* **1987**, *30*, 808–821. (b) Stephenson, R. C.; Clarke, S. *J. Biol. Chem.* **1989**, *264*, 6164–6170. (c) Robinson, N. E. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5283–5288.
- (17) Tyler-Cross, R.; Schirch, V. *J. Biol. Chem.* **1991**, *266*, 22549–22556.
- (18) Aspartimide formation at an Asp-Gly site was recently described as a side-reaction during NCL, see: Cowper, B.; Shariff, L.; Chen, W.; Gibson, S. M.; Diab W.-L.; Macmillan, D. *Org. Biomol. Chem.*, **2015**, *13*, 7469–7476.
- (19) Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262*, 785–794.
- (20). (a) Patel, K.; Borchardt, R. T. *Pharm. Res.* **1990**, *7*, 703–711. (b) Xie, M.; Vander Velde, D.; Morton, M.; Borchardt, R. T.; Schowen, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 8955–8956. (c) Song, Y.; Schowen, R. L.; Borchardt, R. T.; Topp, E. M. *J. Pharm. Sci.* **2001**, *90*, 141–156.
- (21) Traces of non-acetylated truncated peptides containing an N-terminal Cys residue were also observed in the LC/MS analysis of the mixture (see supporting information p S22). These co-products were attributed to in situ N→S shift at Xaa-Cys sites followed by trans-thioesterification at high temperature, as reported by Macmillan *et al.*. See (18) and references cited therein.
- (22) AhPDF1.5 sequence: <sup>1</sup>ZLCKRESETWSGRCVNDYQCRDH CINNDRGNDGYCAGGYPWYRGCFCCFFS<sup>51</sup>C, with Z = pyroglutamic acid (41% homology with AhPDF1.1b sequence).

## Références bibliographiques

1. Brazeau P.; Vale N.; Burgus R.; Ling N.; Butcher M.; Rivier, J.; Guillemin R., Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone. *Science*, **1973**, *179*, 77–79.
2. Halai, R.; Craik, D. J., Conotoxins: natural product drug leads. *Nat. Prod. Rep.*, **2009**, *26*, 526–536.
3. Konno K.; Pico G.; Gutierrez V. P.; Brigatte P.; Zambelli V. O.; Camargo A. C. M.; Cury Y., Crotalphine, a novel potent analgesic peptide from the venom of the South American rattlesnake *Crotalus durissus terrificus*. *Peptides*, **2008**, *29*, 1293–1304.
4. Fujiwara Y.; Akaji K.; Kiso Y., Racemization-free synthesis of C-terminal cysteine-peptide using 2-chlorotrityl resin. *Chem. Pharm. Bull.*, **1994**, *42*, 724–726.
5. Lukszo J.; Patterson D.; Albericio F.; Kates S. A., 3-(1-Piperidinyl)alanine formation during the preparation of C-terminal cysteine peptides with the Fmoc/tBu strategy. *Lett. Pept. Sci.*, **1996**, *3*, 157–166.
6. Huang Z.; Derksen D. J.; Vederas J. C., Preparation and use of cysteine orthoesters for solid-supported synthesis of peptides. *Org. Lett.*, **2010**, *12*, 2282–2285.
7. Meindre F.; Lelièvre D.; Loth K.; Mith O.; Aucagne V.; Berthomieu P.; Marquès L.; Delmas A. F.; Landon C.; Paquet F., The nuclear magnetic resonance solution structure of the synthetic AhPDF1.1b plant defensin evidences the structural feature within the  $\gamma$ -motif. *Biochemistry*, **2014**, *53*, 7745–7754.
8. Hackeng T. M.; Griffin J. H.; Dawson P. E., Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. Usa*, **1999**, *96*, 10068–10073.
9. Pollock S. B.; Kent S. B. H., An investigation into the origin of the dramatically reduced reactivity of peptide-prolyl-thioesters in native chemical ligation. *Chem. Commun.*, **2011**, *47*, 2342–2344.
10. Raibaut L.; Seeberger P.; Melnyk O., Bis(2-sulfanylethyl)amido peptides enable native chemical ligation at proline and minimize deletion side-product formation. *Org. Lett.*, **2013**, *15*, 5516–5519.
11. Nakamura T.; Shigenaga A.; Sato K.; Tsuda Y.; Sakamoto K.; Otaka A., Examination of native chemical ligation using peptidyl prolyl thioesters. *Chem. Commun.*, **2014**, *50*, 58–60.
12. Miranda L. P.; Meutermans W. D.; Smythe M. L.; Alewood P. F., An activated  $O\rightarrow N$  acyl transfer auxiliary: efficient amide-backbone substitution of hindered "difficult" peptides. *J. Org. Chem.*, **2000**, *65*, 5460–5468.
13. Diaz-Mochón J. J.; Bialy L.; Bradley M., Full orthogonality between Dde and Fmoc: the direct synthesis of PNA-peptide conjugates. *Org. Lett.*, **2004**, *6*, 1127–1129.

14. Clarke S., Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. *Int. J. Pept. Prot. Res.*, **1987**, *30*, 808-821.
15. Stephenson R. C.; Clarke S., Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J. Biol. Chem.*, **1989**, *264*, 6164-6170.
16. Robinson N. E., Protein deamidation. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 5283-5288.
17. Tyler-Cross R.; Schirch V., Effects of amino acid sequence, buffers, and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides. *J. Biol. Chem.*, **1991**, *266*, 22549-22556.
18. Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., Expanding the scope of *N* → *S* acyl transfer in native peptide sequences. *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.
19. Geiger T.; Clarke S., Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *J. Biol. Chem.*, **1987**, *262*, 785-794.
20. Patel K.; Borchardt R. T., Chemical pathways of peptide degradation. II. Kinetics of deamidation of an asparaginyl residue in a model hexapeptide. *Pharm. Res.*, **1990**, *7*, 703-711.
21. Xie M.; Vander Velde D.; Morton M.; Borchardt R. T.; Schowen R. L., pH-Induced change in the rate-determining step for the hydrolysis of the Asp/Asn-derived cyclic-imide intermediate in protein degradation. *J. Am. Chem. Soc.*, **1996**, *118*, 8955-8956.
22. Song Y.; Schowen R. L.; Borchardt R. T.; Topp E. M., Effect of 'pH' on the rate of asparagine deamidation in polymeric formulations: 'pH'-rate profile. *J. Pharm. Sci.*, **2001**, *90*, 141-156.

# CHAPITRE 4





## Chapitre 4 : Synthèse de peptides cycliques naturels riches en cystéine

Les peptides cycliques riches en ponts disulfure (cDRP : *cyclic disulfide-rich peptides*) sont trouvés chez de très nombreuses espèces végétales ou animales. Ils représentent une classe importante de mini-protéines caractérisées par une large gamme d'activités biologique.<sup>1</sup> En complément de leur structure compacte due aux ponts disulfure, leur squelette cyclique leur confère une stabilité accrue face à la dégradation enzymatique, en comparaison de leurs analogues linéaires. Ces peptides constituent donc des modèles très intéressants pour l'ingénierie de molécules d'intérêt biologique.<sup>2</sup> Leur grande diversité chimique est illustrée par de très nombreuses séquences allant de 14 à 37 acides aminés (aa), contraintes par 1 ou 3 ponts disulfure. Il a été démontré, par ailleurs, que la cyclisation de toxines peptidiques naturelles acycliques résultait en une augmentation significative de leur stabilité en milieu biologique, leur conférant ainsi un potentiel thérapeutique accru.<sup>3,4</sup> La mise au point de méthodes de synthèse rapide et efficace de ce type de cible est donc un but important pour les chimistes des peptides.

L'approche historiquement utilisée reposait sur la cyclisation de séquences peptidiques protégées à l'aide d'agents de couplage en milieu organique.<sup>5,6</sup> Néanmoins, l'activation de la fonction carboxylique peut conduire à une épimérisation de l'acide aminé C-terminal. De plus, comme cela a déjà été mentionné, l'utilisation de peptides protégés est souvent rendue difficile à cause d'une faible solubilité dans le milieu réactionnel. Ces problèmes limitaient la synthèse de peptides cycliques de taille moyenne.

La NCL utilisée, de façon intramoléculaire est une réaction de choix pour permettre la cyclisation d'une séquence peptidique, à partir d'un précurseur linéaire contenant les deux fonctions nécessaires à cette réaction (figure 1).<sup>7-9</sup>

---

<sup>1</sup> Cezamar M.; Kwon S.; Mahatmanto T.; Ravipati A. S.; Craik D. J., *Curr. Top. Med. Chem.*, **2012**, *12*, 1534-1545.

<sup>2</sup> Craik D. J., *Science*, **2006**, *311*, 1563-1564.

<sup>3</sup> Clark R. J.; Fischer H.; Dempster L.; Daly N. L.; Rosengren K. J.; Nevin S. T.; Meunier F. A.; Adams D. J.; Craik D. J., *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 13767-13772.

<sup>4</sup> Akcan M.; Stroud M. R.; Hansen S. J.; Clark R. J.; Daly N. L.; Craik D. J.; Olson J. M., *J. Med. Chem.*, **2011**, *54*, 782-787.

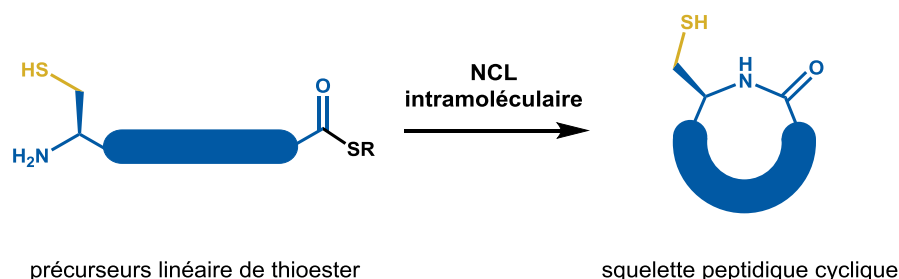
<sup>5</sup> Albericio F.; Hammer R. P.; García-Echevarría C.; Molins M. A.; Chang J. L.; Munson M. C.; Pons M.; Giralt E.; Barany G., *Int. J. Pept. Protein Res.*, **1991**, *37*, 402-413.

<sup>6</sup> Lambert J. N.; Mitchell J. P.; Roberts K. D., *J. Chem. Soc., Perkin Trans. 1*, **2001**, 471-484.

<sup>7</sup> Zhang L.; Tam J. P., *J. Am. Chem. Soc.*, **1997**, *119*, 2363-2370.

<sup>8</sup> Camarero J. A.; Muir T. W., *Chem. Commun.*, **1997**, 1369-1370.

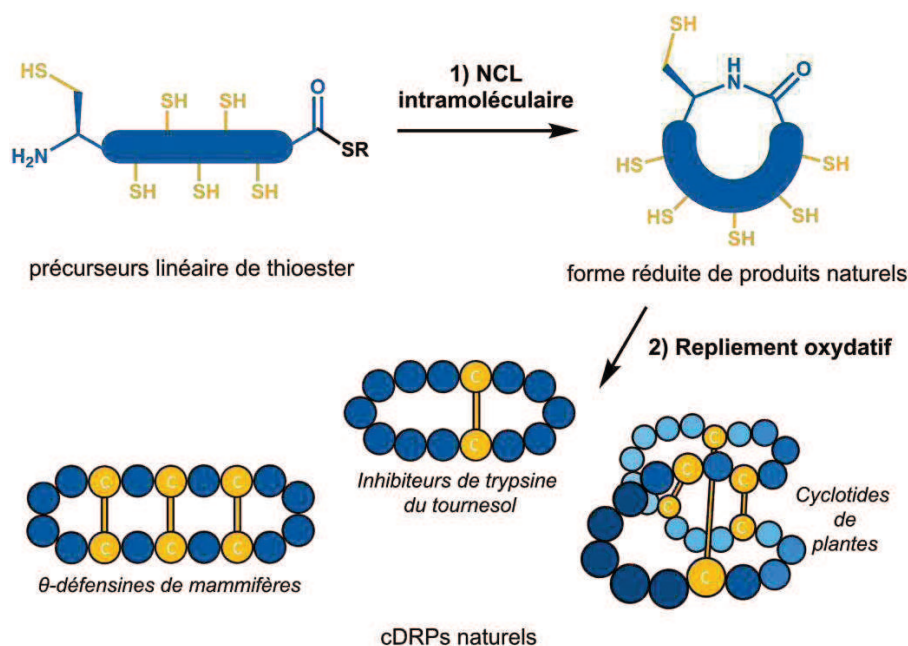
<sup>9</sup> Camarero J. A.; Cotton G. J.; Adeva A.; Muir T. W., *J. Pept. Res.*, **1998**, *51*, 303-316.



**Figure 1 :** Cyclisation du squelette peptidique par NCL intramolécuaire

Il faut noter que la NCL a récemment été appliquée à la synthèse de peptide cyclique de grande taille par une première ligation entre deux fragments suivie de la cyclisation du fragment obtenu par une NCL intramolécuaire entre ses extrémités N- et C-terminales.<sup>10</sup>

De façon évidente, la NCL intramolécuaire est également très adaptée à la synthèse de cDRPs en raison du nombre important de cystéines de leur séquence.<sup>11,12</sup> La cyclisation peut être suivie d'une étape de repliement oxydatif pour conduire à l'obtention de la forme native du produit ciblé. Il a d'ailleurs été montré que la cyclisation préalable au repliement pouvait favoriser la formation correcte des ponts disulfure.<sup>13,14</sup>



**Figure 2 :** Stratégie de synthèse totale de produits naturels cycliques riches en ponts disulfure

<sup>10</sup> Boll E.; Ebran J-P.; Drobecq H.; El-Mahdi O.; Raibaut L.; Ollivier N.; Melnyk O., *Org. Lett.*, **2015**, *17*, 130-133.

<sup>11</sup> Tam J. P.; Lu Y-A., *Tetrahedron Lett.*, **1997**, *38*, 5599-5602.

<sup>12</sup> Camarero J. A.; Pavel J.; Muir T. W., *Angew. Chem. Int. Ed.*, **1998**, *37*, 347-349.

<sup>13</sup> Daly N. L.; Love S.; Alewood P. F.; Craik D. J., *Biochemistry*, **1999**, *38*, 10606-10614.

<sup>14</sup> Aboye T. L.; Li Y.; Majumder S.; Hao J.; Shekhtman A.; Camarero J. A., *Bioorg. Med. Chem. Lett.*, **2012**, *22*, 2823-2826.

Cette utilisation de la NCL n'échappe pas au problème de la synthèse délicate de la fonction thioester C-terminale par Fmoc-SPPS. Plusieurs méthodes de synthèse de peptides thioester ont été appliquées à la cyclisation de cDRPs par NCL comme la stratégie du *safety catch linker* de type sulfonamide,<sup>14-16</sup> les peptides hydrazide,<sup>17,18</sup> les peptides Nbz<sup>19</sup> et MeNbz,<sup>20</sup> le dispositif TEA<sup>21</sup> ou encore l'utilisation de peptides de séquence native avec une cystéine C-terminale.<sup>22,23</sup> Par sa mise en œuvre simple, la méthode des peptides hydrazide paraît particulièrement efficace, néanmoins, les limitations soulignées lors de la description de ces différentes stratégies s'appliquent également à cette approche de cyclisation.

Nous avons décidé d'examiner la cyclisation du squelette peptidique de cinq cDRPs naturels, représentatifs des classes les plus étudiées à partir de peptides crypto-thioester *N*-Hnb-Cys.

Pour rappel, les cyclotides constituent la plus grande famille de cDRPs trouvés chez les plantes (28 à 37 aa), ils sont caractérisés par un motif à nœud de cystine formé par 3 ponts disulfure.<sup>24</sup> Kalata B1 (kB1, 29 aa) est le premier cyclotide à avoir été découvert, il a été isolé de la plante *Oldenlandia affinis* (café).<sup>13</sup> La cycloviolacine O2 (cO2, 30 aa) est un autre cyclotide, isolé de la plante *Viola odorata* (violette).<sup>25</sup> Cter M (29 aa) a récemment été découvert dans les feuilles de la plante *Clitoria ternatea* (pois bleu).<sup>26</sup>

Les inhibiteurs de trypsine du tournesol (*sunflower trypsin inhibitors* – SFTI) font partie d'une autre classe de cDRPs isolés des graines de tournesol *Helianthus annuus* dont la séquence est plus courte (14 aa) et possédant un seul pont disulfure.<sup>27</sup> SFTI-1 est le premier membre de cette famille à avoir été découvert.

<sup>15</sup> Thongyoo P.; Tate E. W.; Leatherbarrow R. J., *Chem. Commun.*, **2006**, 2848-2850.

<sup>16</sup> Thongyoo P.; Roqué-Rosell N.; Leatherbarrow R. J.; Tate E. W., *Org. Biomol. Chem.*, **2008**, 6, 1462-1470.

<sup>17</sup> Zheng J-S.; Tang S.; Guo Y. Chang H-N.; Liu L., *ChemBioChem*, **2012**, 13, 542-546.

<sup>18</sup> Chen Y-Q.; Chen C-C.; He Y.; Yu M.; Xu L.; Tian C-I.; Guo Q-X.; Shi J.; Zhang M.; Li Y-M., *Tetrahedron Lett.*, **2014**, 55, 2883-2886.

<sup>19</sup> Gunasekera S.; Aboye T. L.; Madian W. A.; El-Seedi H. R.; U Göransson, *Int. J. Pept. Res. Ther.*, **2013**, 19, 43-54.

<sup>20</sup> Blanco-Canosa J. B.; Nardone B.; Albericio F.; Dawson P. E., *J. Am. Chem. Soc.*, **2015**, 137, 7197-7209.

<sup>21</sup> Taichi M.; Hemu X.; Qiu Y.; Tam J. P., *Org. Lett.*, **2013**, 15, 2620-2623.

<sup>22</sup> Shariff L.; Zhu Y.; Cowper B.; Di W-L.; Macmillan D., *Tetrahedron*, **2014**, 70, 7675-7680.

<sup>23</sup> Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., *Org. Biomol. Chem.*, **2015**, 13, 7469-7476.

<sup>24</sup> Craik D. J.; Daly N. L.; Bond T.; Waine C., *J. Mol. Biol.*, **1999**, 294, 1327-1336.

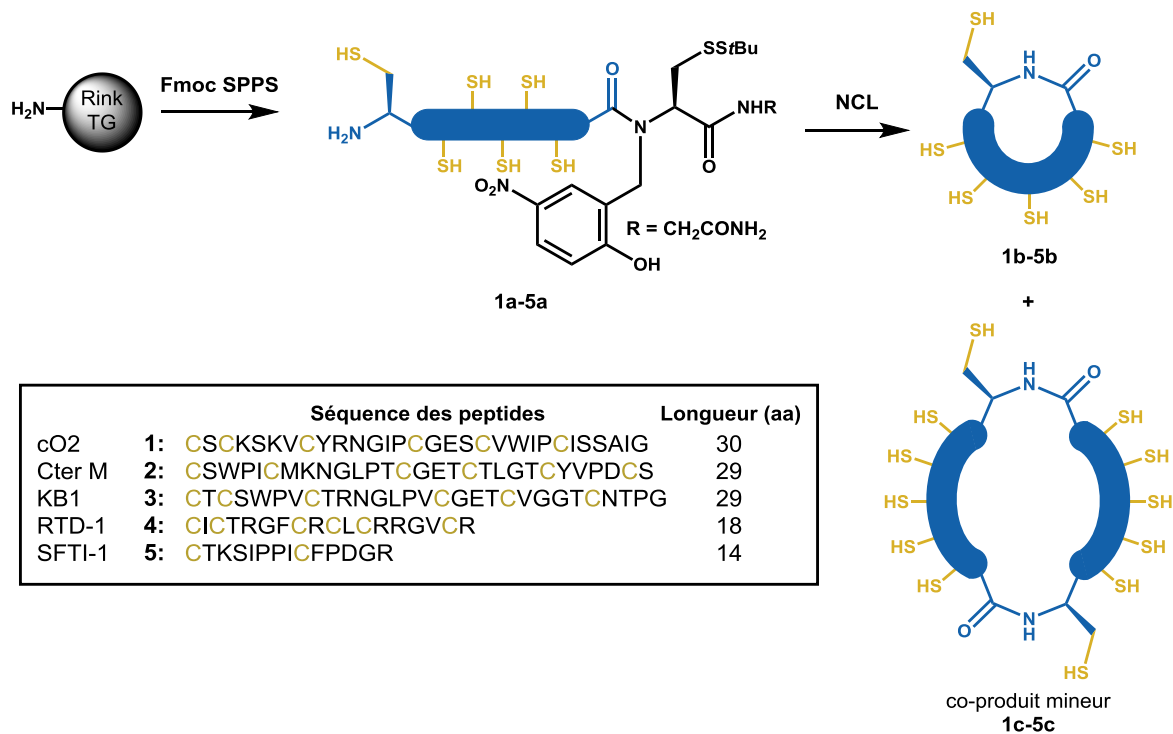
<sup>25</sup> Aboye T. L.; Clark R. J.; Craik D. J.; Göransson U., *ChemBioChem*, **2008**, 9, 103-113.

<sup>26</sup> Poth A. G.; Colgrave M. L.; Lyons R. E.; Daly N. L.; Craik D. J., *Proc. Natl. Acad. Sci. USA*, **2011**, 108, 10127-10132.

<sup>27</sup> Luckett S.; Santiago Garcia R.; Barker J. J.; Konarev A. V.; Shewry P. R.; Clarke A. R.; Brady R. L., *J. Mol. Biol.*, **1999**, 290, 525-533.

Les  $\theta$ -défensines constituent l'unique classe de cDRPs trouvés chez les mammifères, leur structure, composée de 18 acides aminés, est caractérisée par une forme d'échelle et comprend 3 ponts disulfure.<sup>28</sup> RTD-1, issu des leucocytes du macaque rhésus (*Macaca mulatta*), est le premier membre décrit de la famille des  $\theta$ -défensines.<sup>29</sup>

Notre stratégie consiste donc en la synthèse des précurseurs linéaires *N*-Hnb-Cys par Fmoc-SPPS de ces différentes séquences puis leur cyclisation par NCL intramoléculaire (figure 3).



**Figure 3 :** Cyclisation du squelette peptidique des peptides *N*-Hnb-Cys **1a-5a**

Après Fmoc-SPPS et purification HPLC, les précurseurs linéaires **1a-5a** ont été obtenus avec des rendements de 10 à 20 % sans optimisation de l'assemblage de la séquence, des conditions de traitement TFA ou de purification. Ces rendements moyens, en considérant les petites séquences synthétisées, ont été attribués à des couplages difficiles potentiellement incomplets, l'oxydation de la méthionine (dans le cas de **2a**), ou à l'alkylation possible des résidus Trp ou Cys par capture de groupements *t*Bu libérés lors du traitement TFA. Ce dernier problème a déjà été décrit lors de l'utilisation des peptides Nbz et MeNbz pour la synthèse de peptides riches en cystéine.<sup>19,20</sup>

<sup>28</sup> Conibear A. C.; Craik D. J., *Angew. Chem. Int. Ed.*, **2014**, 53, 10612-10623.

<sup>29</sup> Tang Y-Q.; Yuan J.; Ösapay G.; Ösapay K.; Tran D.; Miller C. J.; Ouellette A. J.; Selsted M. E., *Science*, **1999**, 286, 498-502.

Les peptides **1a-5a** ont été engagés dans des conditions de NCL à une concentration de 1 mM. Dans tous les cas, la cyclisation est très rapide (taux de conversion HPLC supérieur à 90 % après 6 h de réaction). Néanmoins, dans certains cas, la formation des produits correspondant au dimère cyclique de la séquence cible a été observée. **3c** et **5c** ont été quantifiés par HPLC analytique à une hauteur de 4 et 5 %. La réalisation des NCLs pour **3a** et **5a** en utilisant une concentration en peptide de 0,1 mM a permis de quasiment supprimer cette réaction secondaire.

Lors de la synthèse de SFTI-1 par utilisation d'un peptide avec une cystéine C-terminale, Macmillan et ses collaborateurs ont observé la formation d'un aspartimide au site Asp-Gly (mécanisme décrit dans le chapitre 3).<sup>23</sup> Dans notre cas, le produit de cette réaction secondaire a été uniquement observé à l'état de trace, lors de l'analyse par LC/MS du mélange brut après NCL. Ceci est en accord avec notre utilisation de conditions plus douces (température de 37 °C) que les conditions utilisées par Macmillan et ses collaborateurs (50 °C).

Finalement, nous avons isolé les formes réduites des peptides cycliques **1b-5b**, par purification HPLC, avec d'excellents rendements (85 à 97 %).

En résumé, nous avons montré que les peptides *N*-Hnb-Cys peuvent être utilisés pour la cyclisation du squelette de peptides cycliques riches en ponts disulfure *via* une NCL intramoléculaire. Les réactions de cyclisation avec ces peptides ont conduit à l'obtention des formes réduites de plusieurs produits naturels avec d'excellents rendements, montrant ainsi le fort potentiel de notre méthodologie pour cette approche. Le repliement oxydatif de ces différentes cibles est actuellement en cours au laboratoire dans le but d'applications biologiques futures.

Ces travaux sont détaillés dans l'article suivant, soumis à l'éditeur pour publication dans le journal *Chemical Communication* : Terrier V. P.; Delmas A. F.; Aucagne V., Highly efficient synthesis of cysteine-rich cyclic peptides through intramolecular native chemical ligation of *N*-Hnb-Cys peptide crypto-thioesters. *Chem. Commun.*, **2016**, article soumis.



## Highly efficient synthesis of cysteine-rich cyclic peptides through intramolecular native chemical ligation of *N*-Hnb-Cys peptide crypto-thioesters

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Victor P. Terrier, Agnès F. Delmas and Vincent Aucagne\*

www.rsc.org/

The potential of the recently described *N*-Hnb-Cys methodology that allows straightforward synthesis of peptide crypto-thioesters is herein demonstrated for the highly efficient native chemical ligation-based backbone cyclization of a representative range of naturally occurring disulfide-rich peptides.

Cyclic disulfide-rich peptides (cDRP) are found in all domains of life, from bacteria to plants and animals. They constitute an important class of naturally occurring peptides that show a very broad range of biological activities, mainly related to host-defense processes.<sup>1</sup> The exceptional chemical and conformational stability of these emerging peptide-based pharmacophores is particularly attractive. In addition to constrained three-dimensional structures imposed by the disulfide bridges, head-to-tail backbones cyclization result in a much enhanced resistance toward enzymatic degradation compared to their non-cyclized counterparts. All together, these features make cDRP promising lead compounds or scaffolds for the development of peptide-based drugs.<sup>2</sup>

Naturally occurring cDRPs are characterized by a high cysteine content in their primary sequences (10–20%) and show a large chemical diversity: the cyclic polypeptide backbone, ranging from 14 to 37 amino acid residues, can be constrained by either one or three disulfide bonds. Biological exploration of these promising molecules has been initiated but is still limited by synthetic accessibility.

The advent of synthetic strategies based on intramolecular native chemical ligation<sup>3</sup>, a very efficient reaction for the backbone cyclization of fully unprotected peptide,<sup>4</sup> has provided a substantial stimulus for the exploration of the “medium-sized” chemical space covered by cDRPs. Linear precursors requiring both a C-terminal thioester and an N-terminal cysteine are indeed direct precursors ideally suited for the synthesis of such cysteine-rich compounds. (Figure 1).<sup>5</sup>

The preparation of peptide C $\alpha$ -thioesters represents the bottleneck of NCL-based synthetic schemes. If the Boc-SPPS strategy was pioneered in the late 90s,<sup>6</sup> the more widely used Fmoc/tBu approach is nowadays largely preferred but yet not totally straightforward.<sup>7</sup> The synthesis of cDRP linear thioester precursors using different Fmoc-based methodologies has been reported, including sulfonamide-based safety-catch linker,<sup>5b,8</sup> Liu’s hydrazides,<sup>9</sup> Tam’s thioethylalkylamide (TEA),<sup>10</sup> and the first (Nbz)<sup>11</sup> and second (MeNBz)<sup>12</sup> generation of Dawson’s *N*-acylurea. Application of a strategy involving a spontaneous *N*→*S* acyl transfer in a native Xaa-Cys sequence has also been recently documented.<sup>13</sup> Nevertheless, these strategies are impaired by slow reaction kinetics leading to low NCL-based cyclization yields in the latter case, and in the other cases by the need of post-SPPS transformations for the preparation of the linear precursors.

We recently introduced a simple, inexpensive and fully automated strategy for the Fmoc-based synthesis of peptide crypto-thioesters. This strategy is based on a *N*-(2-hydroxy-5-nitrobenzyl)cysteine (*N*-Hnb-Cys) C-terminal thioesterification device operating through an intramolecular *N*→*S* acyl transfer

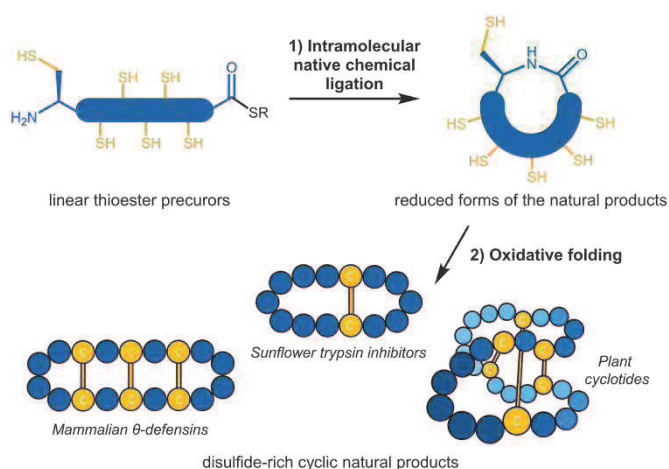


Fig. 1 Synthesis of naturally occurring disulfide-rich cyclic peptides through intramolecular native chemical ligation.

Centre de Biophysique Moléculaire, CNRS UPR 4301, Rue Charles Sadron, 45071 Orléans Cedex 2, France. E-mail: aucagne@cnrs-orleans.fr

† Electronic Supplementary Information (ESI) available: Materials and methods, experimental section, compound characterization, RP-HPLC chromatograms and LC/MS analyses. See DOI: 10.1039/x0xx00000x

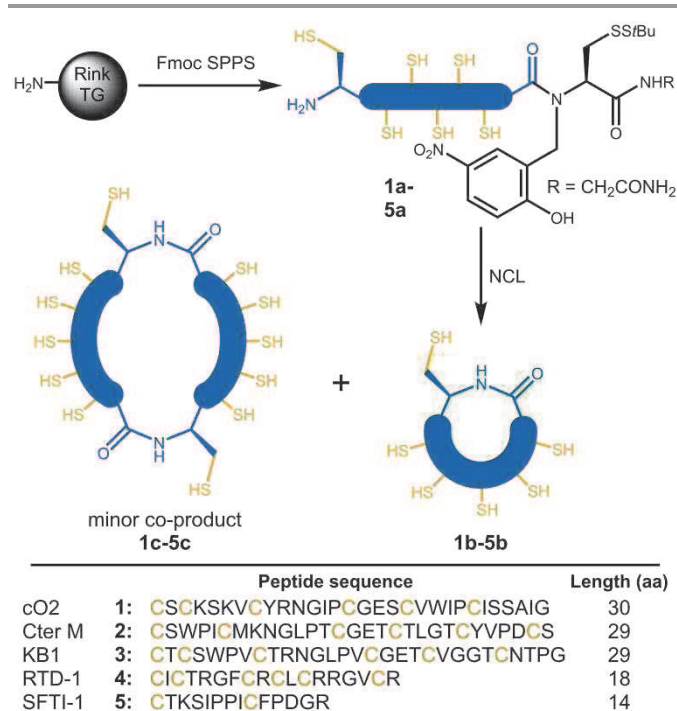


bearing this Hnb-Cys device can thus be directly used in native chemical ligation reactions without requirement of post-SPPS steps. We herein report on the efficient NCL-based backbone cyclization of five naturally occurring cyclic disulfide-rich peptides sequences using the *N*-Hnb-Cys peptides cryptothioesters methodology (Figure 2).

The targets were selected to represent the diversity of naturally occurring cDRPs in terms of length and disulfide arrangement. Kalata B1 (kB1, 29 aa) is the prototypic Möbius cyclotide, isolated from the Rubiaceae *Oldenlandia affinis*.<sup>5a</sup> Cycloviolacin (cO2, 30 aa) is a bracelet cyclotide found in the Violaceae *Viola odorata*.<sup>15</sup> The more recently discovered cyclotide from the Fabaceae *Clitoria ternatea*, Cter M (29 aa), was added to our study to further establish the versatility of our approach.<sup>16</sup> Cyclotides, the major cDRP family found in plants,<sup>17</sup> are characterized by three conserved disulfide bridges that shape a cyclic cystine knot (CCK) mechanically-interlocked motif.<sup>18</sup> Sunflower trypsin inhibitor-1 (SFTI-1, 14 aa) is the first member of a family of smaller plant cyclic peptides, isolated from the Asteraceae *Helianthus annuus*, and constrained by a single disulfide bridge.<sup>19</sup> The  $\theta$ -defensins family is, to date, the only cyclic peptides family found in animals.  $\theta$ -defensins are structurally characterized by their ladder-like motif formed by a 18 aa cyclic backbone and three disulfide bridges.<sup>20</sup> The rhesus  $\theta$ -defensin-1 (RTD-1), expressed in rhesus macaque (*Macaca mulatta*) leukocytes, was the first member of this defending family to be described.<sup>21</sup>

Our synthetic scheme started with the syntheses of the linear precursors **1a-5a** from a Rink Tentagel resin. Syntheses were entirely performed on a parallel peptide synthesizer, including Hnb installation through automated reductive amination.<sup>14</sup> Without further optimization of the sequences assembly or the purification, stable *N*-Hnb-Cys peptides **1a-5a**<sup>22</sup> were isolated after TFA cleavage from the resin and HPLC purification up to 99% purity, with yields ranging from 10 to 21% (see supporting information for chromatograms of crude and purified compounds). These non-optimized SPPS yields are impaired by formation of side products showing an extra mass of 56 Da consistent with *t*Bu adducts frequently encountered for cysteine-rich peptide synthesis,<sup>11,12</sup> but also truncated peptides arising from incomplete couplings unrelated to the methodology, as well as, in the case of **2a**, methionine sulfoxide formation (+16 Da). Note that the intermediary purification of the linear precursors is much probably not necessary for our synthetic scheme, but we wanted here to carefully assess for the efficacy of *N*-Hnb-Cys cryptothioesters for intramolecular NCL through rigorous analysis of the reaction mixtures.

NCL-based cyclization of pure peptides **1a-5a** were then conducted under non-optimized dilute conditions (1 mM peptide concentration) using 25 mM 4-mercaptophenylacetic acid (MPAA) and 50 mM *tris*(2-carboxyethyl)phosphine (TCEP) at 37 °C under argon atmosphere (Table 1). All five reactions showed 90% or higher HPLC conversion after 6 h; in cases of incomplete reactions, samples were incubated overnight to achieve quantitative conversions. Importantly, no side reactions that could be inherent to the *N*-Hnb-Cys



**Fig. 2** Synthesis of naturally occurring cyclic peptide sequences through intramolecular NCL cyclization of *N*-Hnb-Cys peptides **1a-5a**. TG : Tentagel resin ; aa : amino acid.<sup>23</sup>

methodology were observed.<sup>24</sup> In particular, no significant amount of cryptothioester hydrolysis could be detected in the crude mixtures, such a side-reaction being relatively common in NCL-based syntheses.

Cyclodimerization is a frequently encountered problem during macrocycle formation. minor co-products consistent with cyclic dimers **1c** (cO2), **3c** (kB1) and **5c** (SFTI-1) were identified from the LC/MS analyses. **1c** was not detectable by HPLC, whereas the amounts of **3c** and **5c** could be quantified (4 and 5%, respectively) by integration of the UV chromatograms. Gratifyingly, when NCL-cyclizations of **3a** and **5a** were performed at 0.1 mM peptide concentrations in order to disfavor intermolecular reactions, cyclodimerization was almost completely abolished in both cases without, as expected, observing significant changes in cyclization kinetic rates.

Providing this slight adjustment of reaction conditions in the case of **3a** and **5a**, all five reactions proceeded extremely cleanly as illustrated in Figure 3 by the analytical HPLC monitoring of NCL-cyclization of **2a** giving **2b** (Cter M) (see supporting information for HPLC profiles of the four other crude cyclic peptides after reaction work-ups). Finally, reduced cyclic peptides **1b-5b** were isolated through HPLC purification with excellent yields ranging from 85 to 97%.

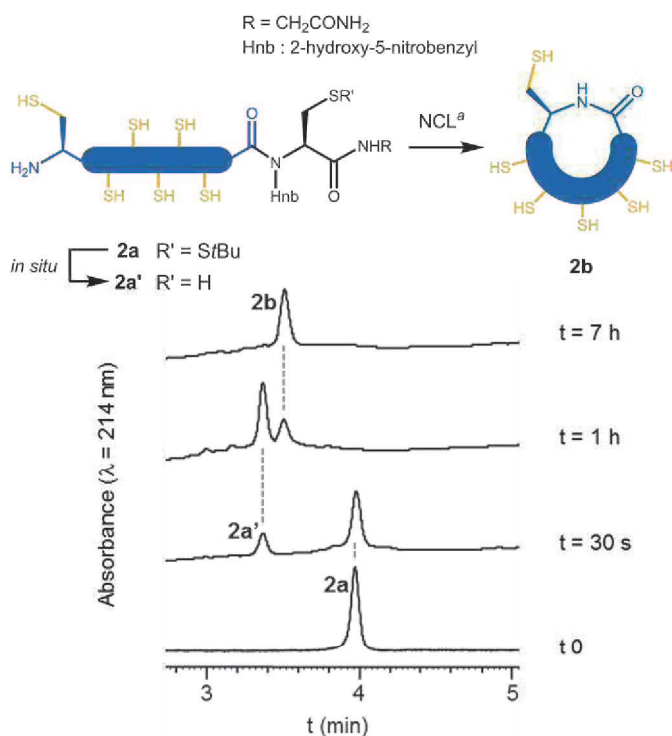
In summary, we demonstrated that *N*-Hnb-Cys cryptothioesters are efficient substrates for the NCL-based backbone cyclization of naturally occurring cysteine-rich peptides. Cyclization of *N*-Hnb-Cys linear precursors operates rapidly and cleanly at low concentrations, affording the target cyclic peptides with excellent yields. This simple strategy may have a high impact for applications to drug discovery using cDRP



**Table 1** Intramolecular NCL cyclization of *N*-Hnb-Cys peptides **1a-5a**.

Linear precursor (Yield %) <sup>a, b</sup>	Concentration (mM) <sup>c</sup>	Time (h)	Yield <b>1b-5b</b> (%) <sup>d</sup>	Cyclic dimer <b>1c-5c</b> (%) <sup>e</sup>
<b>1a</b> (18)	1	24 <sup>f</sup>	97	< 0.5 <sup>g</sup>
<b>2a</b> (10)	1	7	94	0 <sup>h</sup>
<b>3a</b> (16)	1	22 <sup>f</sup>	nd <sup>i</sup>	4
	0.1	24 <sup>f</sup>	85	0 <sup>h</sup>
<b>4a</b> (11)	1	6	90	0 <sup>h</sup>
<b>5a</b> (21)	1	24 <sup>f</sup>	nd <sup>i</sup>	5
	0.1	24 <sup>f</sup>	91	< 0.5 <sup>g</sup>

<sup>a</sup>Linear precursors were synthesized through standard Fmoc-SPPS and automated reductive amination.<sup>14</sup> <sup>b</sup>Isolated yields of HPLC-purified **1-5a** peptides determined from initial resin loading. <sup>c</sup>All reactions were performed at 37 °C in a 0.2 M pH 6.5 phosphate buffer, in the presence of 25 mM TCEP, 50 mM MPAA and 6 M guanidinium chloride. <sup>d</sup>Isolated yield, HPLC purification. <sup>e</sup>Quantified by HPLC, not isolated. <sup>f</sup>HPLC yield between 90 and 95% after 6h. <sup>g</sup>Traces detected by LC-MS analysis. <sup>h</sup>Not detected by LC-MS analysis. <sup>i</sup>> 90% HPLC yields, not isolated. nd: not determined; MPAA: 4-mercaptophenylacetic acid; TCEP: tris(2-carboxyethyl)phosphine.



**Fig. 3** Analytical HPLC monitoring of the cyclization of **2a** through intramolecular NCL. <sup>a</sup>Reaction was performed under the conditions described in table 1

scaffold, considering the straightforward automated synthesis *N*-Hnb-Cys crypto-thioesters amenable to high-throughput parallel synthesis.

We thank Dr Guillaume Gabant, Dr Cyril Colas and the mass spectrometry platforms of CBM and FR2708 federation for the

MS and LC/MS analyses. We also thank Ms. Sandra Petinon, intern student who worked on this project. The Pierre Potier foundation for the chemistry of natural products is gratefully acknowledged for a PhD fellowship for VPT.

## Notes and references

- M. Cemazar, S. Kwon, T. Mahatmanto, A. S. Ravipati and D. J. Craik, *Curr. Top. Med. Chem.*, 2012, **12**, 1534-1545.
- D. J. Craik, *Science*, 2006, **311**, 1563-1564.
- P.E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776-779.
- a) L. Zhang and J. P. Tam, *J. Am. Chem. Soc.*, 1997, **119**, 2363-2370; b) J. A. Camarero and T. W. Muir, *Chem. Commun.*, 1997, 1369-1370; c) J. A. Camarero, G. J. Cotton, A. Adeva and T. W. Muir, *J. Pept. Res.*, 1998, **51**, 303-316.
- a) N. L. Daly, S. Love, P. F. Alewood and D. J. Craik, *Biochemistry*, 1999, **38**, 10606-10614; b) T. L. Aboye, Y. Li, S. Majumder, J. Hao, A. Shekhtman and J. A. Camarero, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2823-2826.
- a) J. P. Tam and Y-A. Lu, *Tetrahedron Lett.*, 1997, **38**, 5599-5602; b) J. P. Tam and Y-A. Lu, *Protein Sci.*, 1998, **7**, 1583-1592; c) J. A. Camarero, J. Pavel and T. W. Muir, *Angew. Chem. Int. Ed.*, 1998, **37**, 347-349.
- F. Mende and O. Seitz, *Angew. Chem. Int. Ed.*, 2011, **50**, 1232-1240.
- a) P. Thongyoo, E. W. Tate and R. J. Leatherbarrow, *Chem. Commun.*, 2006, 2848-2850; b) P. Thongyoo, N. Roqué-Rosell, R. J. Leatherbarrow and E. W. Tate, *Org. Biomol. Chem.*, 2008, **6**, 1462-1470.
- a) J-S. Zheng, S. Tang, Y. Guo, H-N. Chang and L. Liu, *ChemBioChem*, 2012, **13**, 542-546; b) Y-Q. Chen, C-C. Chen, Y. He, M. Yu, L. Xu, C-I. Tian, Q-X. Guo, J. Shi, M. Zhang and Y-M. Li, *Tetrahedron Lett.*, 2014, **55**, 2883-2886.
- M. Taichi, X. Hemu, Y. Qiu and J. P. Tam, *Org. Lett.*, 2013, **15**, 2620-2623.
- S. Gunasekera, T. L. Aboye, W. A. Madian, H. R. El-Seedi and U. Göransson, *Int. J. Pept. Res. Ther.*, 2013, **19**, 43-54.
- J. B. Blanco-Canosa, B. Nardone, F. Albericio and P. E. Dawson, *J. Am. Chem. Soc.*, 2015, **137**, 7197-7209.
- a) L. Shariff, Y. Zhu, B. Cowper, W-L. Di and D. Macmillan, *Tetrahedron*, 2014, **70**, 7675-7680; b) B. Cowper, L. Shariff, W. Chen, S. M. Gibson, W-L. Di and D. Macmillan, *Org. Biomol. Chem.*, 2015, **13**, 7469-7476.
- V. P. Terrier, H. Adihou, M. Arnould, A. F. Delmas and V. Aucagne, *Chem. Sci.*, 2015, DOI: 10.1039/C5SC02630J.
- T. L. Aboye, R. J. Clark, D. J. Craik and U. Göransson, *ChemBioChem*, 2008, **9**, 103-113.
- A. G. Poth, M. L. Colgrave, R. E. Lyons, N. L. Daly and D. J. Craik, *Proc. Natl. Acad. Sci. USA.*, 2011, **108**, 10127-10132.
- a) C. W. Gruber, A. G. Elliott, D. C. Ireland, P. G. Delprete, S. Dessein, U. Göransson, M. Trabi, C. K. Wang, A. B. Kinghorn, E. Robbrecht and D. J. Craik, *Plant Cell*, 2008, **20**, 2471-2483; b) L. Chiche, A. Heitz, J-C. Gelly, J. Gracy, P. T. T. Chau, P. T. Ha, J-F. Hernandez and D. Le-Nguyen, *Curr. Protein Pept. Sci.*, 2004, **5**, 341-349; c) A. G. Poth, M. L. Colgrave, R. Philip, B. Kerenga, N. L. Daly, M. A. Anderson and D. J. Craik, *ACS Chem. Biol.*, 2011, **6**, 345-355.
- D. J. Craik, N. L. Daly, T. Bond and C. Waine, *J. Mol. Biol.*, 1999, **294**, 1327-1336.
- S. Lockett, R. Santiago Garcia, J. J. Barker, Al. V. Konarev, P. R. Shewry, A. R. Clarke and R. L. Brady, *J. Mol. Biol.*, 1999, **290**, 525-533.
- A. C. Conibear and D. J. Craik, *Angew. Chem. Int. Ed.*, 2014, **53**, 10612-10623.

- 21 Y-Q. Tang, J. Yuan, G. Ösapay, K. Ösapay, D. Tran, C. J. Miller, A. J. Ouellette and M. E. Selsted, *Science*, 1999, **286**, 498-502.
- 22 From a retrosynthetic point of view, due to multiple occurrences of Cys in the sequences of the target cDRP, many Xaa-Cys disconnections could have been envisaged. We deliberately focused on disconnections that exclude NCL involving disfavoured Xaa such as Val, Ile, Thr and Pro.
- 23 Oxidative folding of these targets to selectively form the disulfide bridges is well documented, and we accordingly limited the present study to their backbone cyclization. For examples, for folding of KB1, see refs 5a, 9a, 10, 11 or 12; for folding of cO2, ref 9a or 17; for folding of Cter M, see ref 18; for folding of SFTI-1, see ref 9b,11 or 13; for folding of RTD-1, see ref 5b or 20.
- 24 Aspartimide formation at the Asp-Gly site has recently been described as a problematic side-reaction observed during the NCL-based cyclization of the SFTI-1 sequence (see reference 13b). Note that in our case, an aspartimide by-product could only be detected as traces amount from the LC/MS analysis of crude **5a** obtained through a 0.1 mM peptide concentration NCL.

## Références bibliographiques

1. Cezamar M.; Kwon S.; Mahatmanto T.; Ravipati A. S.; Craik D. J., Discovery and applications of disulfide-rich cyclic peptides. *Curr. Top. Med. Chem.*, **2012**, *12*, 1534-1545.
2. Craik D. J., Seamless proteins tie up their loose ends. *Science*, **2006**, *311*, 1563-1564.
3. Clark R. J.; Fischer H.; Dempster L.; Daly N. L.; Rosengren K. J.; Nevin S. T.; Meunier F. A.; Adams D. J.; Craik D. J., Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the  $\alpha$ -conotoxin MII. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 13767-13772.
4. Akcan M.; Stroud M. R.; Hansen S. J.; Clark R. J.; Daly N. L.; Craik D. J.; Olson J. M., Chemical re-engineering of chlorotoxin improves bioconjugation properties for tumor imaging and targeted therapy. *J. Med. Chem.*, **2011**, *54*, 782-787.
5. Albericio F.; Hammer R. P.; García-Echevarría C.; Molins M. A.; Chang J. L.; Munson M. C.; Pons M.; Giralt E.; Barany G., Cyclization of disulfide-containing peptides in solid-phase synthesis. *Int. J. Pept. Protein Res.*, **1991**, *37*, 402-413.
6. Lambert J. N.; Mitchell J. P.; Roberts K. D., The synthesis of cyclic peptides. *J. Chem. Soc., Perkin Trans. 1*, **2001**, 471-484.
7. Zhang L.; Tam J. P., Synthesis and application of unprotected cyclic peptides as building blocks for peptide dendrimers. *J. Am. Chem. Soc.*, **1997**, *119*, 2363-2370.
8. Camarero J. A.; Muir T. W., Chemoselective backbone cyclization of unprotected peptides. *Chem. Commun.*, **1997**, 1369-1370.
9. Camarero J. A.; Cotton G. J.; Adeva A.; Muir T. W., Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.*, **1998**, *51*, 303-316.
10. 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*, 130-133.
11. Tam J. P.; Lu Y-A., Synthesis of large cyclic cystine-knot peptide by orthogonal coupling strategy using unprotected peptide precursors. *Tetrahedron Lett.*, **1997**, *38*, 5599-5602.
12. Camarero J. A.; Pavel J.; Muir T. W., Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization. *Angew. Chem. Int. Ed.*, **1998**, *37*, 347-349.
13. Daly N. L.; Love S.; Alewood P. F.; Craik D. J., Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide Kalata B1. *Biochemistry*, **1999**, *38*, 10606-10614.
14. Aboye T. L.; Li Y.; Majumder S.; Hao J.; Shekhtman A.; Camarero J. A., Efficient one-pot cyclization/folding of Rhesus  $\theta$ -defensin-1 (RTD-1). *Bioorg. Med. Chem. Lett.*, **2012**, *22*, 2823-2826.
15. Thongyoo P.; Tate E. W.; Leatherbarrow R. J., Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II. *Chem. Commun.*, **2006**, 2848-2850.

16. Thongyoo P.; Roqué-Rosell N.; Leatherbarrow R. J.; Tate E. W., Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.*, **2008**, *6*, 1462-1470.
17. Zheng J-S.; Tang S.; Guo Y. Chang H-N.; Liu L., Synthesis of cyclic peptides and cyclic proteins via ligation of peptide hydrazides. *ChemBioChem*, **2012**, *13*, 542-546.
18. Chen Y-Q.; Chen C-C.; He Y.; Yu M.; Xu L.; Tian C-I.; Guo Q-X.; Shi J.; Zhang M.; Li Y-M., Efficient synthesis of trypsin inhibitor SFTI-1 via intramolecular ligation of peptide hydrazide. *Tetrahedron Lett.*, **2014**, *55*, 2883-2886.
19. Gunasekera S.; Aboye T. L.; Madian W. A.; El-Seedi H. R.; U Göransson, 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*, 43-54.
20. 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*, 7197-7209.
21. Taichi M.; Hemu X.; Qiu Y.; Tam J. P., A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift. *Org. Lett.*, **2013**, *15*, 2620-2623.
22. Shariff L.; Zhu Y.; Cowper B.; Di W-L.; Macmillan D., Sunflower trypsin inhibitor (SFTI-1) analogues of synthetic and biological origin via *N*→*S* acyl transfer: potential inhibitors of human Kallikrein-5 (KLK5). *Tetrahedron*, **2014**, *70*, 7675-7680.
23. Cowper B.; Shariff L.; Chen W.; Gibson S. M.; Di W. L.; Macmillan D., Expanding the scope of *N* → *S* acyl transfer in native peptide sequences. *Org. Biomol. Chem.*, **2015**, *13*, 7469-7476.
24. Craik D. J.; Daly N. L.; Bond T.; Wayne C., Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.*, **1999**, *294*, 1327-1336.
25. Aboye T. L.; Clark R. J.; Craik D. J.; Göransson U., Ultra-stable peptide scaffolds for protein engineering-synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. *ChemBioChem*, **2008**, *9*, 103-113.
26. Poth A. G.; Colgrave M. L.; Lyons R. E.; Daly N. L.; Craik D. J., Discovery of an unusual biosynthetic origin for circular proteins in legumes. *Proc. Natl. Acad. Sci. USA*, **2011**, *108*, 10127-10132.
27. Luckett S.; Santiago Garcia R.; Barker J. J.; Konarev Al. V.; Shewry P. R.; Clarke A. R.; Brady R. L., High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.*, **1999**, *290*, 525-533.
28. Conibear A. C.; Craik D. J., The chemistry and biology of theta defensins. *Angew. Chem. Int. Ed.*, **2014**, *53*, 10612-10623.
29. Tang Y-Q.; Yuan J.; Ösapay G.; Ösapay K.; Tran D.; Miller C. J.; Ouellette A. J.; Selsted M. E., A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. *Science*, **1999**, *286*, 498-502.

CONCLUSIONS  
ET  
PERSPECTIVES



## Conclusions et perspectives

La synthèse de peptides et protéines par voie chimique, qui connaît encore actuellement une vive expansion, s'est imposé comme une alternative séduisante à la production par voie recombinante. La synthèse chimique a permis de repousser les limites de l'étude de ces macromolécules. Cela a été rendu possible principalement par deux avancées majeures : premièrement, la synthèse peptidique sur support solide, permettant de populariser la préparation en routine de peptides d'une taille modérée. L'introduction de méthodes de ligation permettant d'augmenter significativement la taille des séquences pouvant être produites représente la seconde avancée. La ligation chimique native (NCL) a démontré un fort potentiel pour coupler, de façon régio- et chimiosélective, deux fragments peptidiques déprotégés, dans le but d'outrepasser la limitation en taille de la SPPS. Cette stratégie a révolutionné ce domaine mais la démocratisation de la NCL est encore limitée par la synthèse délicate d'un des partenaires clés de cette réaction : les peptides C $\alpha$ -thioester. Par ailleurs, les peptides thioester jouent également un rôle central dans plusieurs autres techniques de ligation.

La majeure partie de ma thèse a été focalisée autour de la conception et l'optimisation d'une nouvelle méthode de synthèse bio-inspirée de peptides thioester, ainsi que sur l'étude du potentiel et des limitations de cette approche.

La stratégie que nous avons développée repose sur l'utilisation d'un dispositif de thioestérification C-terminal, consistant en un motif *N*-(2-hydroxy-5-nitrobenzyl)cystéine (*N*-Hnb-Cys), opérant par un transfert d'acyle intramoléculaire *N*→*S*. Ce dispositif simple peut être assemblé sur support solide, de façon totalement automatisée, à partir de réactifs à très bas cout (le motif Hnb s'installe sur une cystéine à l'aide de 10 équivalents d'un aldéhyde coutant moins de 2 €/mmol et de 20 équivalents d'un réducteur coutant moins de 20 cents/mmol). De plus, l'installation de la séquence peptidique sur ce dispositif est, elle aussi, automatisée; grâce à l'assistance de la *N*-acylation de notre dispositif, pour introduire l'acide aminé C-terminal, par un mécanisme de transfert d'acyle intramoléculaire *O*→*N*. Enfin, les peptides portant notre dispositif peuvent être directement utilisés dans une NCL sans nécessiter d'étapes post-SPPS. Nous sommes convaincus que ces avantages constituent une avancée majeure et peuvent inciter des personnes non-spécialistes à utiliser la méthode.

Des études de NCLs avec des peptides modèles composés du dispositif *N*-Hnb-Cys ont permis d'explorer, de façon préliminaire, le mécanisme de cette réaction. Grâce à sa conception

## Conclusions et perspectives

bio-inspirée, le dispositif de thioestérification opère rapidement à pH neutre dans les conditions de NCL. L'étude des cinétiques de réactions modèles ont mis en évidence le rôle crucial que joue le groupement phénol de notre dispositif dans l'étape de thioestérification, appuyant notre hypothèse d'une participation de ce groupement par catalyse acido-basique intramoléculaire, inspirée de celle observée dans la nature chez les intéines.

Nous avons exploré la portée de notre approche par son application à la synthèse des formes réduites de plusieurs peptides naturels riches en ponts disulfure. Notre méthodologie a permis d'obtenir des DRPs linéaires de grande taille à partir de longs peptides *N*-Hnb-Cys (parfois plus de 50 acides aminés). Il faut noter que nous avons décrit le premier exemple de synthèse par voie chimique de la défensine *Cg*-BigDef1, son repliement oxydatif ayant été réalisé avec succès au sein de l'équipe. Nous avons également décrit une nouvelle stratégie de préparation de peptides contenant une cystéine C-terminale, permettant d'éviter les réactions secondaires observées lors de l'utilisation de l'approche historique. Enfin, l'application de notre méthode à la cyclisation de squelettes peptidiques de DRPs a conduit à d'excellents rendements. Ces différents exemples démontrent ainsi le fort potentiel de la stratégie des peptides *N*-Hnb-Cys.

Ce travail de thèse donne lieu à de nombreuses perspectives :

Une exploration théorique du mécanisme de la thioestérification impliquant notre dispositif, à l'aide de modélisation moléculaire par des outils informatiques, est en cours pour d'une part, étudier plus en détails ce mécanisme et permettre de déterminer les différents états de transition et d'autre part, tenter de proposer de façon rationnelle des modifications de la structure du dispositif qui pourraient permettre d'encore accélérer les cinétiques réactionnelles. Une augmentation de la vitesse de réaction, d'un facteur deux ou trois, pourrait être bénéfique, notamment dans le cas de longs peptides difficiles à solubiliser, nécessitant leur emploi à de très faibles concentrations.

Enfin, le développement de stratégies permettant de réaliser des NCLs successives avec les peptides *N*-Hnb-Cys, pour conduire à l'obtention de très longues chaînes polypeptidiques, est une étape future importante. La compatibilité évidente, mais néanmoins non prouvée, avec l'utilisation de protection de cystéine N-terminale (par exemple sous forme de thiazolidine) sur un peptide portant notre dispositif permettrait l'application de cette stratégie à des NCLs successives de C vers N. De façon importante, nous avons vu qu'il est également possible de désactiver le dispositif *N*-Hnb-Cys par protection de la cystéine à l'aide d'un groupement stable



## Conclusions et perspectives

aux conditions de SPPS (traitement TFA inclus) et aux conditions de NCL (Acm par exemple), ce qui rend théoriquement possible les NCLs successives de N vers C, à condition que la déprotection de ce groupement au moment voulu soit efficace. Cette approche originale, car encore peu développée, fait actuellement l'objet de développements au sein de notre laboratoire.



# PARTIES EXPERIMENTALES



# Partie expérimentale article 1 :

## A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters

Terrier V. P.; Adihou H.; Arnould M.; Delmas A. F.; Aucagne V., *Chemical Science*, **2016**, 7, 339-345.

### Supporting Information

#### Table of contents:

1) General information	166
2) General procedures for solid phase peptide synthesis	167
3) Optimization of the benzyl group to maximize the <i>N</i> -acylation yield	168
a- Synthesis of cysteinyl peptide resin <b>1</b>	168
b- Synthesis of peptide resins <b>2a-d</b>	169
c- Study on the solid-supported <i>N</i> -acylation of <b>2a-d</b>	173
4) Evaluation of the phenol pKa of an <i>N</i> -acyl- <i>N</i> -(Hnb)Cys compound <b>S3</b>	176
5) Characterization of a model <i>N</i> -acyl- <i>N</i> -(Hnb)Cys compound <b>S5</b>	178
a- NMR of <b>S5</b>	178
b- Molar extinction coefficient of the Hnb group of <b>S5</b>	188
6) Optimization of the cysteine thiol protecting group	188
a- Synthesis and stability of ( <i>S</i> -Trt) model peptide <b>S7</b>	188
b- Synthesis and stability of ( <i>S</i> -StBu) model peptide <b>7</b>	192
7) Optimization of the spacer between the cysteine and the resin	194
a- Direct attachment to Rink linker – amide hydrolysis co-product	194
b- Direct attachment to PHB Tentagel resin (Wang type linker)	195
c- Introduction of a Gly residue as a spacer	196
8) <i>N</i> -Acylation with the 20 different proteogenic amino acids	198
9) Kinetics studies of the NCL with model Ac-LYRAA(Hnb)C peptides	222
a- Synthesis of model cysteinyl peptide <b>6</b>	222
b- Influence of the nature of the cysteine C-terminus	223
c- Study of the hydrolysis of the thioester precursor <b>9</b> and epimerization	226
d- Influence of the buffer	228
e- Influence of masking phenol or thiol groups	229
f- Influence of the pH	235
10) Influence of the C-terminal aa: NCL with model Ac-LYRAX(Hnb)C peptides	237
a- Synthesis of peptide Ac-LYRAS-(Hnb)C(StBu)G-NH <sub>2</sub> <b>10</b>	237
b- NCL with peptide Ac-LYRAS-(Hnb)C(StBu)G-NH <sub>2</sub> <b>10</b>	238
c- Synthesis of peptide Ac-LYRAV-(Hnb)C(StBu)G-NH <sub>2</sub> <b>11</b>	239
d- NCL with peptide Ac-LYRAV-(Hnb)C(StBu)G-NH <sub>2</sub> <b>11</b>	241
e- NCL with peptide <b>11</b> under optimized conditions	242
f- Determination of apparent second order kinetic constants	243

11) Application to MT7	246
a- Synthesis of cysteinyl peptide <b>19</b>	246
b- Synthesis of crypto-thioester peptide <b>18</b>	249
c- NCL	252
12) Application to Cg-Bigdef1	254
a- Synthesis of cysteinyl peptide <b>22</b>	254
b- Synthesis of crypto-thioester peptide <b>21</b>	256
c- NCL	259

## 1- General information

All reagents and solvents were used without further purification. Protected amino acids, Fmoc-Gly-Wang resin, Fmoc-Rink polystyrene resin, Rink linker, HBTU and HCTU were purchased from Merck Biosciences (Nottingham, UK). Aminomethyl TentaGel R resin and pre-loaded Fmoc-Cys(StBu)-PHB (Wang type) Tentagel resin were purchased from Rapp polymers (Tuebingen, Germany). Fmoc-K(Boc)-Mppa-OH was purchased from Polypeptide laboratories (Strasbourg, France). 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).

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker AVANCE III 600 instrument, at a constant temperature of 25°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.

High resolution ESI-MS analyses were performed on a maXis ultra-high-resolution Q-TOF 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  $\alpha$ -cyano-4-hydroxy-cinnamic acid (HCCA) as a matrix. The reported  $m/z$  values correspond to the monoisotopic ions if not specified otherwise. Peptides incorporating an *N*-(2-hydroxy-5-nitrobenzyl) group displayed a typical MALDI fragmentation pattern, consisting in -16 Da and -34 Da minor peaks in addition to the  $[\text{MH}]^+$  peak.

HPLC analyses and 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 autosampler. Nucleosil C18 (300 Å, 5  $\mu\text{m}$ , 250  $\times$  4.6 mm, 1 mL/min flow rate) or Chromolith HighResolution RP-18e (150 Å, 10  $\times$  4.6 mm, 3 mL/min flow rate) columns were used for analysis and Nucleosil C18 (300 Å, 5  $\mu\text{m}$ ,

250 × 10 mm, 3 mL/min flow rate) for purification. Solvents A and B are 0.1 % TFA in H<sub>2</sub>O and 0.1 % TFA in MeCN, respectively. Each gradient was followed by a washing step (95% B/A over 0.5 min for the HR Chromolith; over 1 min for the Nucleosil C18 column) to identify eventual co-products not eluted during the gradient.

LC-MS analyses were carried out on a Ultimate<sup>®</sup> 3000 RSLC HPLC system (Dionex, Germering, Germany), coupled with the maXis mass spectrometer and fitted with a Zorbax 300 SB-C18 RRHD (300 Å, 1.8 μm, 100 × 2.1 mm, 0.3 mL/min flow rate, 40°C) column. Solvents A and B were 0.1 % formic acid in H<sub>2</sub>O and 0.08 % formic acid in MeCN, respectively. Gradient: 3 % B/A for 1 min, then 3 to 50 % B/A over 19 min.

## 2- General procedures for solid phase peptide synthesis

Fmoc-based solid phase peptide syntheses (SPPS) were carried out on either a Prelude synthesizer from Protein Technologies or a 433A synthesizer from Applied Biosystems. Microwave-assisted Fmoc-SPPS syntheses were carried out on an Initiator+ SP Wave synthesizer from Biotage.

The side-chain protecting groups used were Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Acm), Cys(Trt), Cys(S<sub>t</sub>Bu), Glu(OtBu), Gln(Trt), His(Trt), Lys(Boc), Ser(*t*Bu), Thr(*t*Bu), Trp(Boc) and Tyr(*t*Bu).

Syntheses starting from aminomethyl TentaGel R resin were performed on the Prelude synthesizer, on a 0.025 mmol per reactor scale. Protected amino acids (0.25 mmol, 10 equiv.) were coupled using HCTU (98 mg, 0.238 mmol, 9.5 equiv.) and *i*Pr<sub>2</sub>NEt (87 μL, 0.5 mmol, 20 equiv.) in NMP (3 mL) for 30 min. Capping of eventual unreacted amine groups was achieved by treatment with acetic anhydride (143 μL, 1.51 mmol, 60 equiv.), *i*Pr<sub>2</sub>NEt (68 μL, 0.39 mmol, 15.5 equiv.) and HOBt (6 mg, 0.044 mmol, 1.8 equiv.) in NMP (3 mL) for 7 min. Fmoc group was deprotected by three successive treatments with 20 % piperidine in NMP (3 mL) for 3 min. In the cases where the coupling yield of an amino acid on the Hnb device is low, a washing step of the peptide-resin using a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in a NMP / CH<sub>2</sub>Cl<sub>2</sub> mixture<sup>240</sup> (5:1, 20 min, 3 mL, ×3) is performed before the capping step.

Syntheses starting from Fmoc-Gly-Wang or Rink polystyrene resins were performed on the 433A synthesizer, using the 0.1 mmol scale Fastmoc program purchased from the manufacturer, with a single coupling with HCTU followed by capping.

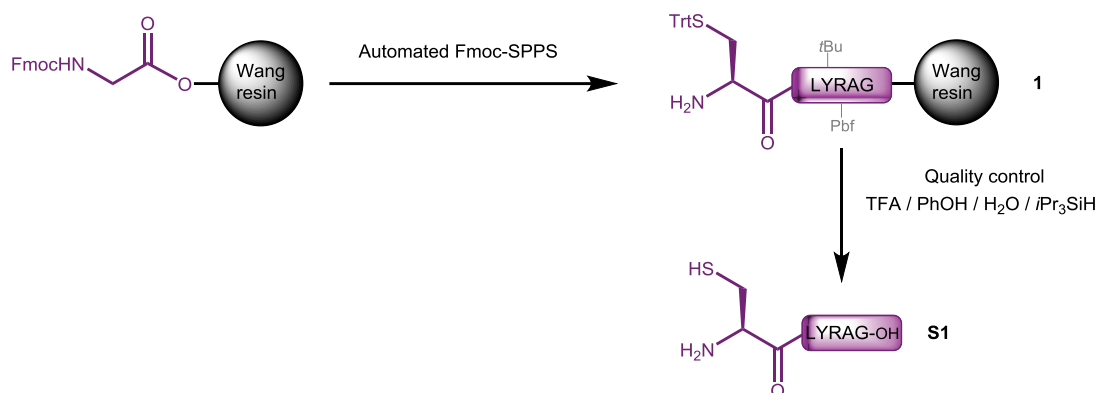
The crude peptide was deprotected and cleaved from the resin through a treatment with TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH/phenol, 88/5/2/5 for 2 h, and the peptide was precipitated by dilution into an ice-cold 1:1 diethyl ether/petroleum ether mixture, recovered by centrifugation and washed twice with diethyl ether.

---

<sup>240</sup> Diaz-Mochon J. J., Bialy L., Bradley M., *Org. Lett.*, 2004, 6, 7, 1127-1129

### 3- Optimization of the benzyl group to maximize the *N*-acylation yield

#### 3a- Synthesis of cysteinyl peptide resin 1



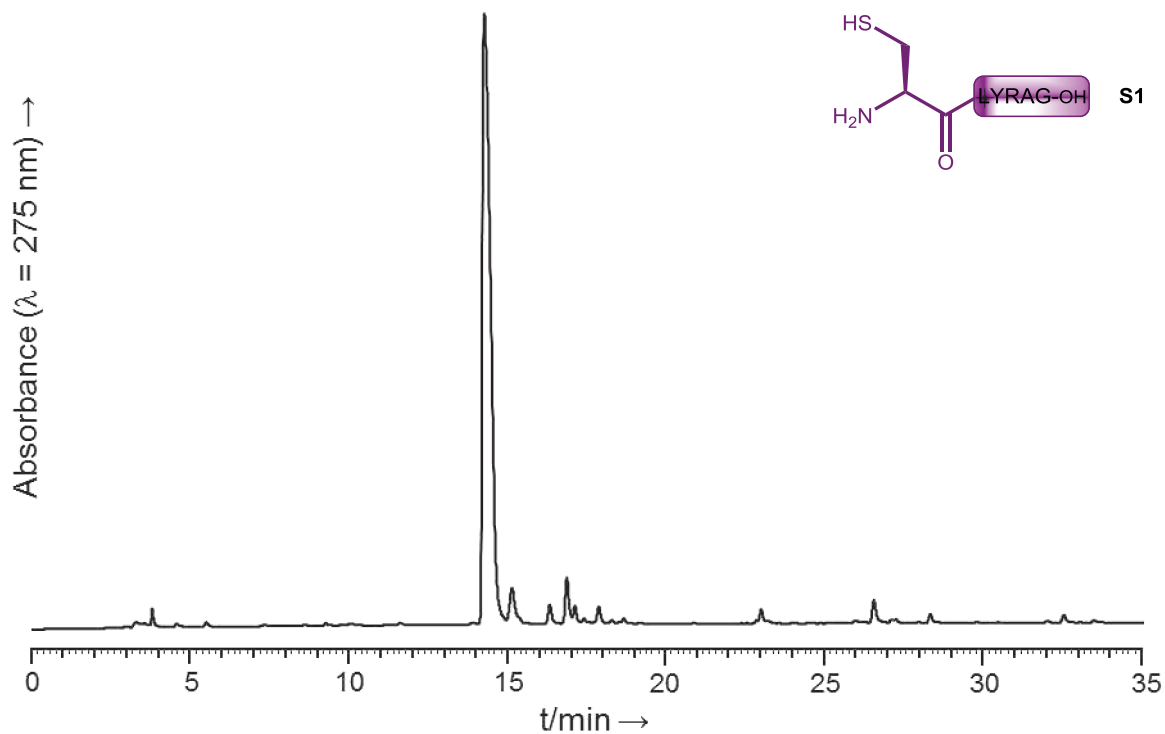
Supplementary scheme S1: Synthesis of peptide resin 1.

Peptide resin **1** was obtained through automated SPPS (protocol p S3) starting from Fmoc-Gly-Wang resin (130 mg, 0.79 mmol/g, 0.1 mmol). An aliquot of the resin was cleaved (protocol p S3) in order to characterize the corresponding peptide **S1**.

#### **S1:**

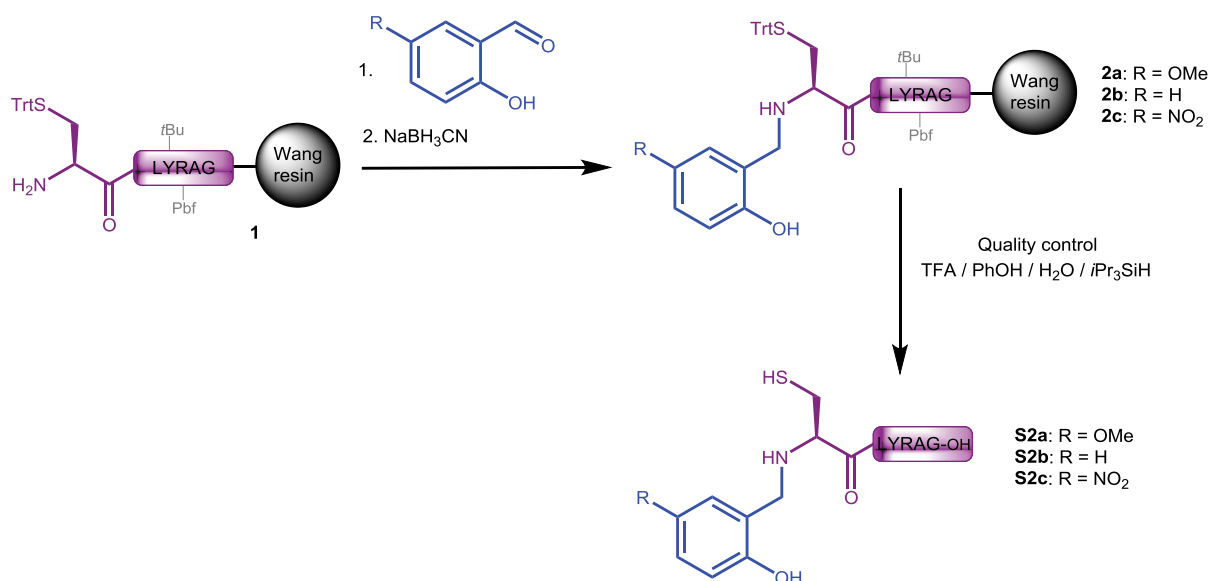
**MALDI-TOF MS** (*m/z*): [M+H]<sup>+</sup> calcd for C<sub>29</sub>H<sub>47</sub>N<sub>9</sub>O<sub>8</sub>S: 682.3, found: 682.3.

**HPLC**: retention time: 14.41 min (Nucleosil, gradient: 5-50 % B/A over 30 min).



Supplementary figure S1: HPLC trace of crude **S1**.



3b- Synthesis of peptide resins **2a-d** by solid-supported *N*-alkylation

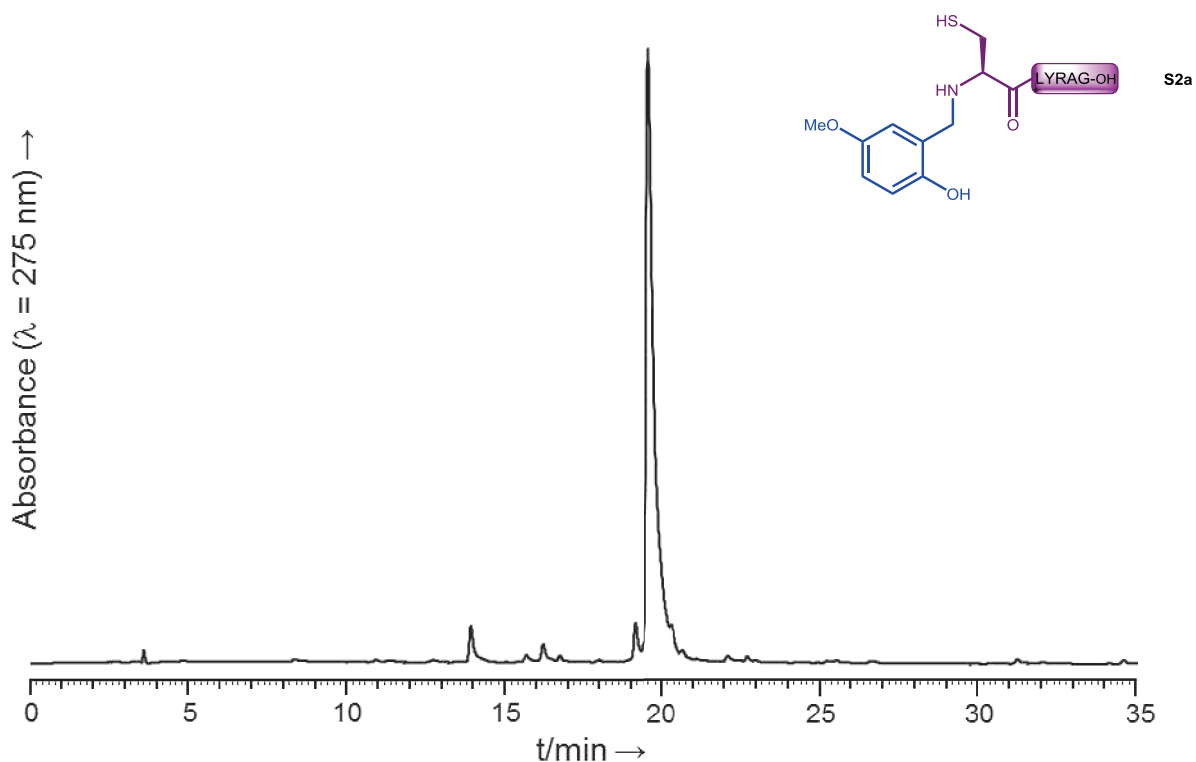
Supplementary scheme S2: Synthesis of *N*-benzyl peptide resins **2a-c** through solid-supported reductive amination.

Peptide resin **1** (10  $\mu$ mol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock, then swollen in a DMF/MeOH/AcOH (9:9:2) mixture for 5 min. The syringe was drained and the resin was washed with a DMF/MeOH (1:1) mixture. The appropriate aldehyde (100  $\mu$ mol, 10 equiv.) dissolved in DMF/MeOH (0.4 mL) was added then the syringe was left for 45 min under gentle stirring. The syringe was drained off and the resin was washed with DMF/MeOH. Without delay, sodium cyanoborohydride (13 mg, 200  $\mu$ mol, 20 equiv.) dissolved in DMF/MeOH/AcOH (0.4 mL) was added and the syringe was left for 30 min under gentle stirring. The syringe was drained and the resin was thoroughly washed with DMF/MeOH/AcOH, NMP, 5% *i*Pr<sub>2</sub>NEt in NMP then NMP. An aliquot of each peptide resin **2a-c** was cleaved (protocol p S3) in order to characterize the corresponding peptides **S2a-c**.

**S2a:**

**MALDI-TOF** ( $m/z$ ): [MH]<sup>+</sup> calcd for C<sub>37</sub>H<sub>55</sub>N<sub>9</sub>O<sub>10</sub>S: 818.4, found: 818.4.

**HPLC:** retention time: 19.59 min (Nucleosil, gradient: 5-50 % B/A over 30 min).

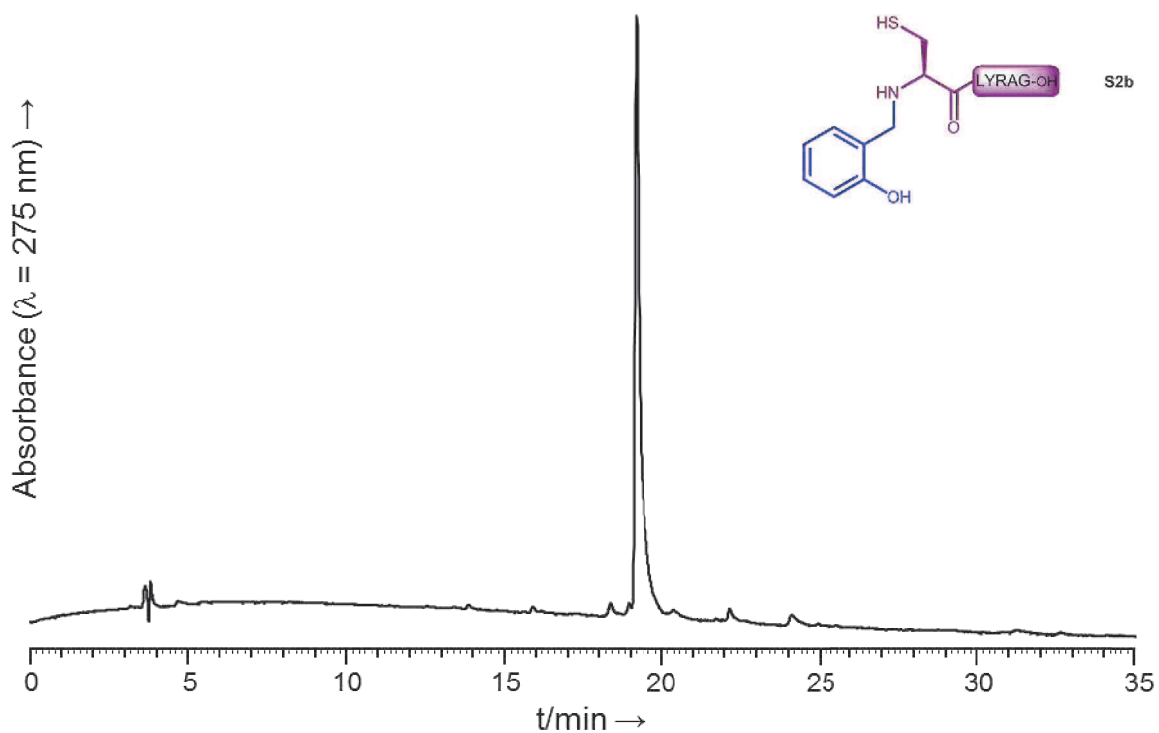


Supplementary figure S2: HPLC trace of crude **S2a**.

**S2b:**

**MALDI-TOF** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{36}\text{H}_{53}\text{N}_9\text{O}_9\text{S}$ : 788.4, found: 788.3.

**HPLC**: retention time: 19.17 min (Nucleosil, gradient: 5-50 % B/A over 30 min).

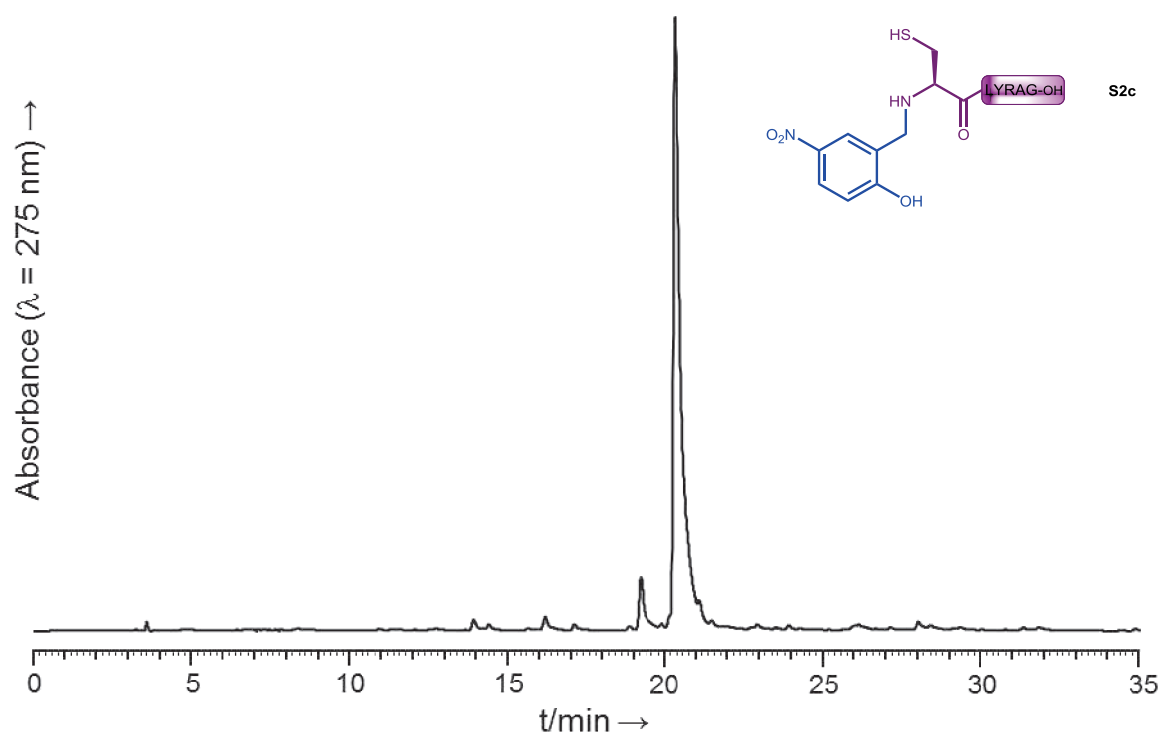


Supplementary figure S3: HPLC trace of crude **S2b**.

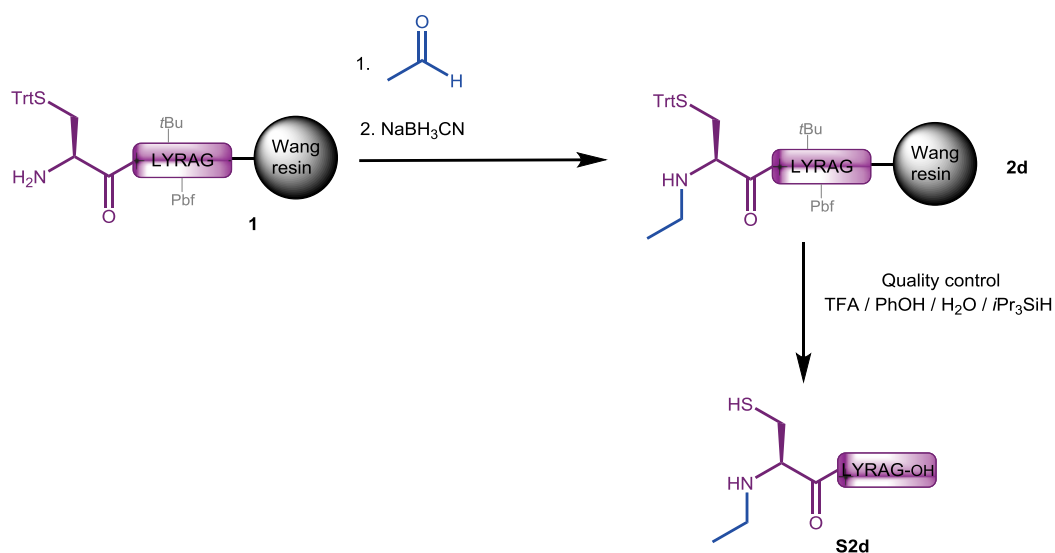
**S2c:**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{36}H_{52}N_{10}O_{11}S$ : 833.4, found: 833.3.

**HPLC**: retention time: 20.44 min (Nucleosil, gradient: 5-50 % B/A over 30 min).



Supplementary figure S4: HPLC trace of crude **S2c**.



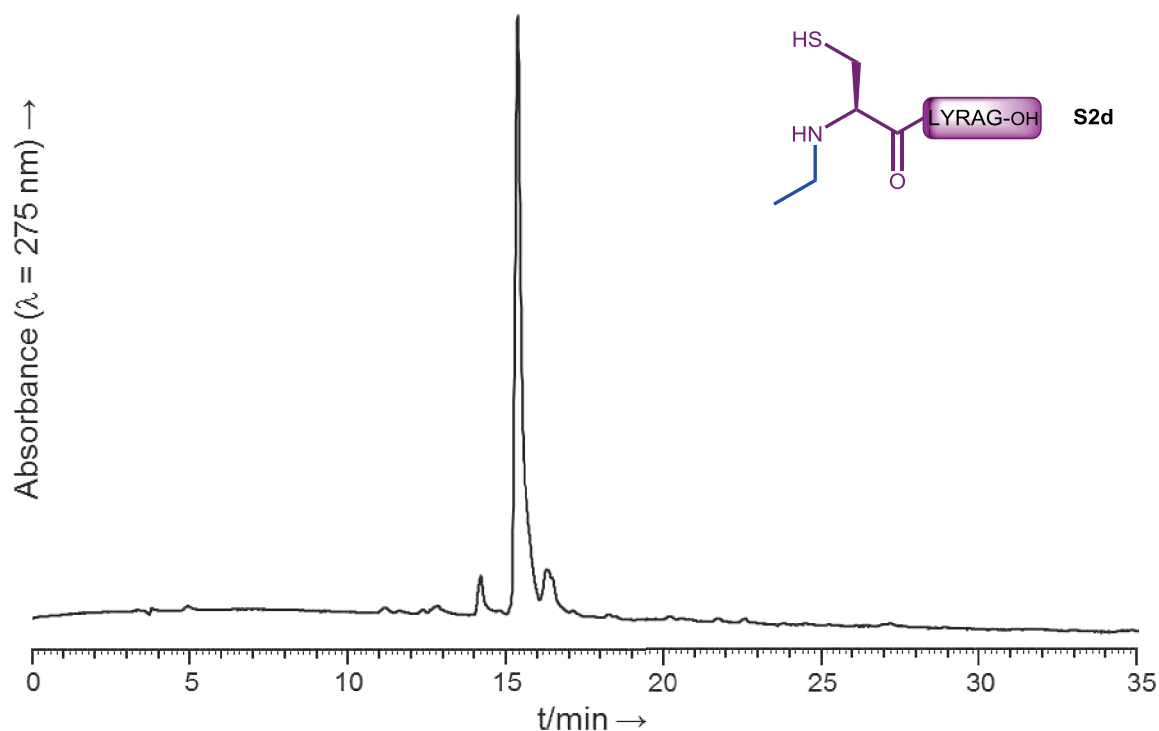
Supplementary scheme S3: Synthesis of *N*-ethyl peptide resin **2d** through solid-supported reductive amination.

Peptide resin **1** (10  $\mu\text{mol}$ , 1 equiv.) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen in a mixture of DMF/MeOH/AcOH (9:9:2) for 5 min. The syringe was drained and the resin was washed with a mixture of DMF/MeOH (1:1). Acetaldehyde (100  $\mu\text{mol}$ , 10 equiv.) dissolved in DMF/MeOH (0.4 mL) was added then the syringe was left for 5 min under gentle stirring. The syringe was drained and the resin was washed with DMF/MeOH. Without delay, sodium cyanoborohydride (13 mg, 200  $\mu\text{mol}$ , 20 equiv.) dissolved in DMF/MeOH/AcOH (0.4 mL) was added and the syringe was left for 30 min under gentle stirring. The syringe was drained and the resin was thoroughly washed with DMF/MeOH/AcOH, NMP, 5% *i*Pr<sub>2</sub>NEt in NMP then NMP. An aliquot of the resulting peptide resin was cleaved (protocol p S3) in order to characterize *N*-ethyl peptide **S2d**.

**S2d:**

**MALDI-TOF** ( $m/z$ ):  $[\text{MH}]^+$  calcd for C<sub>31</sub>H<sub>51</sub>N<sub>9</sub>O<sub>8</sub>S: 710.4, found 710.4.

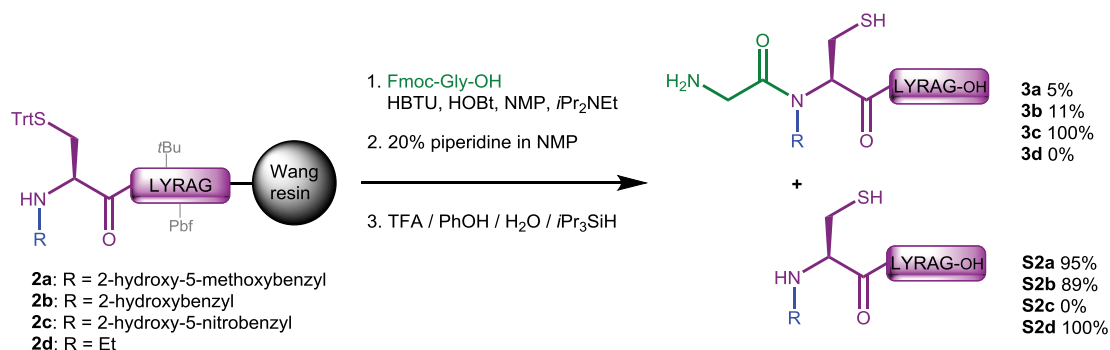
**HPLC**: retention time: 15.36 min (Nucleosil, gradient: 5-50 % B/A over 30 min).



Supplementary figure S5: HPLC trace of crude **S2d**.

3c- Study on the solid-supported *N*-acylation of secondary amines **2a-d**General procedure for solid-supported *N*-acylation

Peptide resin (**2a-d**) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Then, Fmoc-Gly-OH (10 equiv.), HBTU (9.5 equiv.) and HOBT (9.5 equiv.) were dissolved in NMP (0.1 M final concentration in amino acid) prior to addition of *i*Pr<sub>2</sub>NEt (20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 2 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. An aliquot resin was cleaved using the standard procedure (p S3) to determine the coupling yields and characterize the corresponding *N*-acylated peptides.

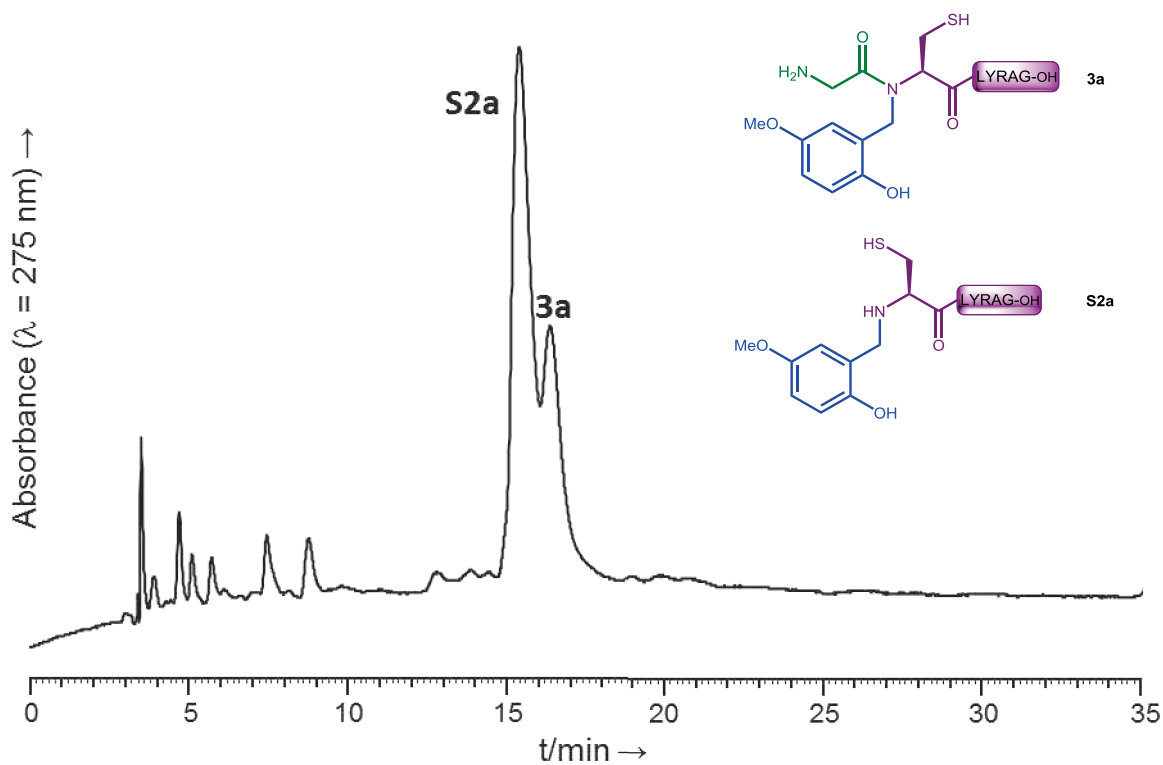


Supplementary scheme S4: Study on the influence of the *N*-alkyl group (R) of peptide resins **2a-d** on the *N*-acylation yield.

**3a:**

**MALDI-TOF** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>39</sub>H<sub>58</sub>N<sub>10</sub>O<sub>11</sub>S: 875.4, found: 875.4.

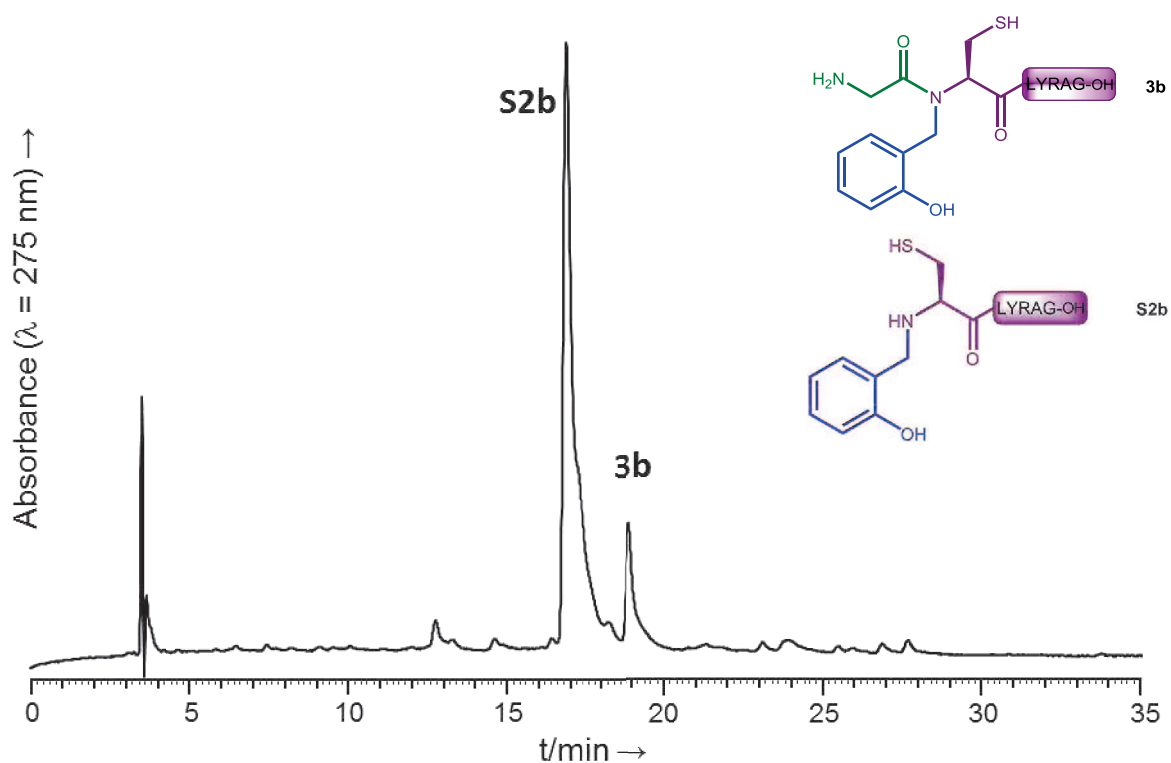
**HPLC:** retention time: 17.24 min (Nucleosil, isocratic 17% B/A over 30 min at 50°C).



**3b:**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{38}H_{56}N_{10}O_{10}S$ : 845.4, found 845.4.

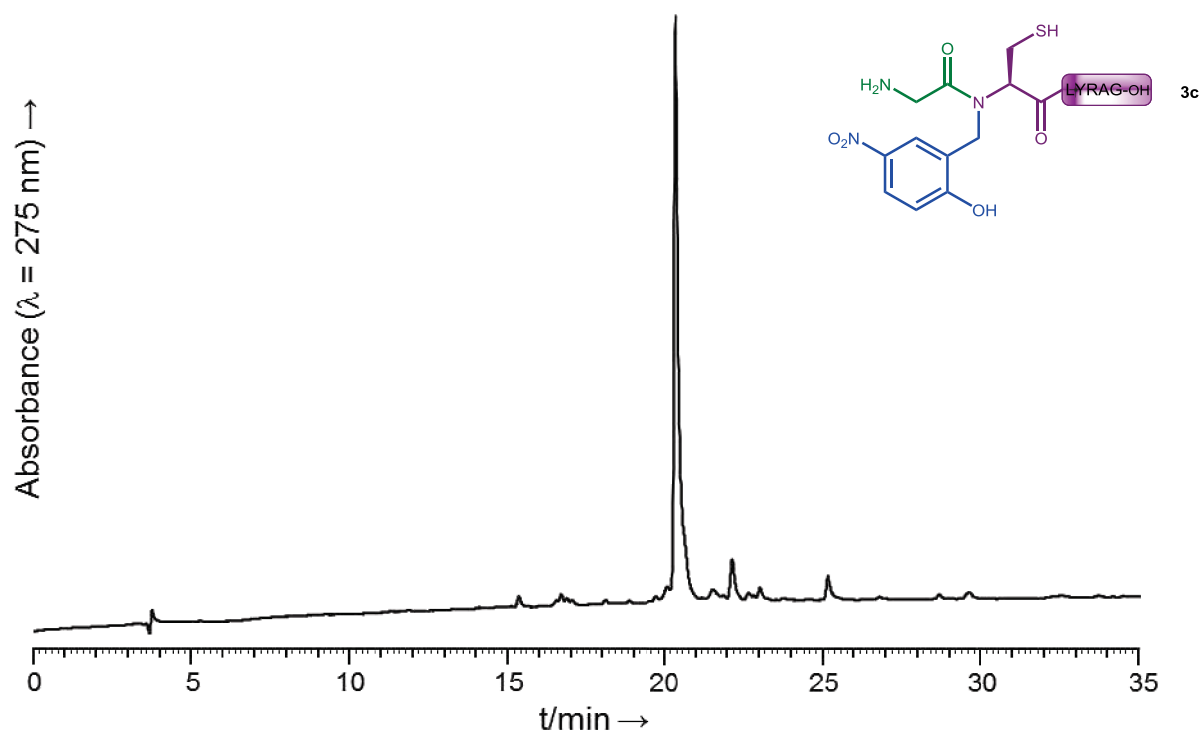
**HPLC**: retention time: 19 min (Nucleosil, gradient: 5-50 % B/A over 30 min).



**3c:**

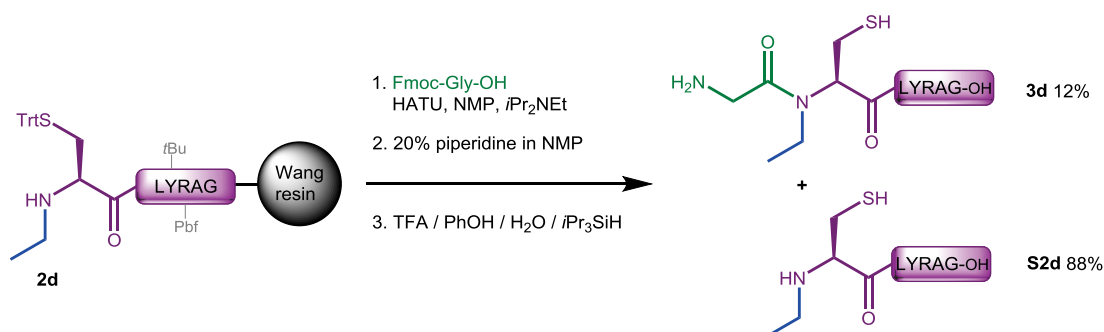
**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{38}H_{55}N_{11}O_{12}S$ : 890.4, found: 890.4.

**HPLC**: retention time: 20.53 min (Nucleosil, gradient: 5-50 % B/A over 30 min).



Supplementary figure S8: HPLC trace of crude **3c**.

*N*-acylation of peptide resin **2d** using HBTU/HOBt followed by deprotection of the Fmoc group then cleavage of the resin did not show any detectable amount of the expected *N*-acylated peptide **3d**. For characterization purpose, the *N*-acylation was performed using HATU as the coupling reagent, yielding to a small amount of **3d** together with the non-acylated peptide **S2d**.

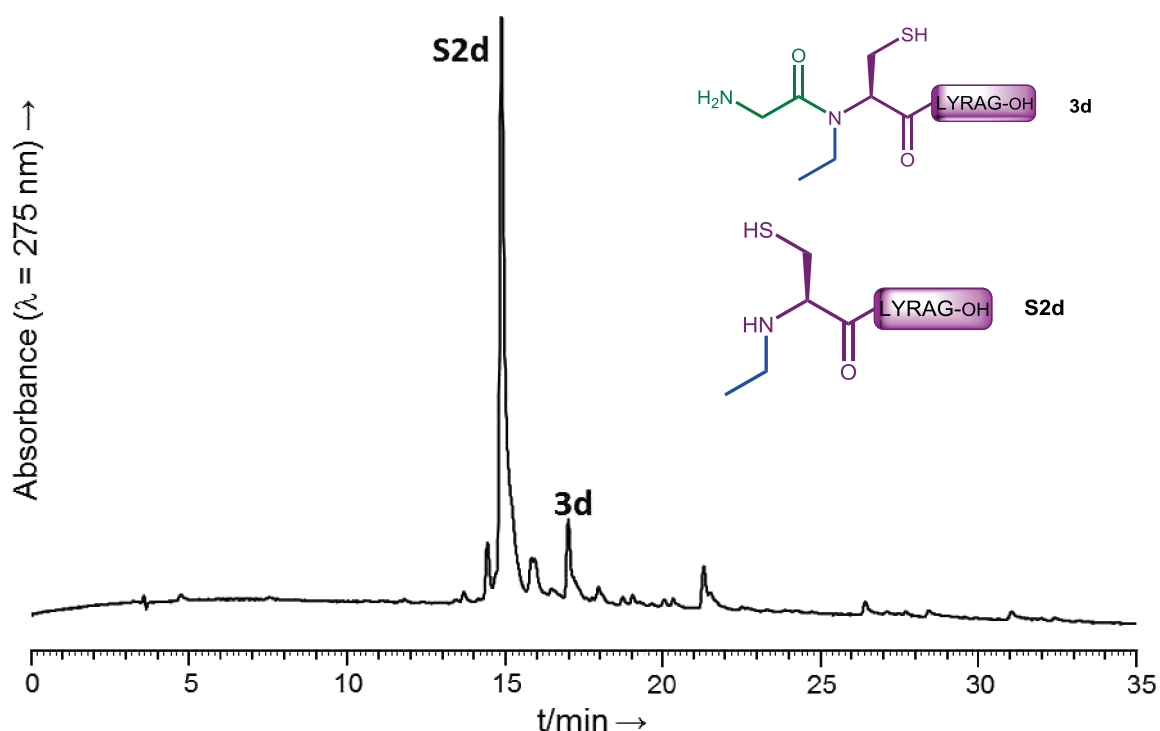


Supplementary scheme S5: *N*-acylation of peptide resin **2d** using HATU as the coupling reagent.

**3d:**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{33}H_{54}N_{10}O_9S$ : 767.4, found 767.4.

**HPLC**: retention time: 17.02 min (Nucleosil, gradient: 5-50 % B/A over 30 min).

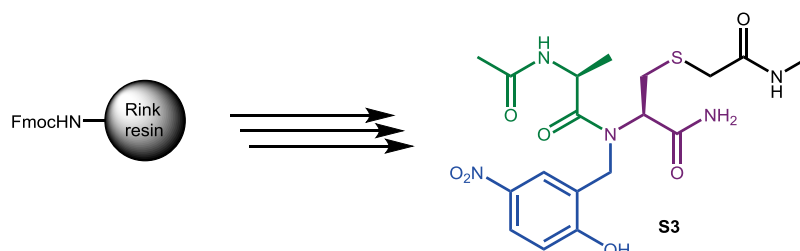


Supplementary figure S9: HPLC trace of crude 3d.

#### 4- Evaluation of the phenol pKa of an *N*-acyl-*N*-(Hnb)*Cys* compound

##### Synthesis of *S*-alkylated compound S3

In order to prevent any *N*-*S* shift that would make uncertain the determination of the pKa of the phenol group, we synthesized a model dipeptide **S3** *S*-alkylated with an acetamidomethyl (Acm) protective group.



Supplementary scheme S6: Synthesis of the model dipeptide S3.

Fmoc-Rink polystyrene resin (182 mg, 0.55 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. The resin was treated with 20% piperidine in NMP (10 mL, 3 × 3 min) and washed with NMP. Then, Fmoc-Cys(Acm)-OH (414 mg, 1 mmol, 10 equiv.), HBTU (360 mg, 0.95 mmol, 9.5 equiv.) and HOBT (128 mg, 0.95 mmol, 9.5 equiv.)

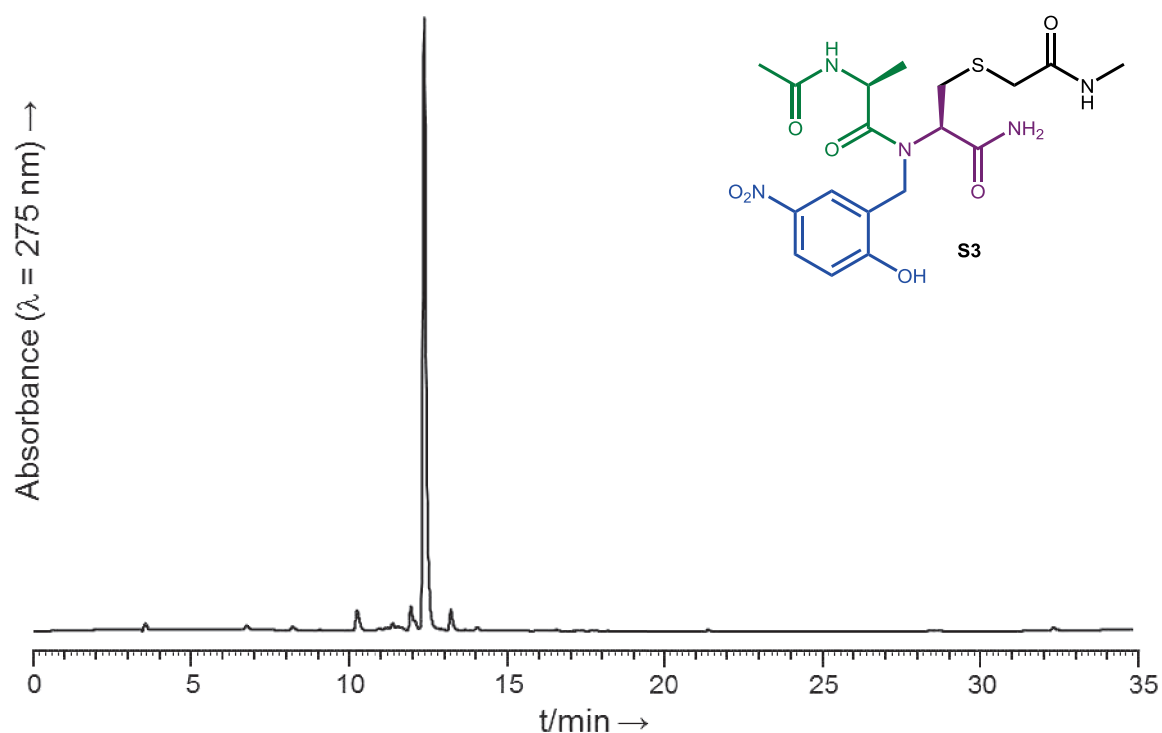


were dissolved in NMP (10 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was deprotected by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). Fmoc-Ala-OH (311 mg, 1 mmol, 10 equiv.) was coupled following the same coupling procedure as described above. Coupling was repeated once. The resin was then treated with 20% piperidine in NMP (10 mL, 3  $\times$  3 min), washed with NMP, then treated with acetic anhydride (571  $\mu$ L, 6 mmol, 60 equiv.), *i*Pr<sub>2</sub>NEt (272  $\mu$ L, 0.39 mmol, 15.5 equiv.) and HOBT (24 mg, 0.044 mmol, 1.8 equiv.) in NMP (12 mL) for 2  $\times$  1 h. Peptide-resin was then treated with 20% piperidine in NMP (10 mL, 3  $\times$  3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>, and finally cleaved using a solution of TFA / H<sub>2</sub>O (9:1) for 2 h. The solvents were removed under reduced pressure to give compound **S3**.

**S3:**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>18</sub>H<sub>26</sub>N<sub>5</sub>O<sub>7</sub>S: 456.1547, found: 456.1544.

**HPLC**: retention time: 12.35 min (Nucleosil, gradient: 5-95 % B/A over 30 min).

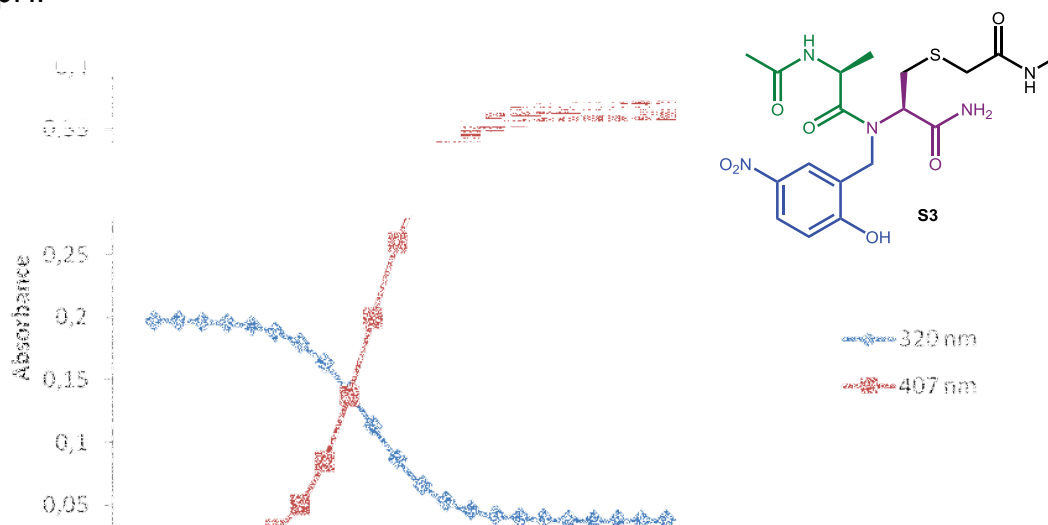


Supplementary figure S10: HPLC trace of crude **S3**.

Determination of the phenol pKa of S-alkylated compound **S3**

The pKa of the phenol group was determined by acid-base titration of a dilute (10  $\mu$ M) solution of compound **S3** in Milli-Q water. The pH was adjusted with 1M HCl and NaOH solutions. The titration was followed by UV spectrophotometry on an Uvikon 923 double beam UV/VIS spectrophotometer. The results were plotted at the  $\lambda_{\text{max}}$  of

the phenol and phenolate species (determined to be 320 and 407 nm, respectively) against pH.



Supplementary figure S11: UV absorbance of compound **S3** as a function of pH.

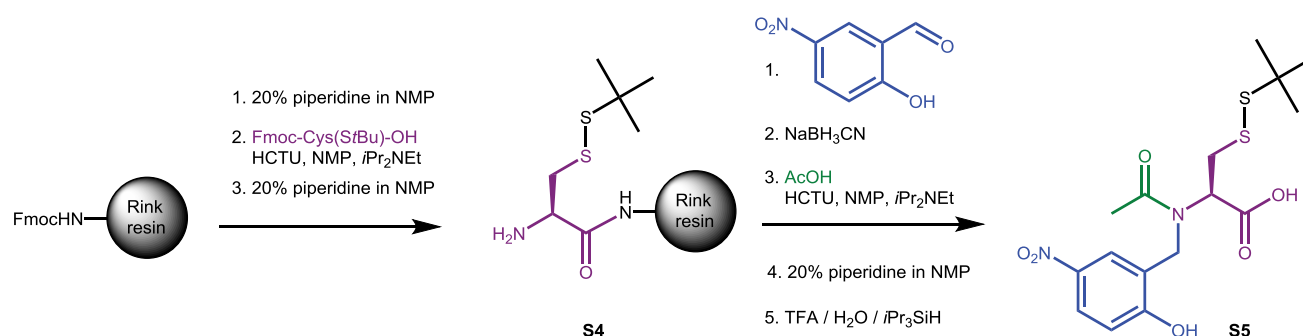
The pKa value ( $6.5 \pm 0.1$ ) was determined fitting with a sigmoidal function and measuring the inflection point of the curve.

$$407 \text{ nm: } y = 0.36324 + (0.00891 - 0.36324) / (1 + \exp((x - 6.51822) / 0.59329)) \quad R^2 = 0.996$$

$$320 \text{ nm: } y = 0.0378 + (0.19762 - 0.0378) / (1 + \exp((x - 6.53917) / 0.60827)) \quad R^2 = 0.993$$

## 5- Characterization of a model *N*-acyl-*N*-(Hnb)Cys compound

### 5a- <sup>1</sup>H and <sup>13</sup>C NMR of **S5** and comparison with <sup>1</sup>H NMR of *O*-methylated **S6**



Supplementary scheme S7: Synthesis of the compound **S5**.

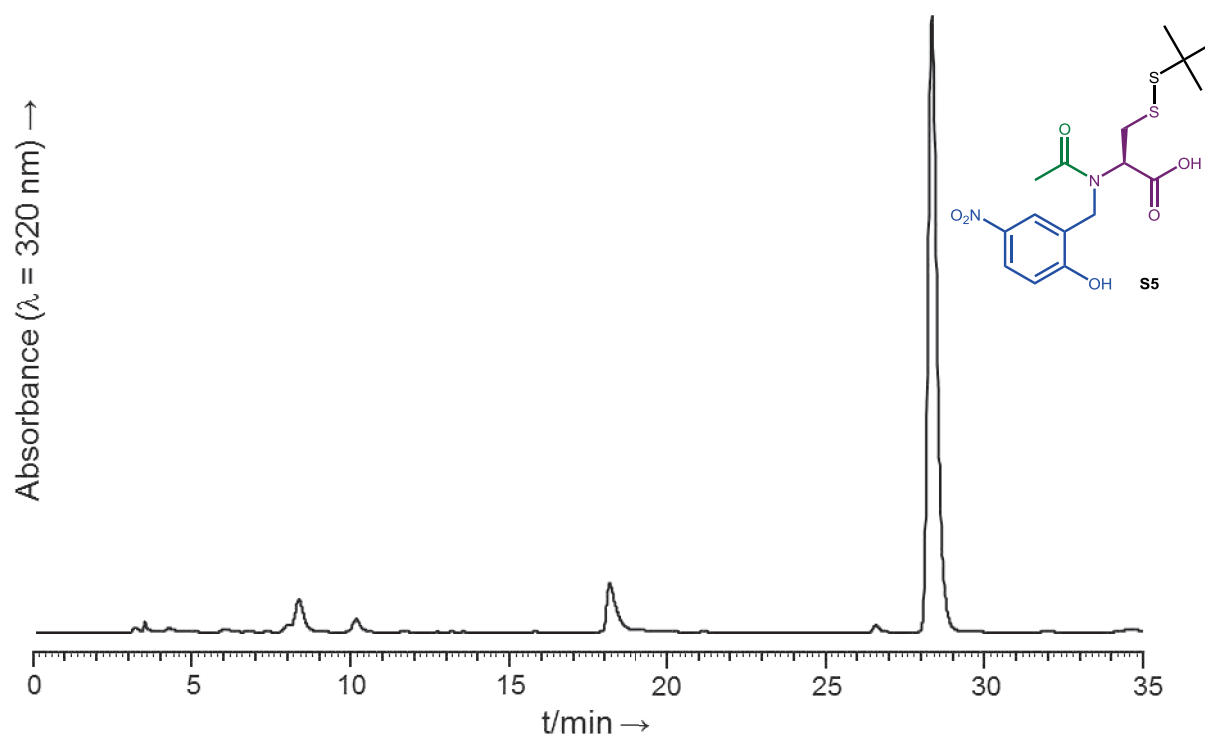
Fmoc-protected Rink polystyrene resin (63.3 mg, 0.79 mmol/g, 0.05 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was deprotected by three successive treatments with 20% piperidine in NMP (5 mL) for 3 min. Then, Fmoc-Cys(SfBu)-OH (216 mg, 0.5 mmol, 10 equiv.) and HCTU (186 mg, 0.45 mmol, 9 equiv.) were

dissolved in NMP (1 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175 μL, 1 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (5 mL) for 3 min to give resin **S4**. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5 for 0.05 mmol). Then, acetic acid (30 μL, 0.5 mmol, 10 equiv.) and HCTU (186 mg, 0.45 mmol, 9 equiv.) were dissolved in NMP (1 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175 μL, 1 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Amino acid-resin was then treated with 20% piperidine in NMP (5 mL, 3 × 3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>. Finally, the compound **S5** was obtained after treatment with TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH, 93/5/2 (4 mL) for 2 h and characterized. Note that the hydrolysis of the C-terminal amide into the acid carboxylic was quantitative under the cleavage conditions. The product was purified by flash reverse phase chromatography (45 mg crude in 1 mL MeCN, column: RP18 25-40 μm - 20 g, Götec-Labortechnik GmbH, gradient: 30-50% B/A over 20 min, 35 mL/min).

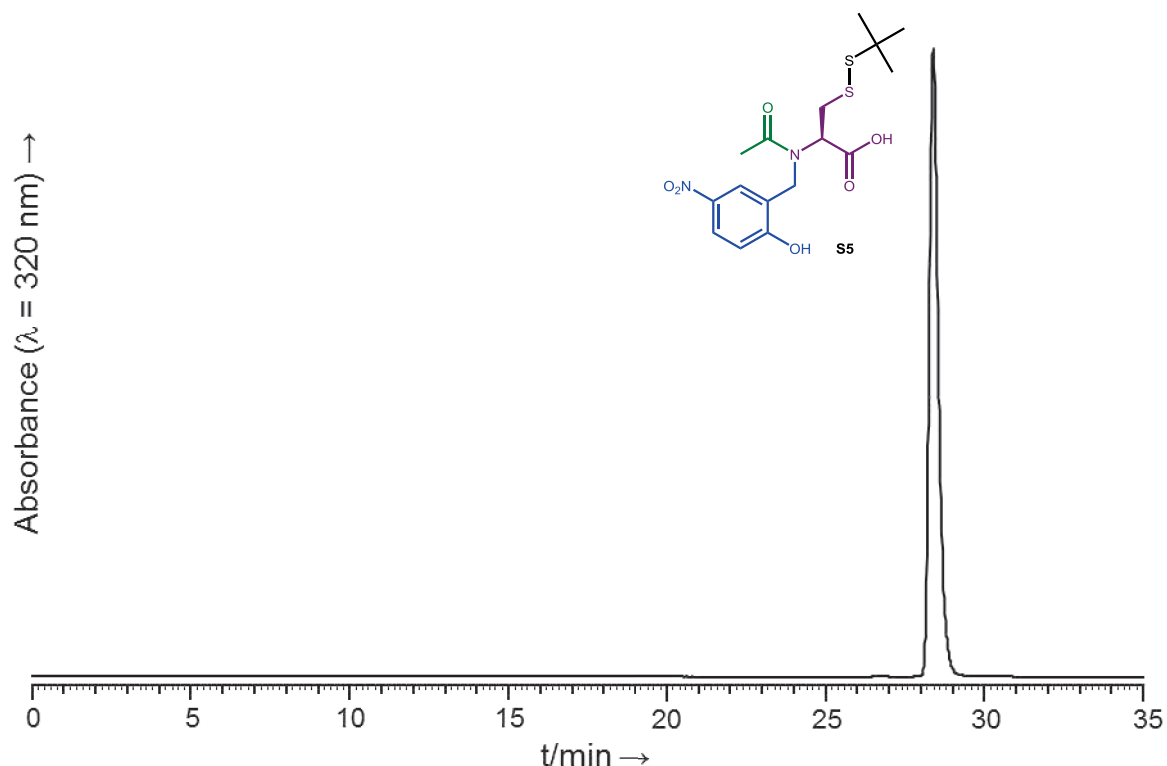
**S5:**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>16</sub>H<sub>23</sub>N<sub>2</sub>O<sub>6</sub>S<sub>2</sub>: 403.0998, found: 403.0993.

**HPLC**: retention time: 28.36 min (Nucleosil, gradient: 20-50% B/A over 30 min).

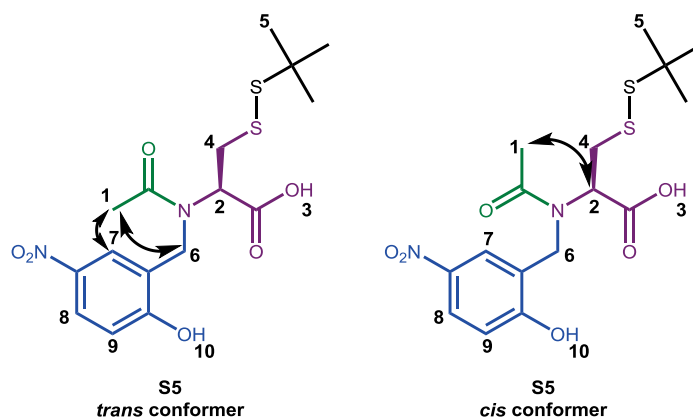


Supplementary figure S12: HPLC trace of crude **S5**.



Supplementary figure S13: HPLC trace of pure **S5**.

Pure **S5** has been analyzed by <sup>1</sup>H NMR in DMSO-*d*<sub>6</sub>. Two conformers are observed: *trans* and *cis* isomers (evaluated ratio 73:27). On the supplementary figure S14, arrows indicate observed NOE on a ROESY 2D spectrum (mixing time = 200 ms) used to characterize the two conformers.



Supplementary figure S14: *Cis* and *trans* isomers of **S5** and observed ROESY connectivities.

### **Trans** isomer of **S5**

**<sup>1</sup>H NMR (600 MHz, DMSO-*d*<sub>6</sub>):** 13.40 (1H, H<sub>3</sub>, s), 11.48 (1H, H<sub>10</sub>, s), 8.42 (1H, H<sub>7</sub>, d, *J* = 2.8 Hz), 8.10 (1H, H<sub>8</sub>, dd, *J* = 8.9, 2.9 Hz), 7.01 (1H, H<sub>9</sub>, d, *J* = 8.9 Hz), 4.64-5.52 (2H, H<sub>6a</sub> and H<sub>6b</sub>, m), 4.26 (1H, H<sub>2</sub>, t, *J* = 6.7 Hz), 3.40-3.36 (1H, H<sub>4b</sub>, m), 2.87 (1H, H<sub>4a</sub>, dd, *J* = 13.5, 7.5 Hz), 2.06 (3H, H<sub>1</sub>, s), 1.21 (9H, H<sub>5</sub>, s).

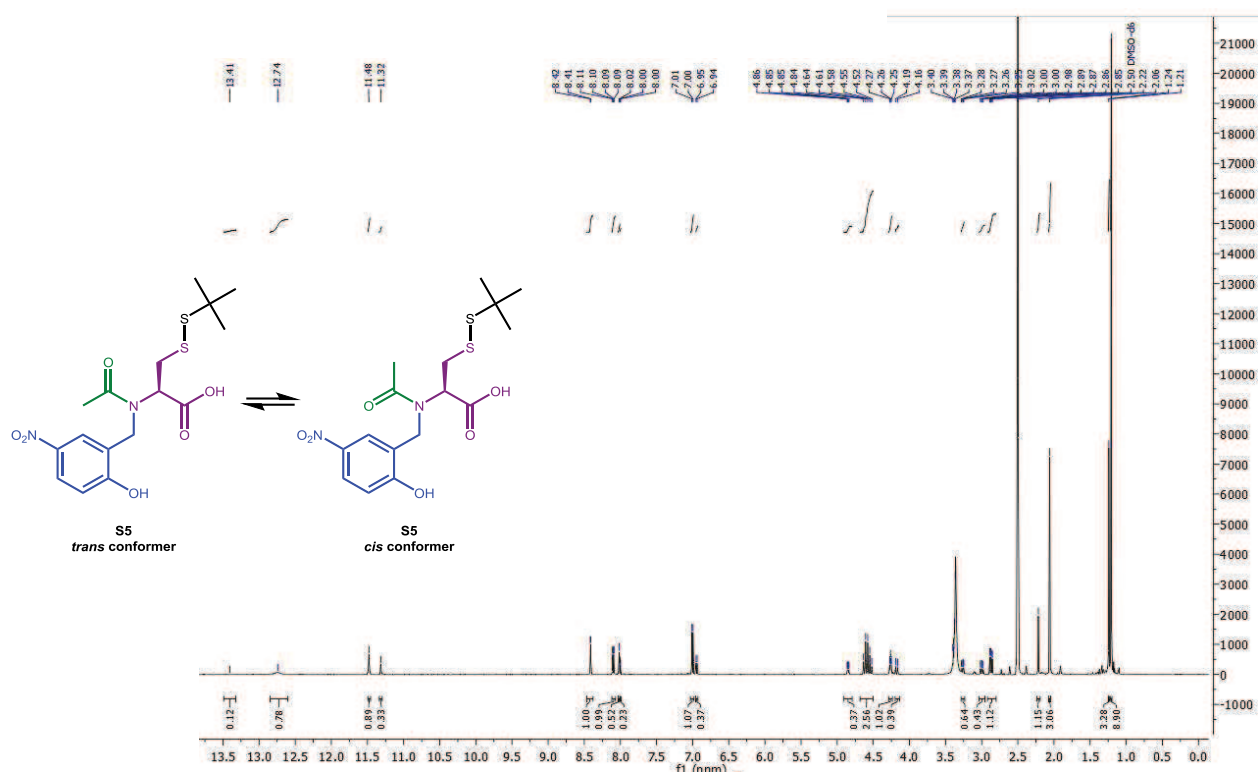
**Cis isomer of S5**

**<sup>1</sup>H NMR (600 MHz, DMSO-*d*<sub>6</sub>):** 12.74 (1H, H<sub>3</sub>, br), 11.32 (1H, H<sub>10</sub>, s), 8.04-8.00 (1H, H<sub>8</sub>, m), 8.02 (1H, H<sub>7</sub>, br), 6.95 (1H, H<sub>9</sub>, d, *J* = 8.7 Hz), 4.85 (1H, H<sub>2</sub>, dd, *J* = 9.6, 4.8 Hz), 4.55-4.52 (1H, H<sub>6b</sub>, m), 4.18 (1H, H<sub>6a</sub>, d, *J* = 16.6 Hz), 3.26 (1H, H<sub>4b</sub>, dd, *J* = 14, 4.8 Hz), 3.01 (1H, H<sub>4a</sub>, dd, *J* = 13.9, 9.6 Hz), 2.22 (3H, H<sub>1</sub>, s), 1.24 (9H, H<sub>5</sub>, s).

**Trans and cis isomers of S5 - 27 <sup>13</sup>C peaks observed** (over 28, the remaining signal probably being hidden by another peak at the same chemical shift and not discernable at the 600 MHz resolution)

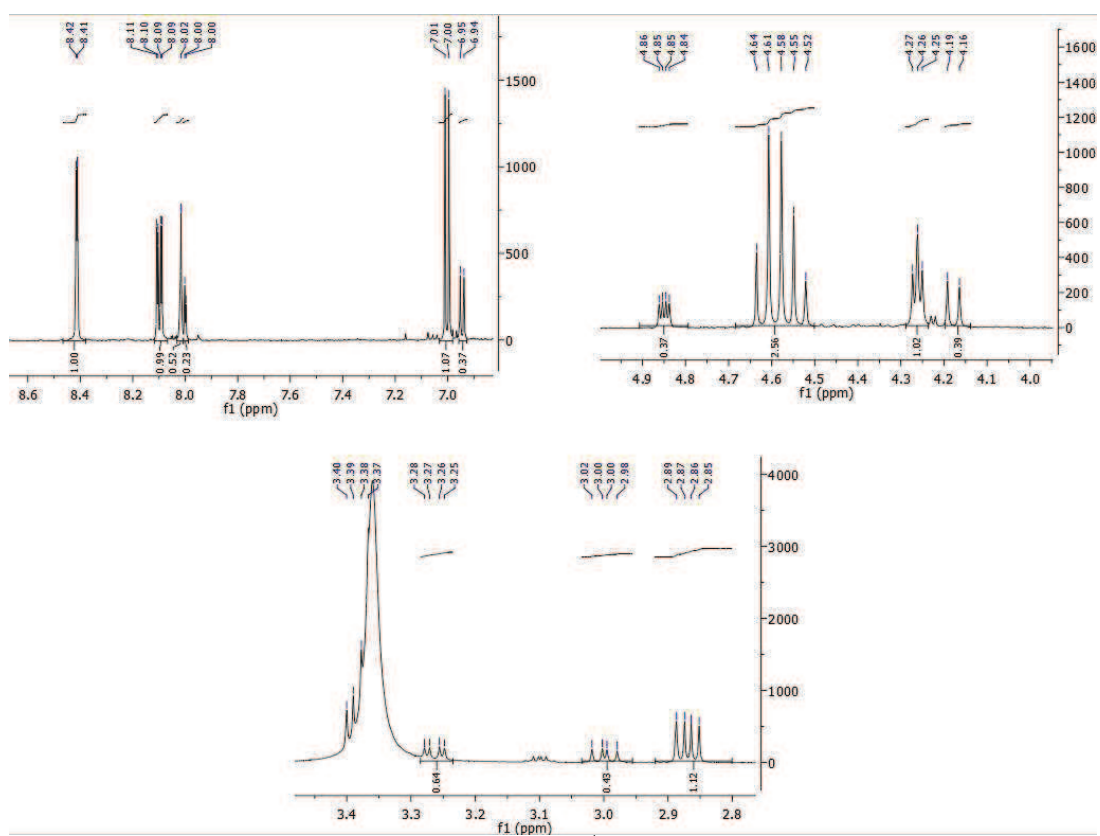
**<sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>):** 171.72, 170.85, 170.80, 161.75, 160.80, 139.63, 139.39, 125.27, 125.16, 125.12, 124.65, 124.54, 124.21, 115.15, 114.98, 99.51, 60.38, 59.42, 48.37, 48.18, 47.60, 45.76, 40.16, 40.06, 29.47, 29.42, 22.04, 21.59.

Partie expérimentale - article 1



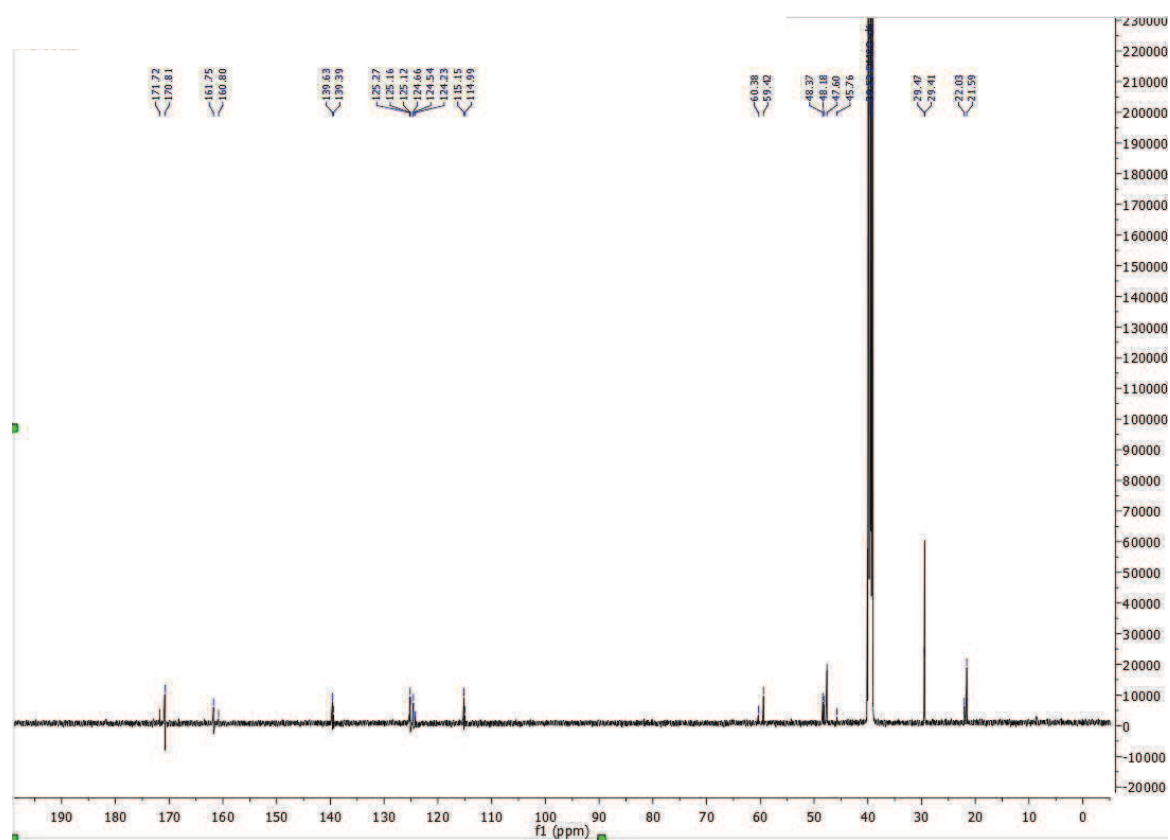
Supplementary figure S15: Copy of the <sup>1</sup>H NMR spectrum of S5.

Partie expérimentale - article 1



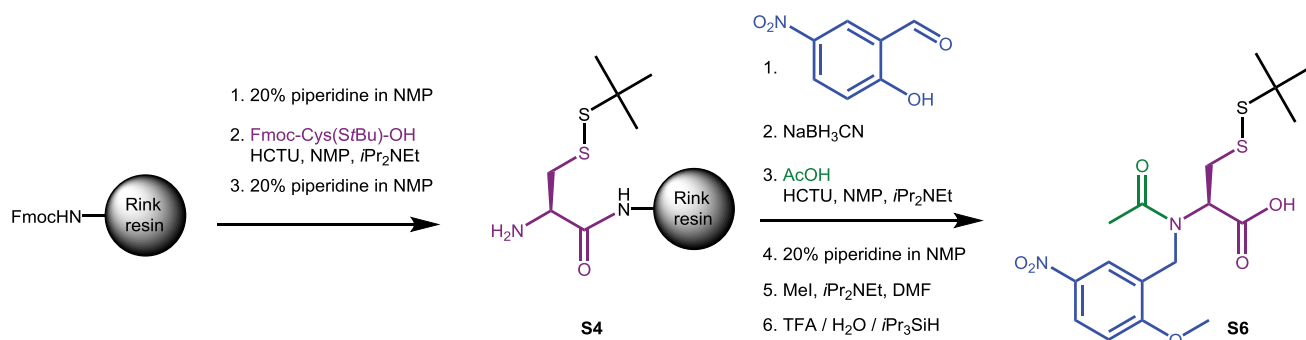
Supplementary figure S16: Copy of the  $^1\text{H}$  NMR spectrum (zoom) of S5.

Deuterium (400 MHz) - 1000000



Supplementary figure S17: Copy of the  $^{13}\text{C}$  NMR spectrum of **S5**.





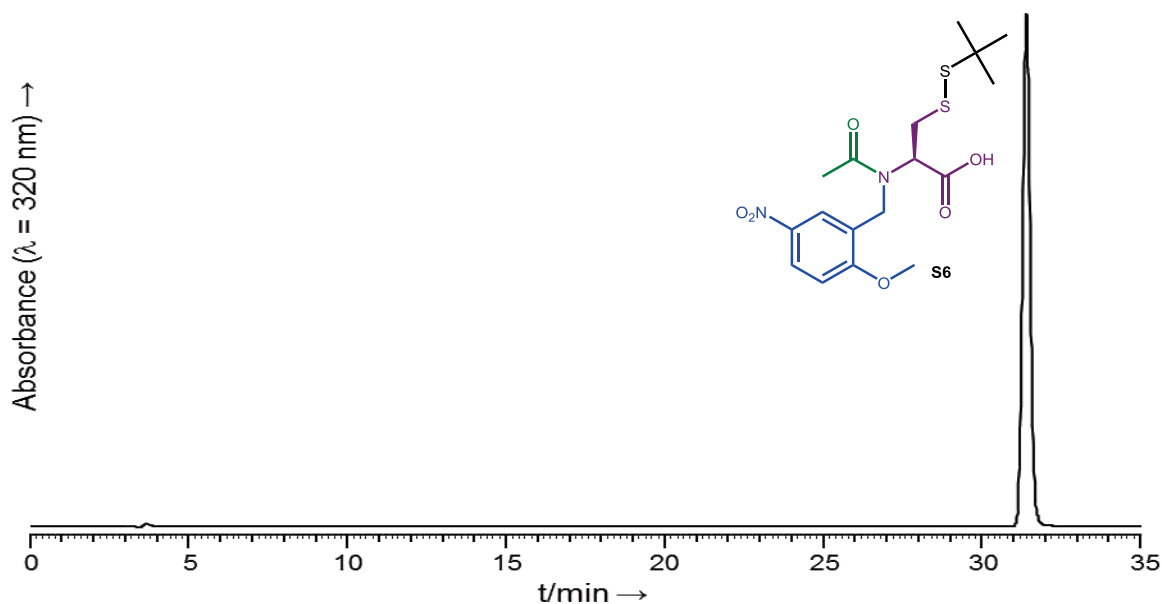
**Supplementary scheme S8: Synthesis of the compound S6.**

Fmoc-protected Rink polystyrene resin (63.3 mg, 0.79 mmol/g, 0.05 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was deprotected by three successive treatments with 20% piperidine in NMP (5 mL) for 3 min. Then, Fmoc-Cys(S $t$ Bu)-OH (216 mg, 0.5 mmol, 10 equiv.) and HCTU (186 mg, 0.45 mmol, 9 equiv.) were dissolved in NMP (1 mL) prior to addition of  $i$ Pr $_2$ NEt (175  $\mu$ L, 1 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (5 mL) for 3 min to give resin **S4**. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5 for 0.05 mmol). Then, acetic acid (30  $\mu$ L, 0.5 mmol, 10 equiv.) and HCTU (186 mg, 0.45 mmol, 9 equiv.) were dissolved in NMP (1 mL) prior to addition of  $i$ Pr $_2$ NEt (175  $\mu$ L, 1 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Amino acid-resin was then treated with 20% piperidine in NMP (5 mL, 3  $\times$  3 min), washed with NMP then CH $_2$ Cl $_2$ . Methylation of the phenol group was performed using a solution of MeI (50 equiv.),  $i$ Pr $_2$ NEt (50 equiv.) in DMF for 4 h and the resin was washed with NMP and CH $_2$ Cl $_2$ . Finally, the compound **S6** was obtained after treatment with TFA/H $_2$ O/ $i$ Pr $_3$ SiH, 93/5/2 (4 mL) for 2 h and characterized. Note that the hydrolysis of the C-terminal amide into the acid carboxylic was quantitative under the cleavage conditions. The product was purified by semi-preparative HPLC (Nucleosil, 40-55 % B/A over 15 min).

**S6:**

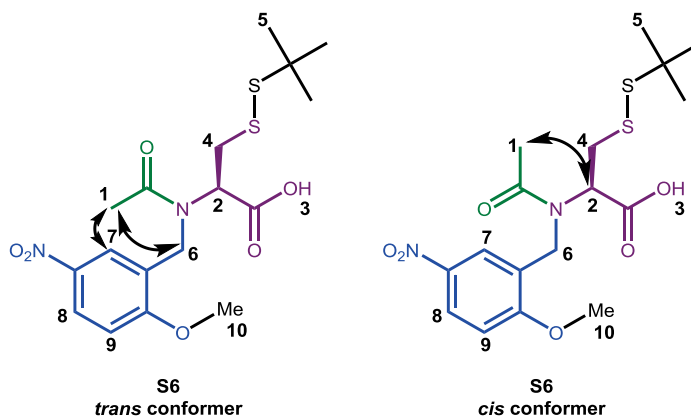
**ESI** ( $m/z$ ): [MH] $^+$  calcd for C $_{17}$ H $_{25}$ N $_2$ O $_6$ S $_2$ : 417.1, found: 417.1.

**HPLC**: retention time: 31.28 min (Nucleosil, gradient: 20-50% B/A over 30 min).



Supplementary figure S18: HPLC trace of pure **S6**.

Pure **S6** has been analyzed by  $^1\text{H}$  NMR in  $\text{DMSO-}d_6$  for comparison of the *cis-trans* ratio with **S5**. Two conformers are observed: *trans* and *cis* isomers (evaluated ratio 72:28, nearly identical with **S5** ratio). On the supplementary figure S19, arrows indicate observed NOE on a ROESY 2D spectrum (mixing time = 200 ms) used to characterize the two conformers.



Supplementary figure S19: *Cis* and *trans* isomers of **S6** and observed ROESY connectivities.

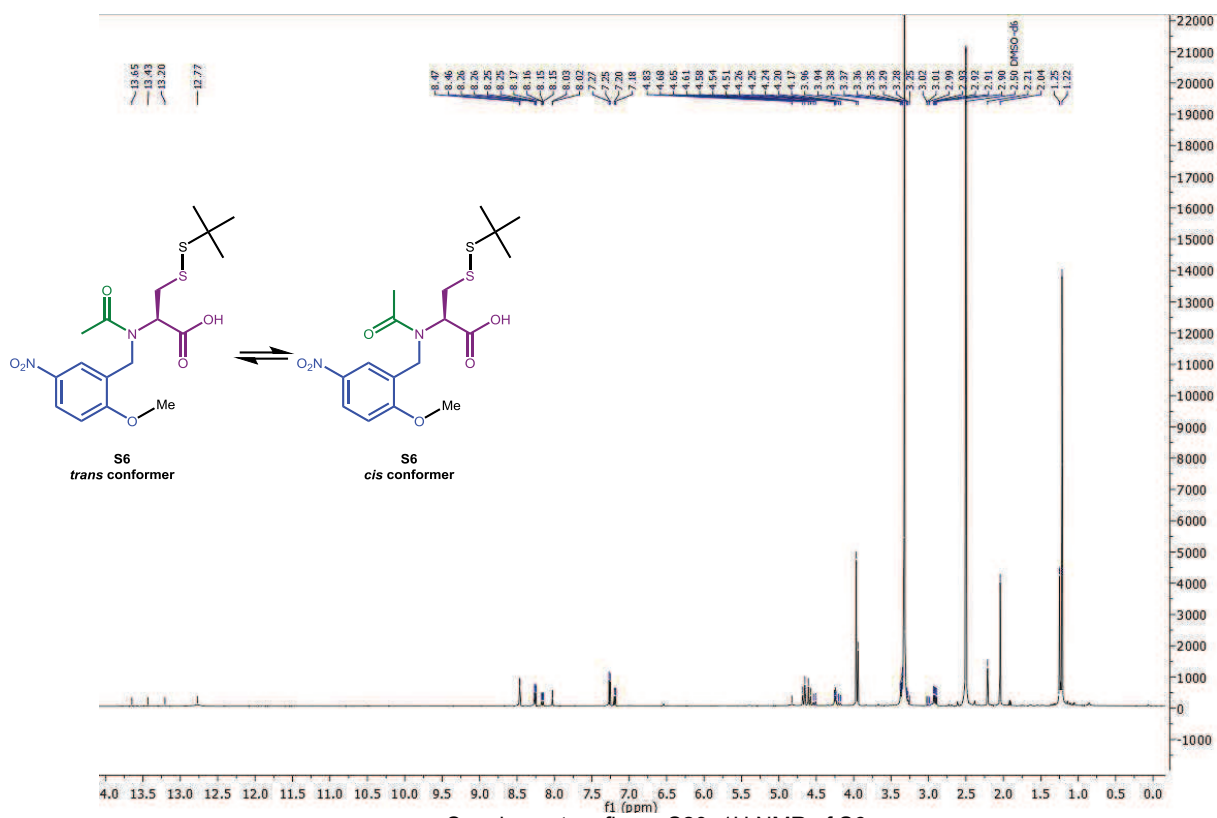
#### **Trans isomer (major) of S6**

**$^1\text{H}$  NMR (600 MHz,  $\text{DMSO-}d_6$ ):** 12.77 (1H,  $\text{H}_3$ , bs), 8.46 (1H,  $\text{H}_7$ , d,  $J = 2.7$  Hz), 8.26 (1H,  $\text{H}_8$ , dd,  $J = 9.1, 2.8$  Hz), 7.26 (1H,  $\text{H}_9$ , d,  $J = 9.1$  Hz), 4.6 (2H,  $\text{H}_{6a}$  and  $\text{H}_{6b}$ , dd,  $J = 45.3, 17.3$  Hz), 4.25 (1H,  $\text{H}_2$ , t,  $J = 6.5$  Hz), 3.96 (3H,  $\text{H}_{10}$ , s), 3.36 (1H,  $\text{H}_{4b}$ , dd,  $J = 13.5, 5.8$  Hz), 2.91 (1H,  $\text{H}_{4a}$ , dd,  $J = 13.7, 7.8$  Hz), 2.04 (3H,  $\text{H}_1$ , s), 1.22 (9H,  $\text{H}_5$ , s).

#### **Cis isomer (minor) of S6**

**$^1\text{H}$  NMR (600 MHz,  $\text{DMSO-}d_6$ ):** 13.43 (1H,  $\text{H}_3$ , bs), 8.16 (1H,  $\text{H}_8$ , dd,  $J = 9.0, 2.9$  Hz), 8.03 (1H,  $\text{H}_7$ , d,  $J = 2.7$  Hz), 7.19 (1H,  $\text{H}_9$ , d,  $J = 9.1$  Hz), 4.86-4.79 (1H,  $\text{H}_2$ , m), 4.52 (1H,  $\text{H}_{6b}$ , t,  $J = 16.9$  Hz), 4.19 (1H,  $\text{H}_{6a}$ , d,  $J = 17.4$  Hz), 3.94 (3H,  $\text{H}_{10}$ , s), 3.29-3.25 (1H,  $\text{H}_{4b}$ , m), 3.02-2.99 (1H,  $\text{H}_{4a}$ , m), 2.21 (3H,  $\text{H}_1$ , s), 1.25 (9H,  $\text{H}_5$ , s).

Partie expérimentale - article 1



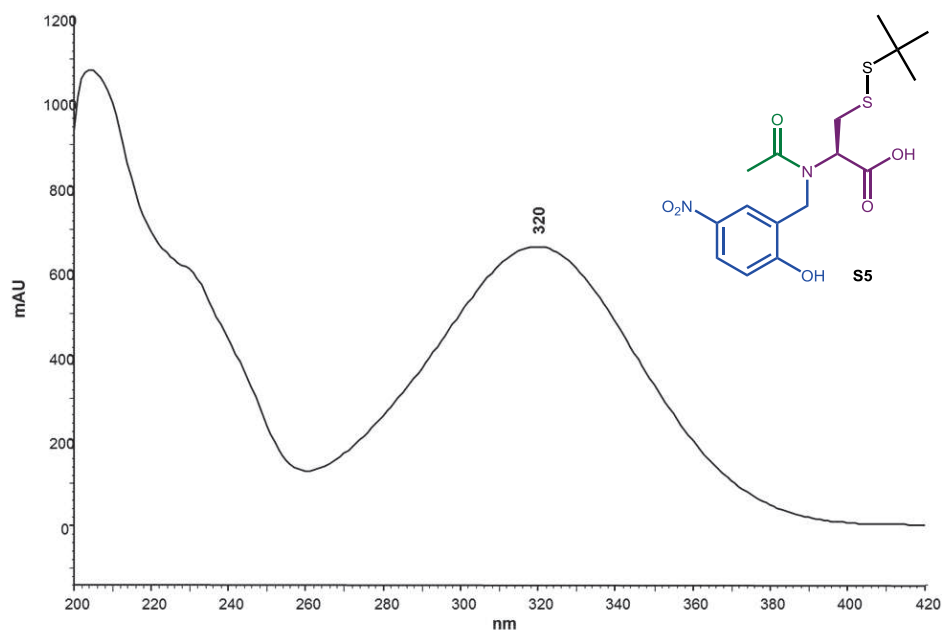
Supplementary figure S20: <sup>1</sup>H NMR of S6.

5b- Molar extinction coefficient of the Hnb group of S5

The molar extinction coefficient ( $\epsilon$ ) of the Hnb group of **S5** was determined by UV spectroscopy at 275, 280, 320 nm in a H<sub>2</sub>O/MeCN/TFA (8:2:0.001) solution.

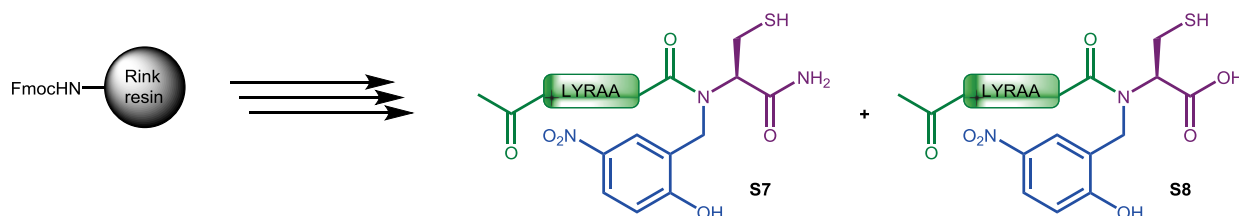
Wavelength (nm)	275	280	320
$\epsilon$ (L.mol <sup>-1</sup> .cm <sup>-1</sup> )	1846	2359	6325

Supplementary table S1: Molar extinction coefficient of **S5**.



Supplementary figure S21: UV spectrum of **S5**.

## 6- Optimization of the cysteine thiol protecting group

6a- Synthesis and stability of (S-Trt) model peptide S7

Supplementary scheme S9: Synthesis of the model peptide **S7** showing hydrolysis co-product **S8**.

Fmoc-protected Rink polystyrene resin (126.6 mg, 0.79 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was deprotected by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Cys(Trt)-OH

(586 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348 μL, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). The Ac-LYRAA sequence was installed by standard SPPS procedure (protocol p S3), an extended double coupling (2 × 2h) was performed for the first alanine and a normal double coupling (2 × 30 min) was performed for the arginine. Peptide resin was then treated with 20% piperidine in NMP (10 mL, 3 × 3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>, and finally cleaved following the general procedure p S3 to give compound **S7** and side product **S8** (89:11) (Measured by HPLC at 275nm). Peptides were obtained as a mixture of C-terminal amide and acid arising from hydrolysis of the amide under the TFA cleavage conditions.<sup>1</sup> Their corresponding thioester forms arising from a premature *N*-S shift was also observed (compounds **S7'** and **S8'**).

**S7:**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>39</sub>H<sub>58</sub>N<sub>11</sub>O<sub>11</sub>S: 888.4038, found: 888.4028.

**HPLC**: retention time: 4.36 min (Chromolith, gradient: 5-50 % B/A over 5 min).

**S8:**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>39</sub>H<sub>57</sub>N<sub>10</sub>O<sub>12</sub>S: 889.3878, found: 889.3881.

**HPLC**: retention time: 4.44 min (Chromolith, gradient: 5-50 % B/A over 5 min).

**S7':**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>39</sub>H<sub>58</sub>N<sub>11</sub>O<sub>11</sub>S: 888.4035, found: 888.4028.

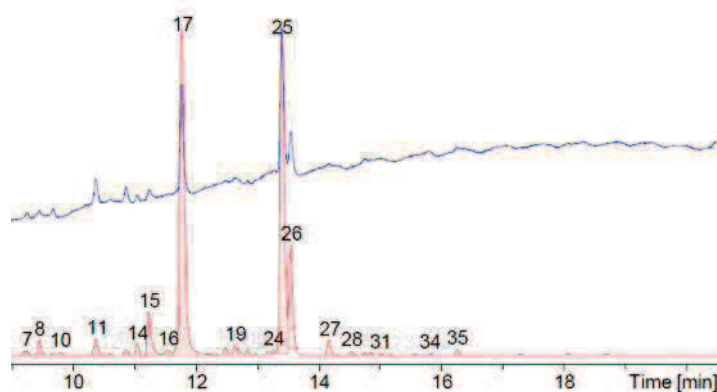
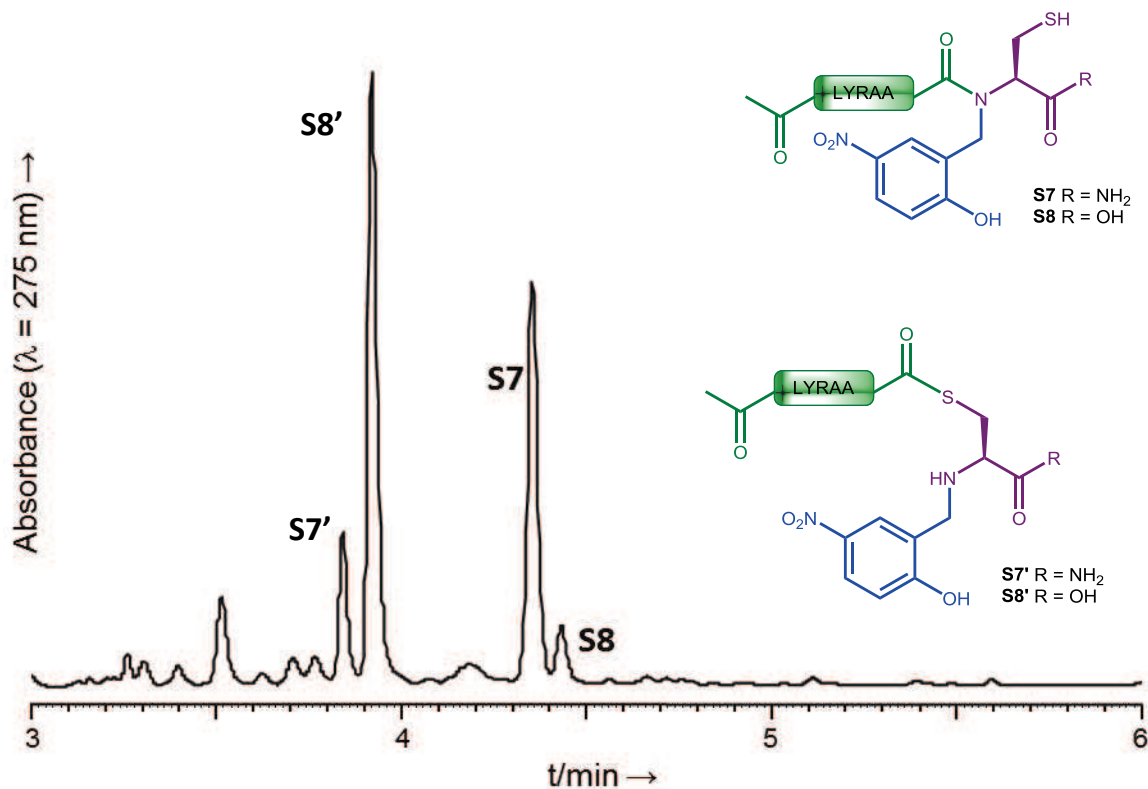
**HPLC**: retention time: 3.82 min (Chromolith, gradient: 5-50 % B/A over 5 min).

**S8':**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>39</sub>H<sub>57</sub>N<sub>10</sub>O<sub>12</sub>S: 889.3884, found: 889.3881.

**HPLC**: retention time: 3.92 min (Chromolith, gradient: 5-50 % B/A over 5 min).

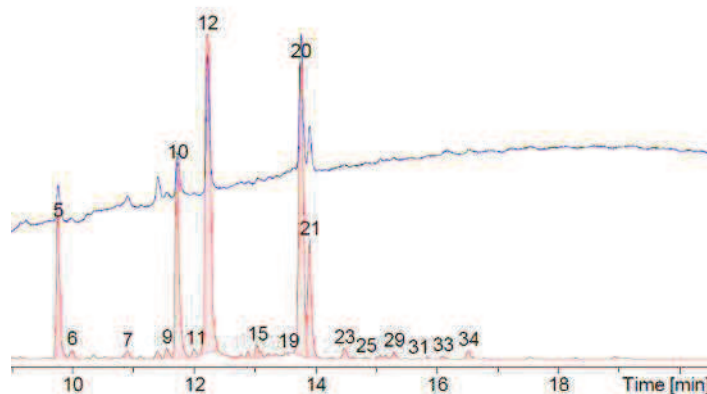
<sup>1</sup> a) C. J. Creighton, T. T. Romoff, J. H. Bu, M. Goodman, *J. Am. Chem. Soc.* **1999**, *121*, 6786–6791; b) M. Teixidó, F. Albericio, E. Giralt, *J. Peptide Res.* **2005**, *65*, 153–166.



Peak (tr (min))	$[MH]^+$ ( $m/z$ ) Calc.	$[MH]^+$ ( $m/z$ ) found	Attributed to
15 (11.24)	888.4038	888.4035	Premature $N \rightarrow S$ shift of <b>S7</b> (thioester form <b>S7'</b> )
17 (11.77)	889.3878	889.3884	Premature $N \rightarrow S$ shift of <b>S8</b> (thioester form <b>S8'</b> )
25 (13.40)	888.4038	888.4028	<b>S7</b>
26 (13.55)	889.3878	889.3881	<b>S8</b>

Supplementary table S2: Attribution of selected peaks observed in LC/MS analysis of crude **S7**.

To test the stability of this mixture of peptides, 1 mg of crude product was solubilized in 1 mL of a 8:2:0.01 H<sub>2</sub>O/MeCN/TFA mixture. The stability of **S7** was studied by analyzing this solution by LC-MS after overnight incubation. Formation of Ac-LYRAA-OH was observed, probably arising from hydrolysis of **S7'** and **S8'**.

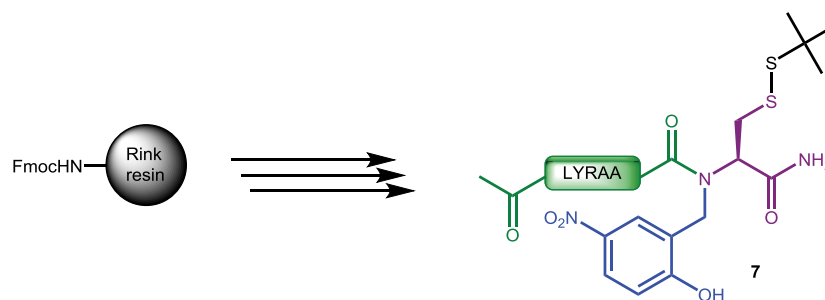


**Supplementary figure S24:**

LC/MS trace of crude **S7** dissolved in 8:2:0.01 H<sub>2</sub>O/MeCN/TFA (1mg/mL), after t = 17.5 h incubation at room temperature. Blue trace UV 214 nm, red trace TIC.

Peak (tr (min))	[MH] <sup>+</sup> (m/z) Calc.	[MH] <sup>+</sup> (m/z) found	Attributed to
5 (9.76)	635.3517	635.3516	Ac-LYRAA-OH
10 (11.72)	888.4038	888.4040	Premature N→S shift of <b>S7</b> (thioester form <b>S7'</b> )
12 (12.21)	889.3878	889.3888	Premature N→S shift of <b>S8</b> (thioester form <b>S8'</b> )
20 (13.75)	888.4038	888.4028	<b>S7</b>
22 (13.90)	889.3878	889.3881	<b>S8</b>

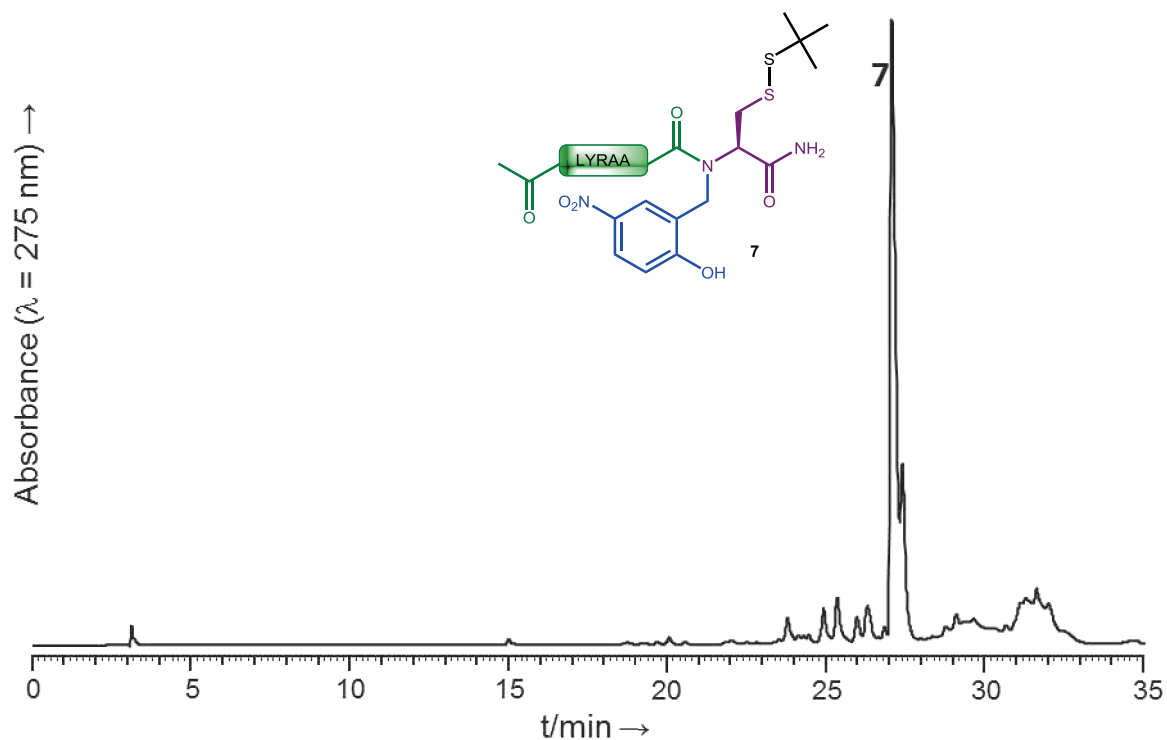
**Supplementary table S3:** Attribution of selected peaks observed in LC/MS analysis of crude **S7** after 17.5 h in solvents A/B.

6b- Synthesis and stability of (S-StBu) model peptide 7Supplementary scheme S10: Synthesis of model peptide 7.

Fmoc-protected Rink polystyrene resin (126.6 mg, 0.79 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Cys(S $t$ Bu)-OH (432 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of  $i$ Pr $_2$ NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). The Ac-LYRAA sequence was installed by standard SPPS procedure (protocol p S3), an extended double coupling (2  $\times$  2h) was performed for the first alanine and a normal double coupling (2  $\times$  30 min) was performed for the arginine. Peptide resin was then treated with 20% piperidine in NMP (10 mL, 3  $\times$  3 min), washed with NMP then CH $_2$ Cl $_2$ , and finally cleaved following the general procedure p S3 to give compound **7**. Peptide was obtained as a mixture of C-terminal amide and acid arising from hydrolysis of the amide under the TFA cleavage conditions.<sup>2</sup>

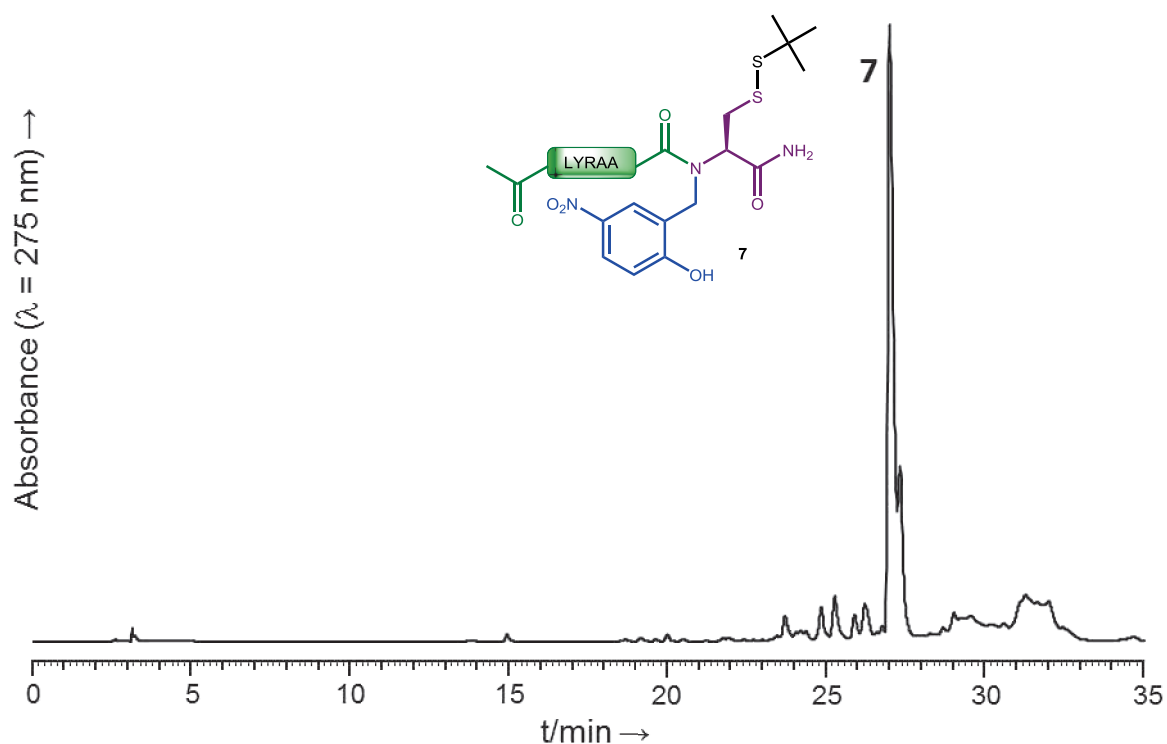
**7:****MALDI-TOF** ( $m/z$ ): [MH]<sup>+</sup> calcd for C $_{43}$ H $_{66}$ N $_{11}$ O $_{11}$ S $_2$ : 976.44, found: 976.45.**HPLC**: retention time: 30.89 min (Nucleosil, gradient: 5-50% B/A over 30 min).





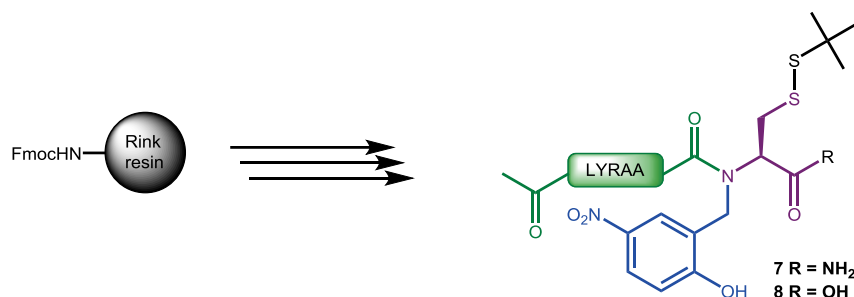
Supplementary figure S25: HPLC trace of crude **7**.

To test the stability of this mixture of peptides, 1 mg of crude product was solubilized in 1 mL of 8:2:0.01  $\text{H}_2\text{O}/\text{MeCN}/\text{TFA}$ . The stability of **7** was studied by analyzing this solution by HPLC at different times. Those products are perfectly stable in acidic aqueous conditions.



Supplementary figure S26: HPLC trace of crude **7** dissolved in 8:2:0.01  $\text{H}_2\text{O}/\text{MeCN}/\text{TFA}$  (1mg/mL), after  $t = 17.5 \text{ h}$  incubation at room temperature.

## 7- Optimization of the spacer between the cysteine and the resin

7a- Direct attachment to Rink linker – amide hydrolysis co-product<sup>2</sup>Supplementary scheme S11: Synthesis of the model peptide 7.

During the cleavage step (using procedure p S3) in order to obtain peptide **7**, the formation of peptide **8** was also observed (12 % determined by HPLC integration at 275 nm) from the hydrolysis of the C-terminal amide function into a C-terminal carboxylic acid.

**7:**

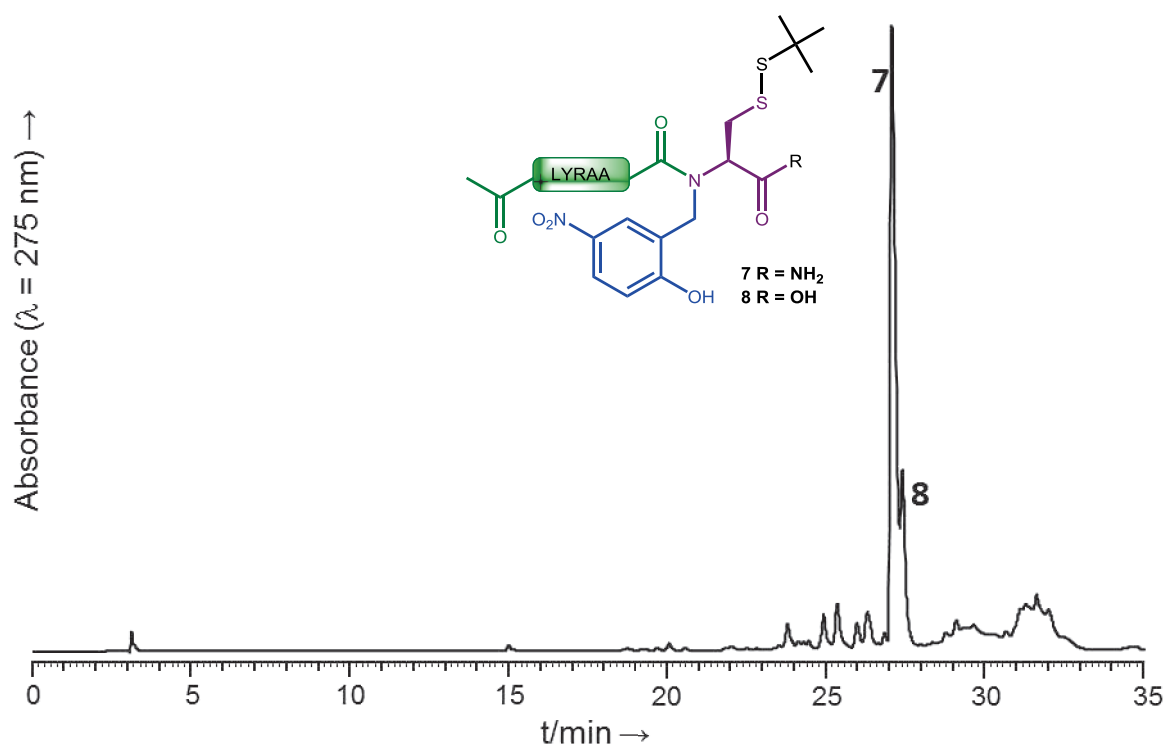
**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for C<sub>43</sub>H<sub>66</sub>N<sub>11</sub>O<sub>11</sub>S<sub>2</sub>: 976.44, found: 976.45.

**HPLC**: retention time: 30.89 min (Nucleosil, gradient: 5-50% B/A over 30 min).

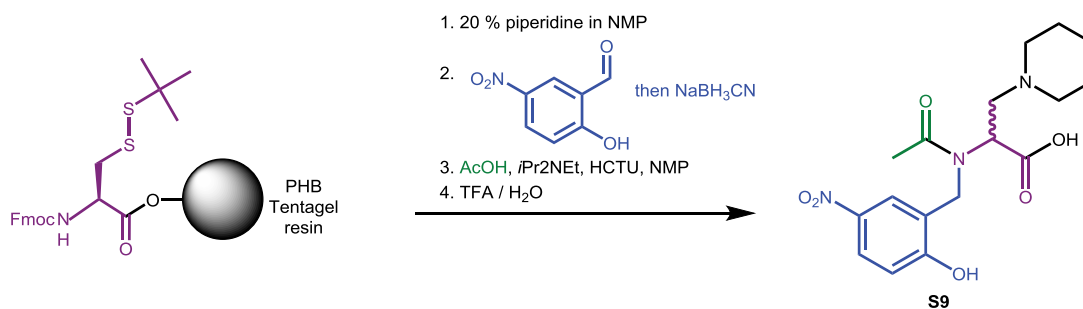
**8 – C-terminal hydrolysis co-product:**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for C<sub>43</sub>H<sub>65</sub>N<sub>10</sub>O<sub>12</sub>S<sub>2</sub>: 977.42, found: 977.43.

**HPLC**: retention time: 31.27 min (Nucleosil, gradient: 5-50% B/A over 30 min).



Supplementary figure S27: HPLC trace of crude **7** showing hydrolysis side product **8**.

7b- Direct attachment to PHB Tentagel resin (Wang type linker)Supplementary scheme S12: Formation of the C-terminal piperidine adduct **S9**.<sup>2</sup>

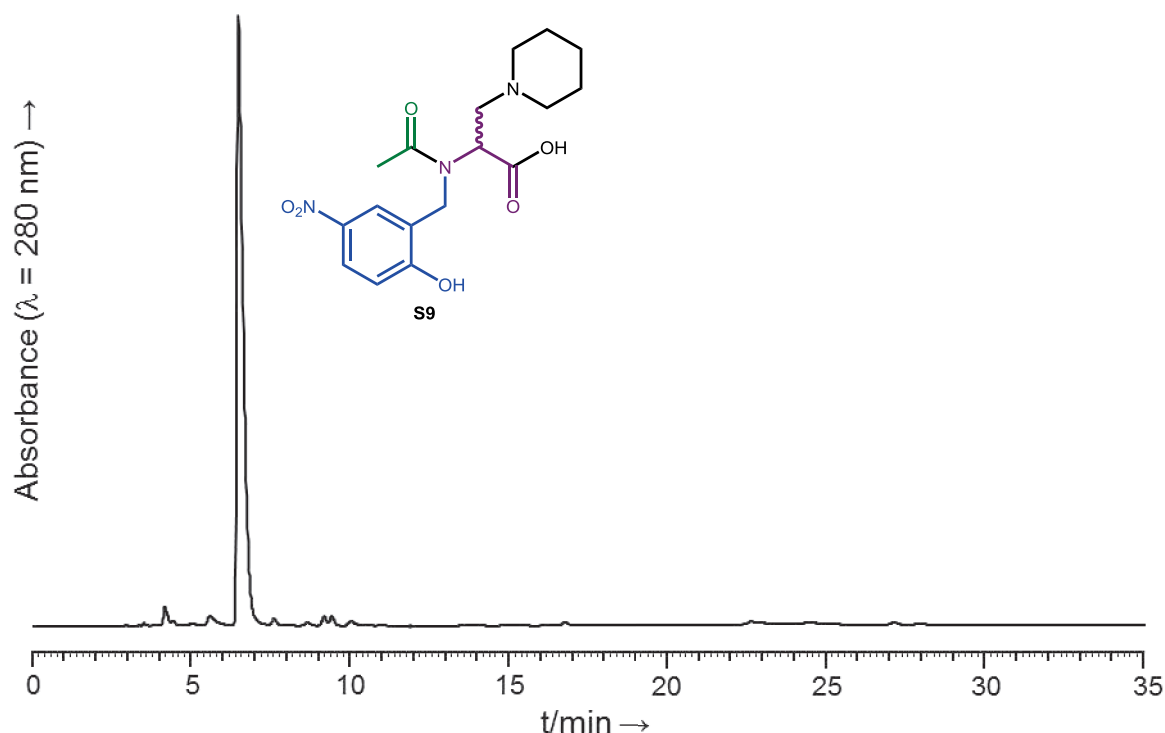
Pre-loaded Fmoc-Cys(SfBu)-PHB Tentagel resin (527 mg, 0.19 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). For a 10  $\mu$ mol aliquot, acetic acid (6  $\mu$ L, 0.1 mmol, 10 equiv.) and HCTU (37.2 mg, 0.09 mmol, 9 equiv.) were dissolved in NMP (0.2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (35  $\mu$ L, 0.2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Amino acid-resin was then treated with 20% piperidine in NMP (1 mL, 3  $\times$  3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>. Finally, the product was cleaved using TFA / H<sub>2</sub>O (95:5) for 2h. The solvents were removed by evaporation and analysis showed the complete formation of the C-terminal piperidine adduct **S9** and no trace of the expected Cys(SfBu) compound.

**S9:**

**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>17</sub>H<sub>24</sub>N<sub>3</sub>O<sub>6</sub>: 366.1665, found: 366.1662.

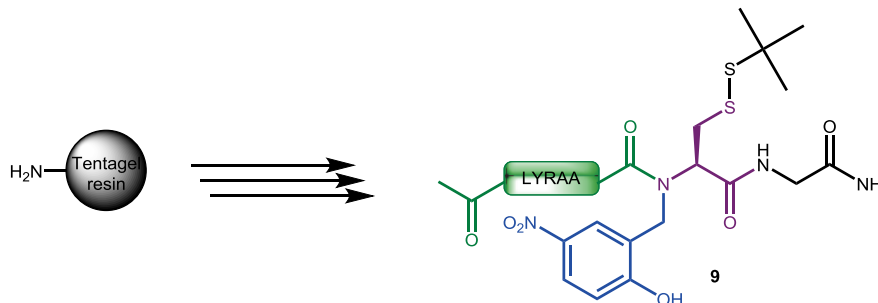
**HPLC**: retention time: 6.50 min (Nucleosil, gradient: 20-50% B/A over 30 min).

<sup>2</sup> Lukszo J., Patterson D., Albericio F., Kates S. A., *Letters in Peptide Science*, 1996, **3**, 157-166



Supplementary figure S28: HPLC trace of crude **S9**.

### 7c- Introduction of a Gly residue as a spacer



Supplementary scheme S13: Synthesis of the model peptide **9**.

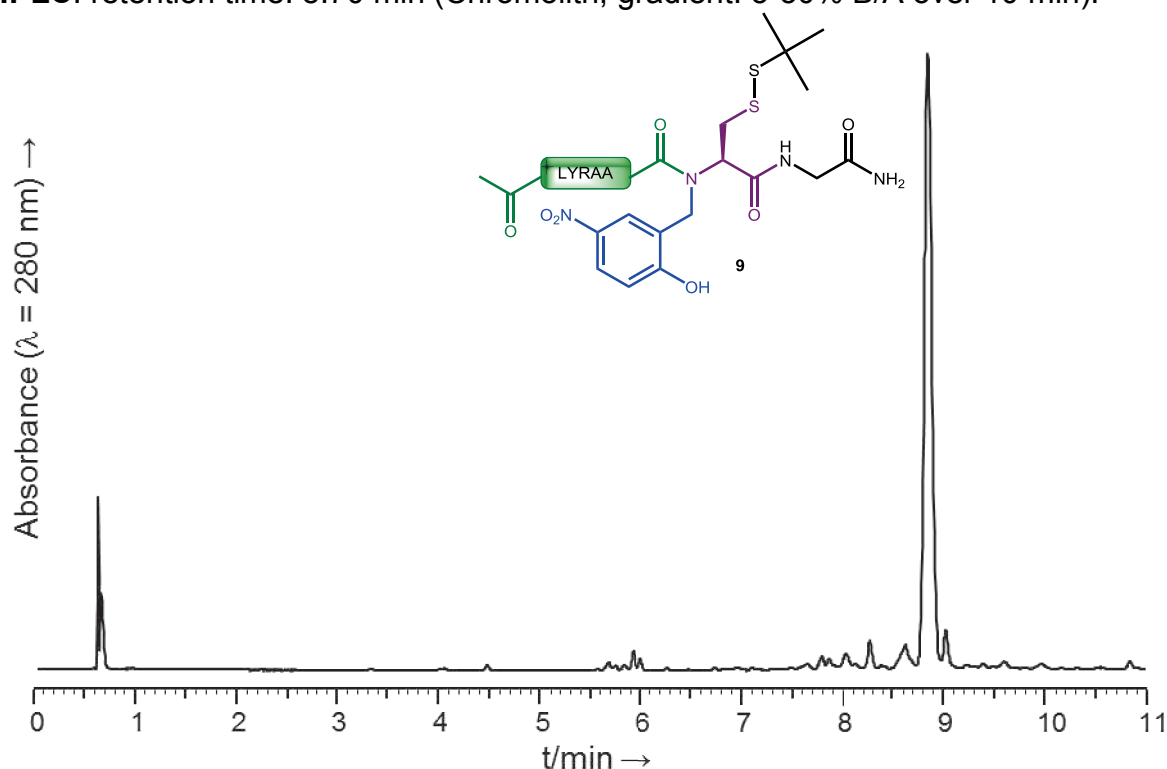
Tentagel R resin (476 mg, 0.21 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Rink-OH (270 mg, 0.5 mmol, 5 equiv.) and HATU (190 mg, 0.5 mmol, 5 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175 μL, 1 mmol, 10 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Gly-OH (297 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348 μL, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was

washed with NMP. Then, Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Cys(*St*Bu)-OH (432 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (10 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). The Ac-LYRAA sequence was installed by standard SPPS procedure (protocol p S3), an extended double coupling (2  $\times$  2h) was performed for the first alanine and a normal double coupling (2  $\times$  30 min) was performed for the arginine. Peptide resin was then treated with 20% piperidine in NMP (10 mL, 3  $\times$  3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>, and finally cleaved following the general procedure p S3 to give compound **9**. The product was purified by HPLC (Nucleosil, 40-45 % B/A over 10 min).

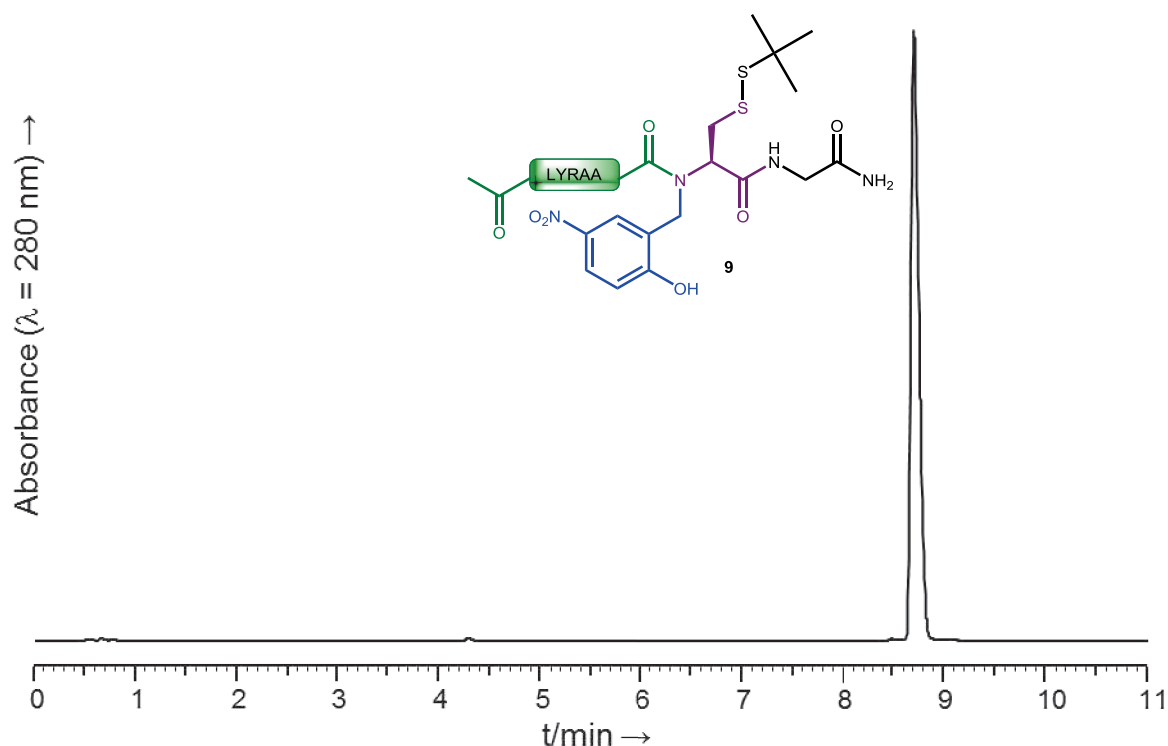
**9:**

**MALDI-TOF** ( $m/z$ ): [MH]<sup>+</sup> calcd for C<sub>45</sub>H<sub>69</sub>N<sub>12</sub>O<sub>12</sub>S<sub>2</sub>: 1033.45, found: 1033.49.

**HPLC**: retention time: 8.70 min (Chromolith, gradient: 5-50% B/A over 10 min).

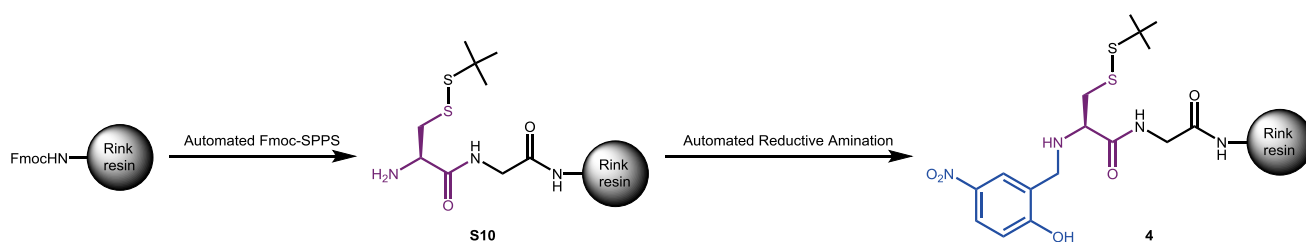


Supplementary figure S29: HPLC trace of crude **9**.



Supplementary figure S30: HPLC trace of pure 9.

### 8- N-Acylation with the 20 different proteogenic amino acids



Supplementary scheme S14: Synthesis of (Hnb)-dipeptide 4.

The Rink linker, Gly and Cys(*St*Bu) were installed through automated SPPS (protocol p S3) starting from Tentagel R resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol) in order to obtain peptide resin **S10**.

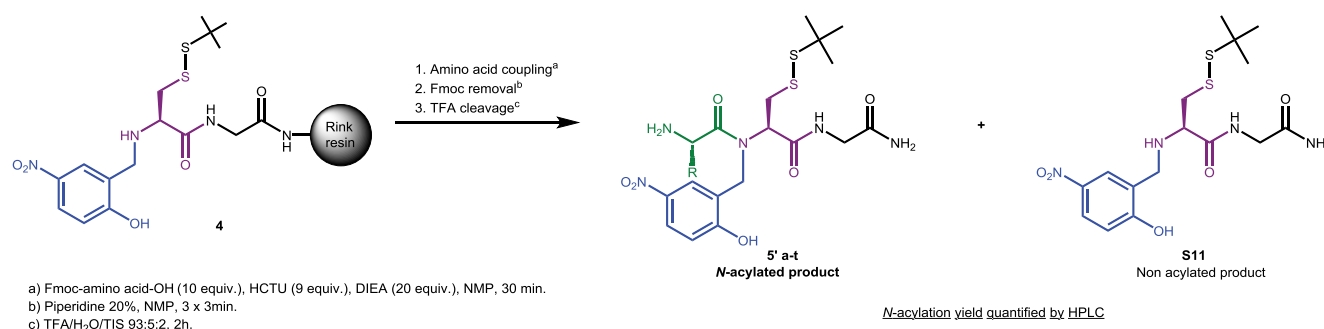
#### Automated reductive amination procedure for a peptide synthesizer

Peptide resin **S10** (25  $\mu$ mol) was introduced into a Prelude synthesizer reactor, washed two times with 3 mL of a DMF/MeOH (1:1) mixture for 30 sec, then swollen in 3 mL of a DMF/MeOH/AcOH (9:9:2) mixture for 5 min. The reactor was drained off and the resin was washed four times with 3 mL of a DMF/MeOH (1:1) mixture for 30 s. 2 mL of a solution of 2-hydroxy-5-nitrobenzaldehyde in DMF/MeOH (1:1) (125 mM, 10 equiv.) were added then the reactor was left for 1h under stirring by nitrogen bubbling. The reactor was drained off and the resin was washed four times with 3 mL of DMF/MeOH (1:1) for 15 sec. Without delay, 2 mL of a fresh solution of sodium

cyanoborohydride in a DMF/MeOH/AcOH (9:9:2) mixture (250 mM, 20 equiv.) were added and the reactor was left for 1 h under stirring by nitrogen bubbling. The reactor was drained and the resin was washed with DMF/MeOH (1:1) (3 mL, 30 sec, × 4), NMP (3 mL, 30 s, × 3), 20% piperidine in NMP (3 mL, 30 s, × 3), NMP (3 mL, 30 sec, × 3), dichloromethane (5 mL, 30 s, × 3) and NMP (3 mL, 30 s, × 2).

Conditioning for reductive amination			
Operation	Solvent	Volume	Time
1) DMF/MeOH Wash	DMF/MeOH (1:1)	3 mL × 2	0.5 min × 2
2) DMF/MeOH/AcOH Wash	DMF/MeOH/AcOH (9:9:2)	3 mL	5 min
3) DMF/MeOH Wash	DMF/MeOH (1:1)	3 mL × 4	0.5 min × 4
Imine formation			
Operation	Solvent	Volume	Time
1) Add 2-hydroxy- 5-nitrobenzaldehyde (Hnba)	125 mM Hnba in DMF/MeOH (1:1)	2 mL	60 min
2) DMF/MeOH Wash	DMF/MeOH (1:1)	3 mL × 4	0.25 min × 4
Reduction			
Operation	Solvent	Volume	Time
1) Add NaBH <sub>3</sub> CN	250 mM NaBH <sub>3</sub> CN in DMF/MeOH/AcOH (9:9:2)	2 mL	60 min
2) DMF/MeOH Wash	DMF/MeOH (1:1)	3 mL × 4	0.5 min × 4
3) NMP Wash	NMP	3 mL × 3	0.5 min × 3
4) Base wash	20 % piperidine in NMP	3 mL × 3	3 min × 3
5) NMP Wash	NMP	3 mL × 3	0.5 min × 3
6) Dichloromethane Wash	CH <sub>2</sub> Cl <sub>2</sub>	5 mL × 3	0.5 min × 3
7) NMP Wash	NMP	3 mL × 2	0.5 min × 2

### N-acylation yield determination



### Supplementary scheme S15: N-Acylation of Hnb-peptide-resin 4.

The 20 different proteogenic amino acids were then coupled on **4** using a single standard automated coupling on the prelude (protocol p S3), after that the peptide resins were treated with 20 % piperidine in NMP (3 × 3 min); finally, an aliquot (5 μmol) of each resins was cleaved using 1mL of a TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH (93:5:2) solution for 2 h. After evaporation of TFA, the samples were analyzed by HPLC; the *N*-acylation yields were quantified by integration at 315 nm, not taking into account eventual differences in molar extinction coefficient of the products.

**Non acylated product - S11:**

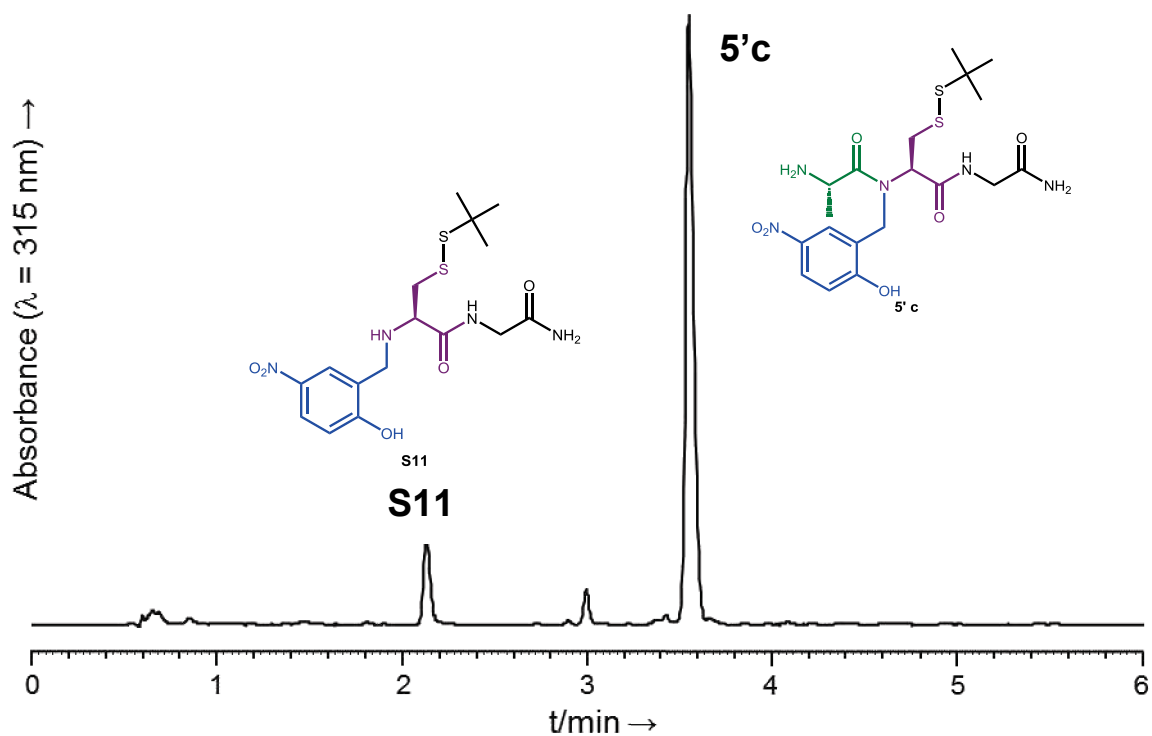
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{16}H_{25}N_4O_5S_2$ : 417.1266, found: 417.1261.

**HPLC**: retention time: 2.12 min (Chromolith, gradient: 20-50 % B/A over 5 min).

**Ala - 5'c:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{19}H_{30}N_5O_6S_2$ : 488.1637, found: 488.1636.

**HPLC**: retention time: 3.55 min (Chromolith, gradient: 20-50 % B/A over 5 min).



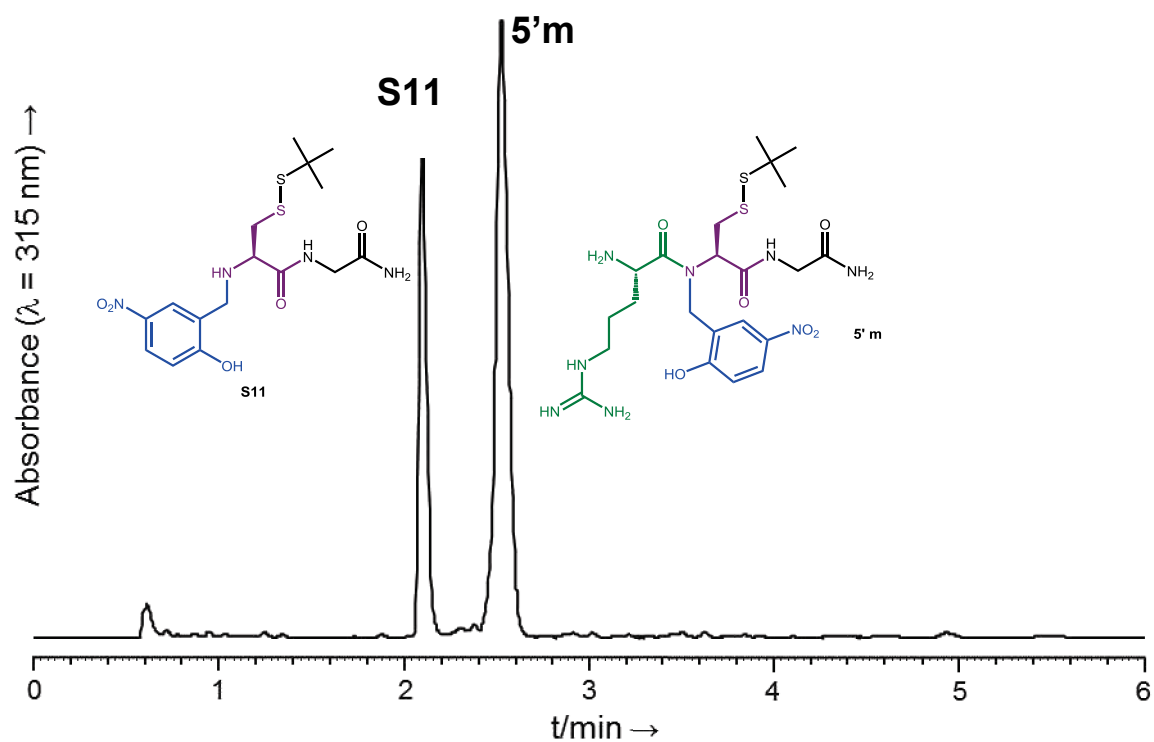
Supplementary figure S31: HPLC trace of crude **5'c**.



**Arg - 5'm:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{22}H_{37}N_8O_6S_2$ : 573.2277, found: 573.2272.

**HPLC**: retention time: 2.54 min (Chromolith, gradient: 20-50 % B/A over 5 min).

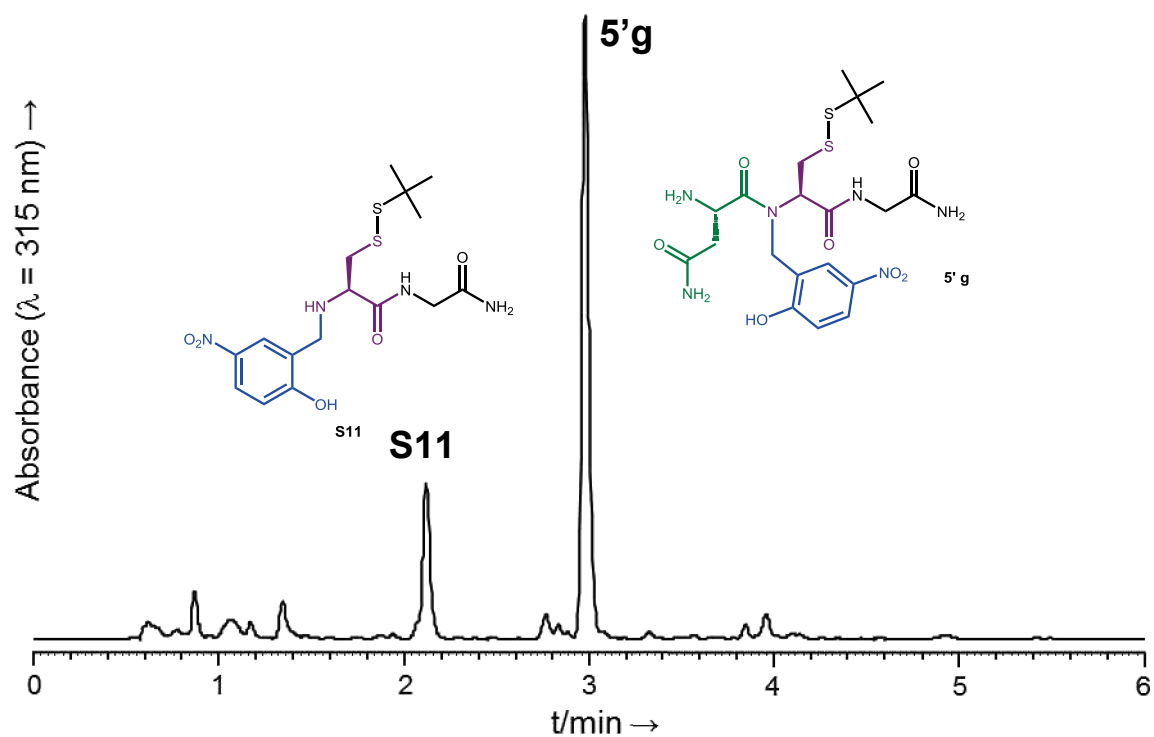


Supplementary figure S32: HPLC trace of crude **5'm**.

**Asn - 5'g:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{20}H_{31}N_6O_7S_2$ : 531.1696, found: 531.1690.

**HPLC**: retention time: 2.97 min (Chromolith, gradient: 20-50 % B/A over 5 min).

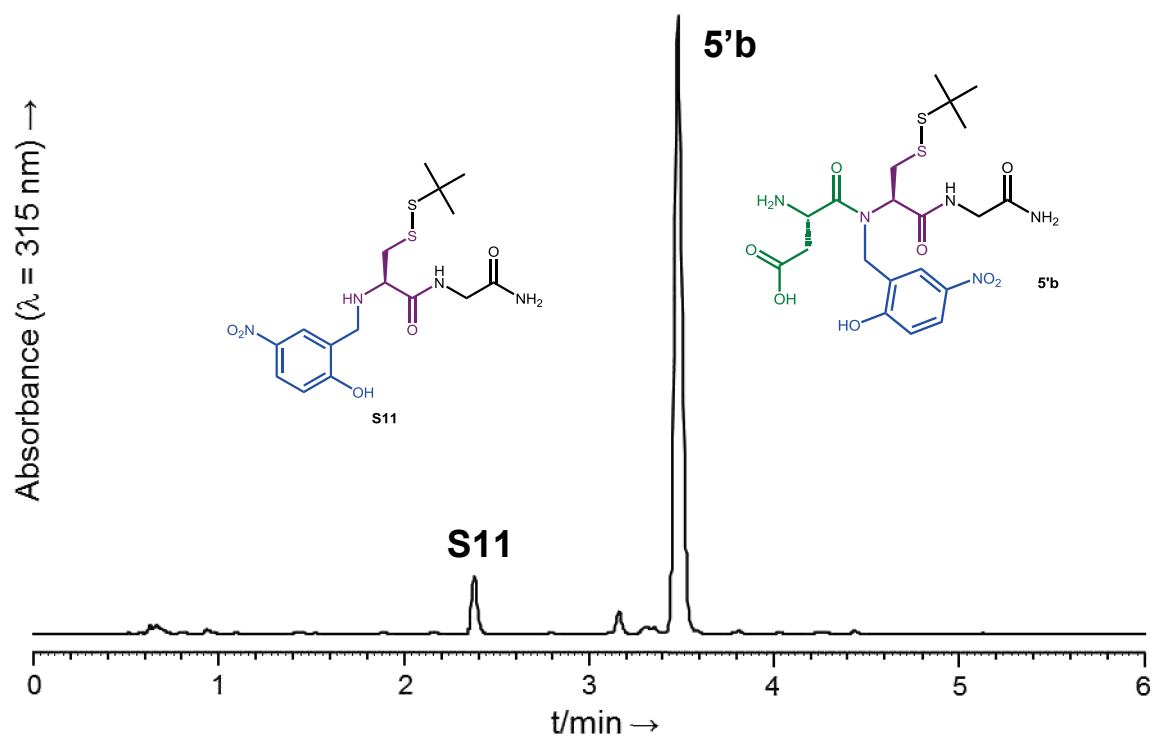


Supplementary figure S33: HPLC trace of crude **5'g**.

**Asp - 5'b:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{20}H_{30}N_5O_8S_2$ : 532.1536, found: 532.1528.

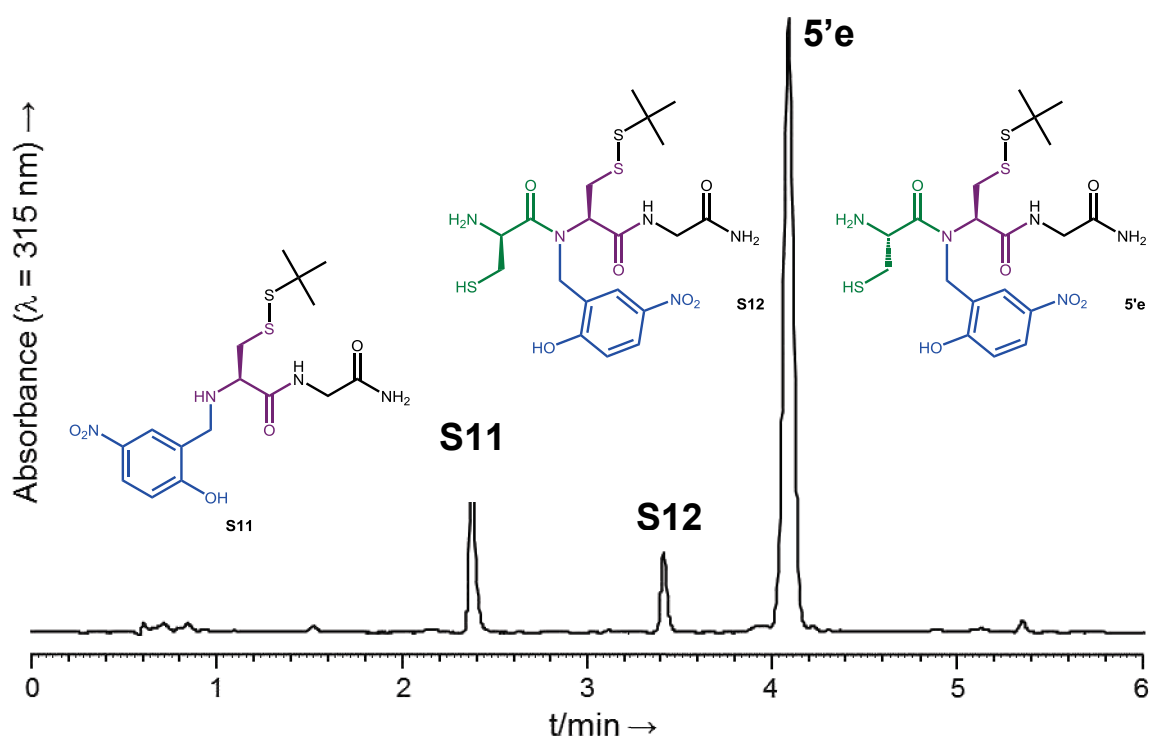
**HPLC**: retention time: 3.48 min (Chromolith, gradient: 20-50 % B/A over 5 min).



Supplementary figure S34: HPLC trace of crude **5'b**.

**Cys - 5'e:****ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{19}H_{30}N_5O_6S_3$ : 520.1358, found: 520.1351.**HPLC**: retention time: 4.07 min (Chromolith, gradient: 20-50 % B/A over 5 min).

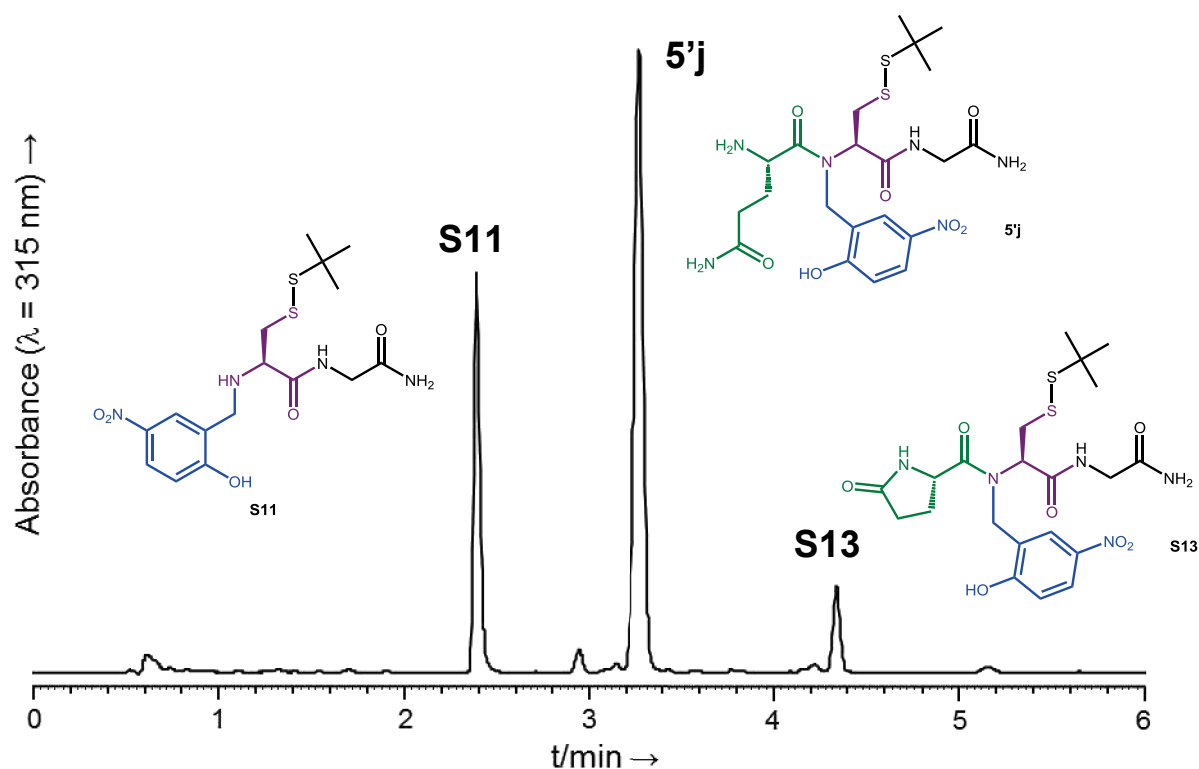
In the case of compound **5'e**, a minor peak was observed showing the same  $m/z$  as expected for compound **5'e**. It was attributed to epimerization of Fmoc-Cys(Trt) during coupling, yielding compound **S12** after TFA cleavage (ratio: L-Cys / D-Cys 93:7).

**D-Cys - S12:****ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{19}H_{30}N_5O_6S_3$ : 520.1358, found: 520.1353.**HPLC**: retention time: 3.41 min (Chromolith, gradient: 20-50 % B/A over 5 min).

Supplementary figure S35: HPLC trace of crude **5'e**.

**Gln - 5'j:****ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{33}N_6O_7S_2$ : 545.1852, found: 545.1847.**HPLC**: retention time: 3.27 min (Chromolith, gradient: 20-50 % B/A over 5 min).

In the case of compound **5'j**, concomitant formation of pyroglutamate (**S13**) was observed during the TFA treatment.

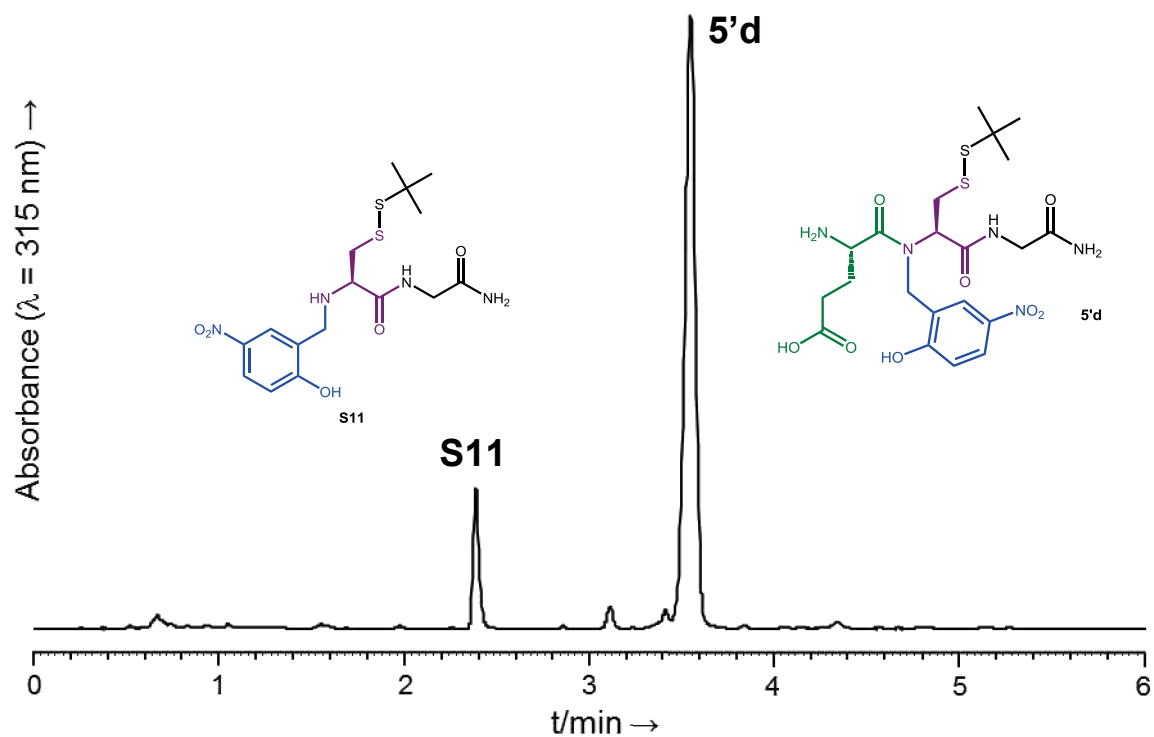
**PyroGlu - S13:****ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{30}N_5O_7S_2$ : 528.1587, found: 528.1579.**HPLC**: retention time: 4.34 min (Chromolith, gradient: 20-50 % B/A over 5 min).

Supplementary figure S36: HPLC trace of crude **5'j**.

**Glu 5'd:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{32}N_5O_8S_2$ : 546.1692, found: 546.1684.

**HPLC**: retention time: 3.55 min (Chromolith, gradient: 20-50 % B/A over 5 min).

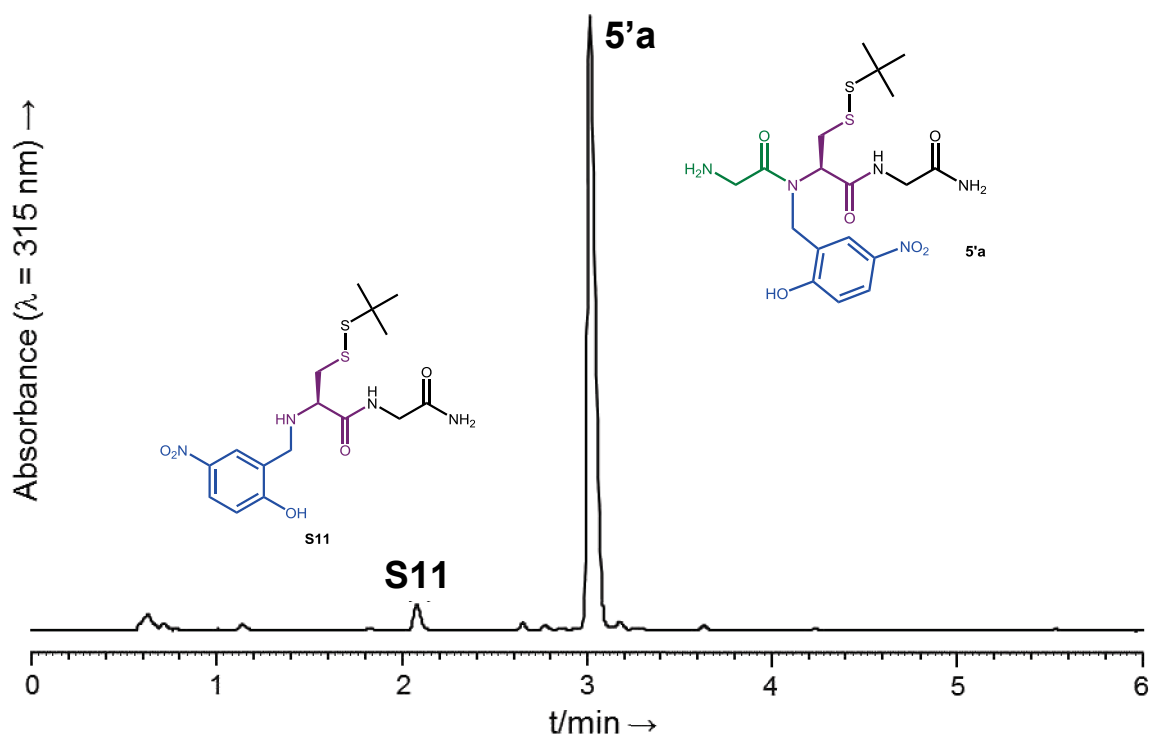


Supplementary figure S37: HPLC trace of crude **5'd**.

**Gly - 5'a:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{18}H_{28}N_5O_6S_2$ : 474.1481, found: 474.1476.

**HPLC**: retention time: 3.01 min (Chromolith, gradient: 20-50 % B/A over 5 min).

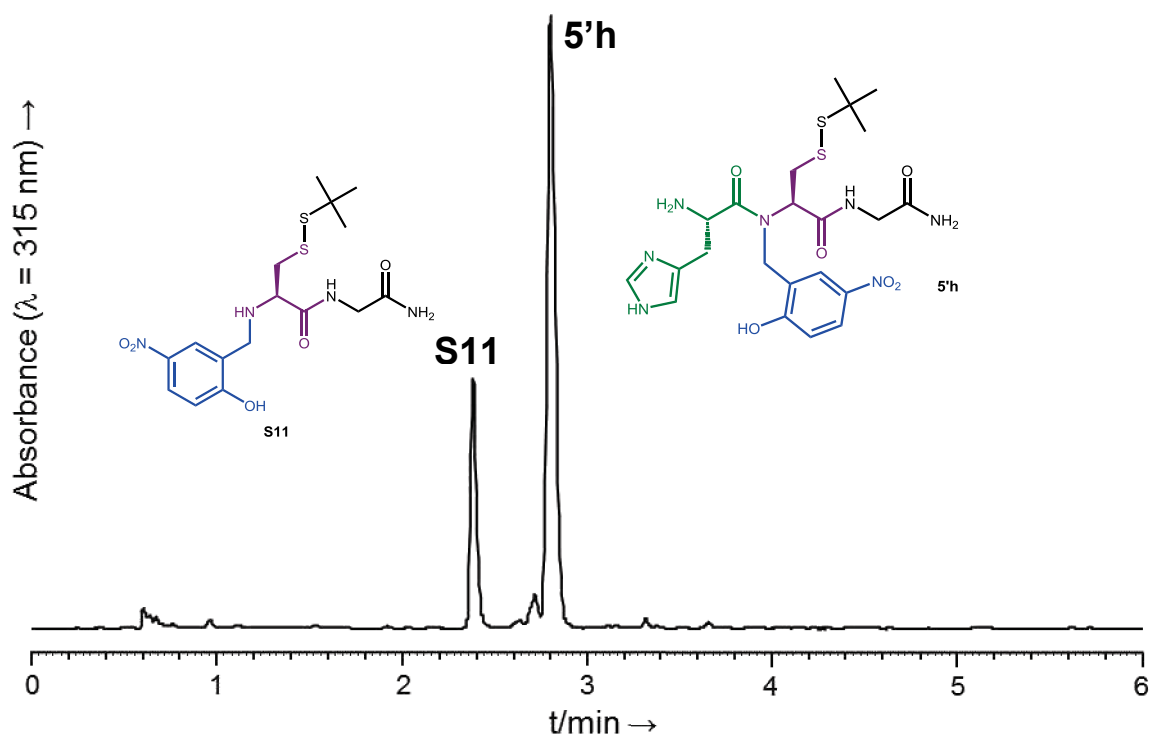


Supplementary figure S38: HPLC trace of crude **5'a**.

**His - 5'h:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{22}H_{32}N_7O_6S_2$ : 554.1855, found: 554.1848.

**HPLC**: retention time: 2.80 min (Chromolith, gradient: 20-50 % B/A over 5 min).



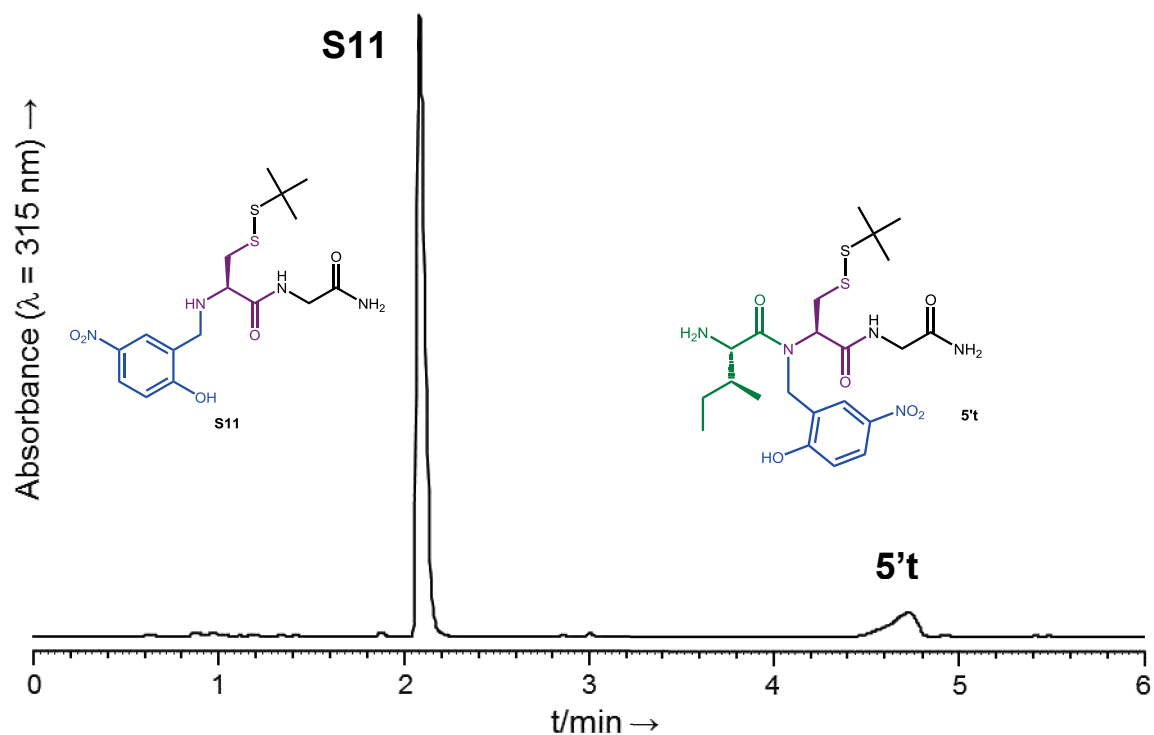
Supplementary figure S39: HPLC trace of crude **5'h**.



Ile - 5't:

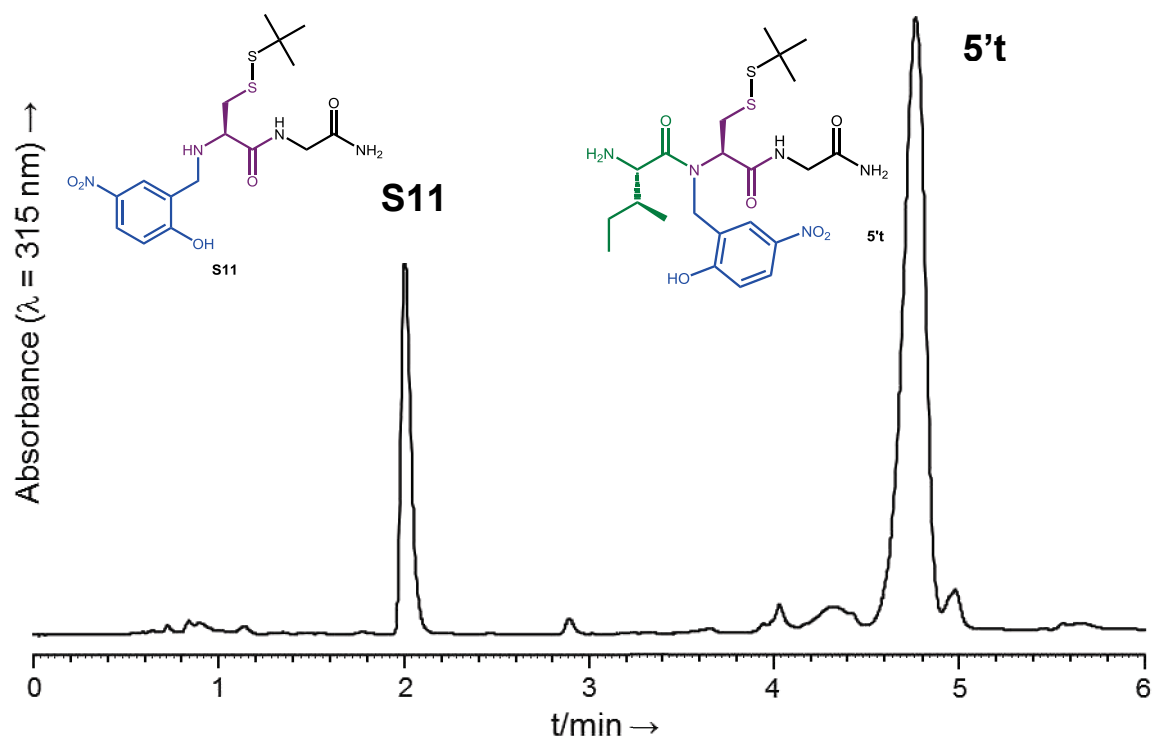
ESI-HRMS ( $m/z$ ):  $[MH]^+$  calcd for  $C_{22}H_{36}N_5O_6S_2$ : 530.2107, found: 530.2104.

HPLC: retention time: 4.72 min (Chromolith, gradient: 20-50 % B/A over 5 min).



Supplementary figure S40: HPLC trace of crude 5't.

Fmoc-Ile-OH (0.05 mmol, 10 equiv.) was coupled using a Biotage Initiator+ SP Wave, on 5  $\mu$ mol of resin **4** using HCTU (19.6 mg, 0.0476 mmol, 9.5 equiv.) and *i*Pr<sub>2</sub>NEt (17.4  $\mu$ L, 0.1 mmol, 20 equiv.) in NMP (0.6 mL) under microwave heating at 70°C for 30 min. After that the peptide resin was treated with 20 % piperidine in NMP (3  $\times$  3 min); finally, resin was cleaved using 1mL of a TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH (93:5:2) solution for 2 h. After evaporation of TFA, the sample was analyzed by HPLC; the *N*-acylation yield was quantified by integration at 315 nm, not taking into account eventual differences in molar extinction coefficient of the products.

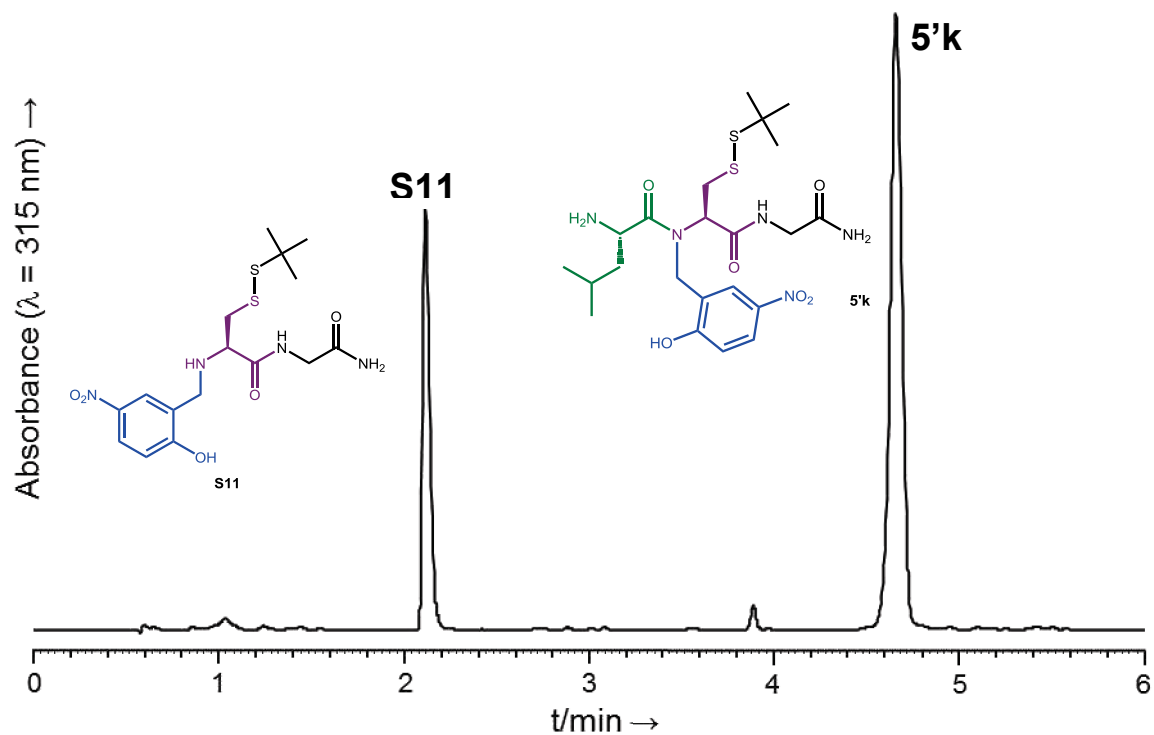


Supplementary figure S41: HPLC trace of crude **5't** after microwave-assisted coupling.

**Leu - 5'k:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{22}H_{36}N_5O_6S_2$ : 530.2107, found: 530.2106.

**HPLC**: retention time: 4.66 min (Chromolith, gradient: 20-50 % B/A over 5 min).

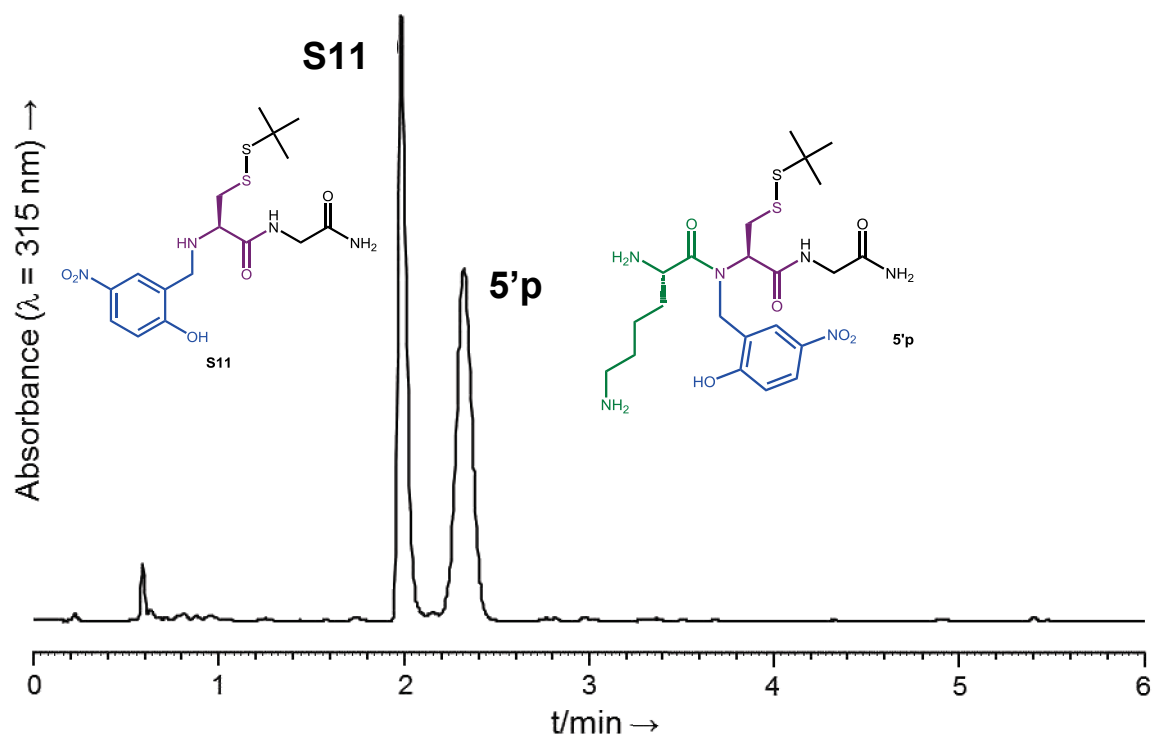


Supplementary figure S42: HPLC trace of crude **5'k**.

**Lys - 5'p:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{22}H_{37}N_6O_6S_2$ : 545.2216, found: 545.2210.

**HPLC**: retention time: 2.40 min (Chromolith, gradient: 20-50 % B/A over 5 min).

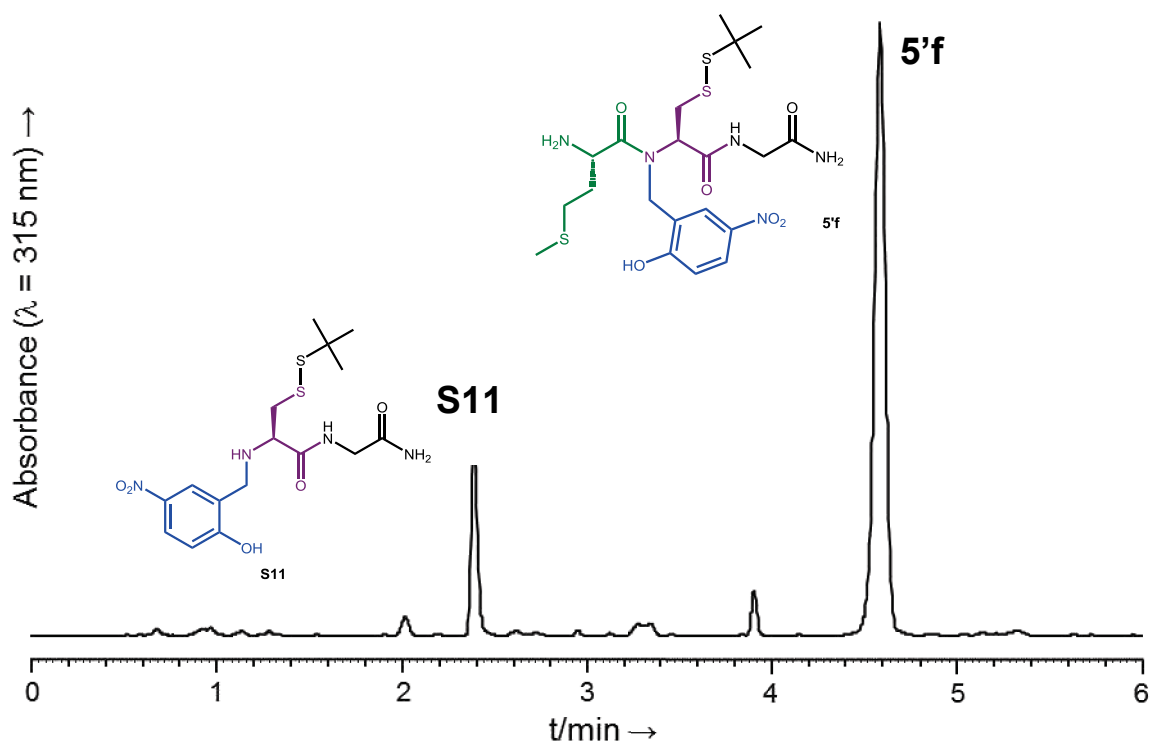


Supplementary figure S43: HPLC trace of crude **5'p**.

**Met - 5'f:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{34}N_5O_6S_3$ : 548.1671, found: 548.1665.

**HPLC**: retention time: 4.56 min (Chromolith, gradient: 20-50 % B/A over 5 min).

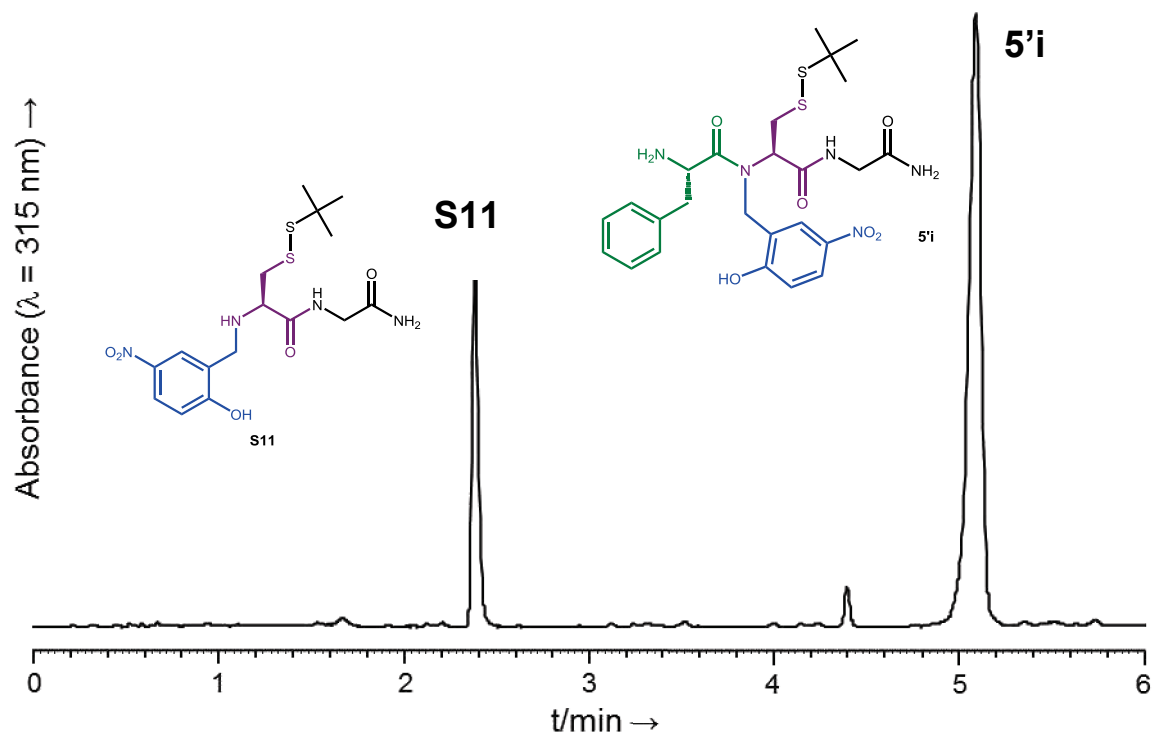


Supplementary figure S44: HPLC trace of crude **5'f**.

**Phe - 5'i:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{25}H_{34}N_5O_6S_2$ : 564.1950, found: 564.1942.

**HPLC**: retention time: 5.07 min (Chromolith, gradient: 20-50 % B/A over 5 min).

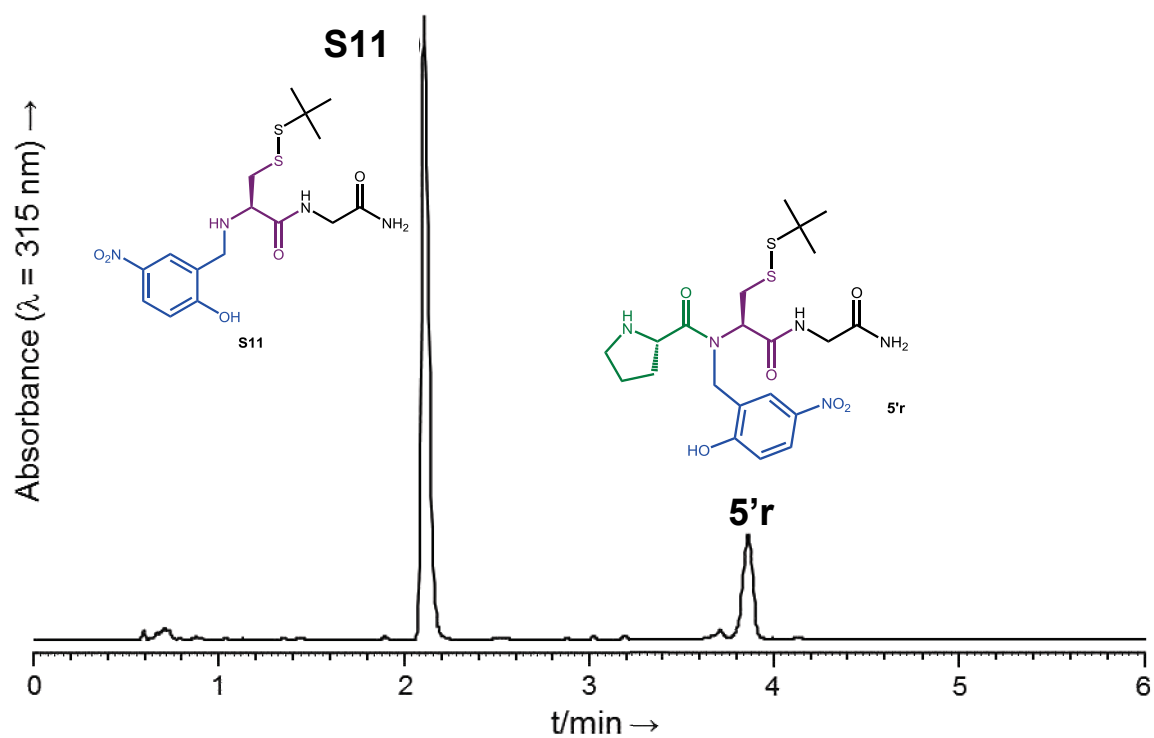


Supplementary figure S45: HPLC trace of crude **5'i**.

**Pro - 5'r:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{32}N_5O_6S_2$ : 514.1794, found: 514.1788.

**HPLC**: retention time: 3.86 min (Chromolith, gradient: 20-50 % B/A over 5 min).

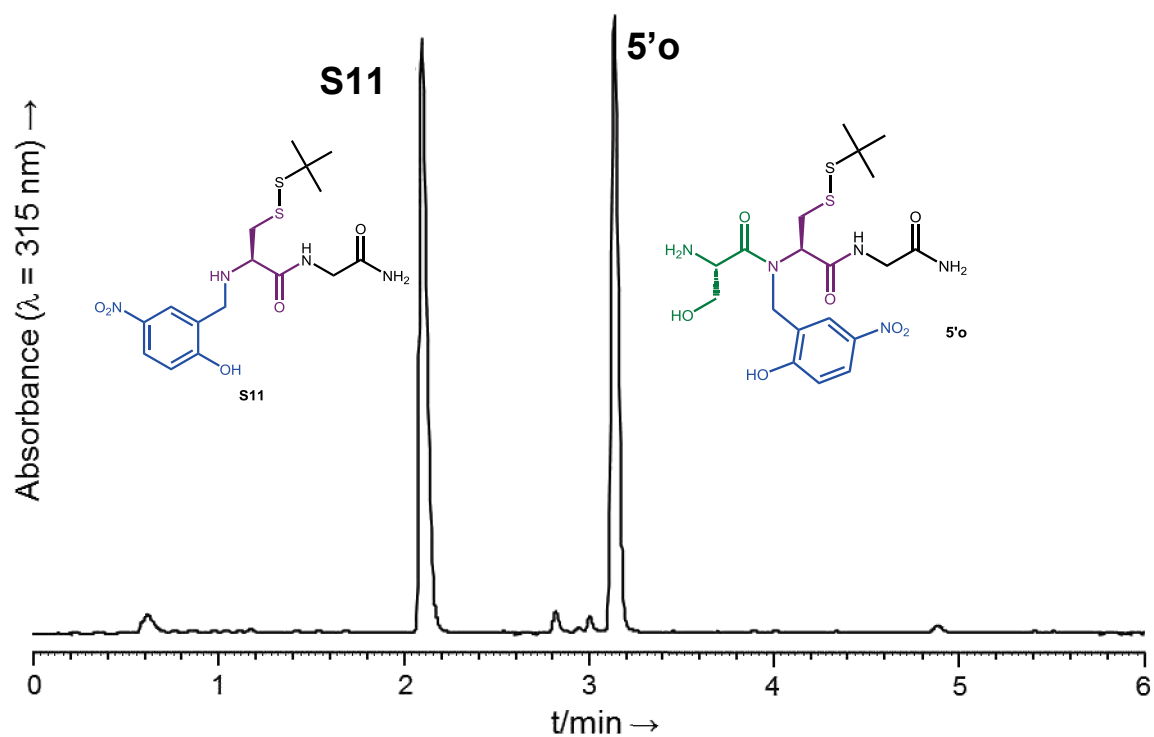


Supplementary figure S46: HPLC trace of crude **5'r**.

**Ser - 5'o:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{19}H_{30}N_5O_7S_2$ : 504.1587, found: 504.1580.

**HPLC**: retention time: 3.13 min (Chromolith, gradient: 20-50 % B/A over 5 min).



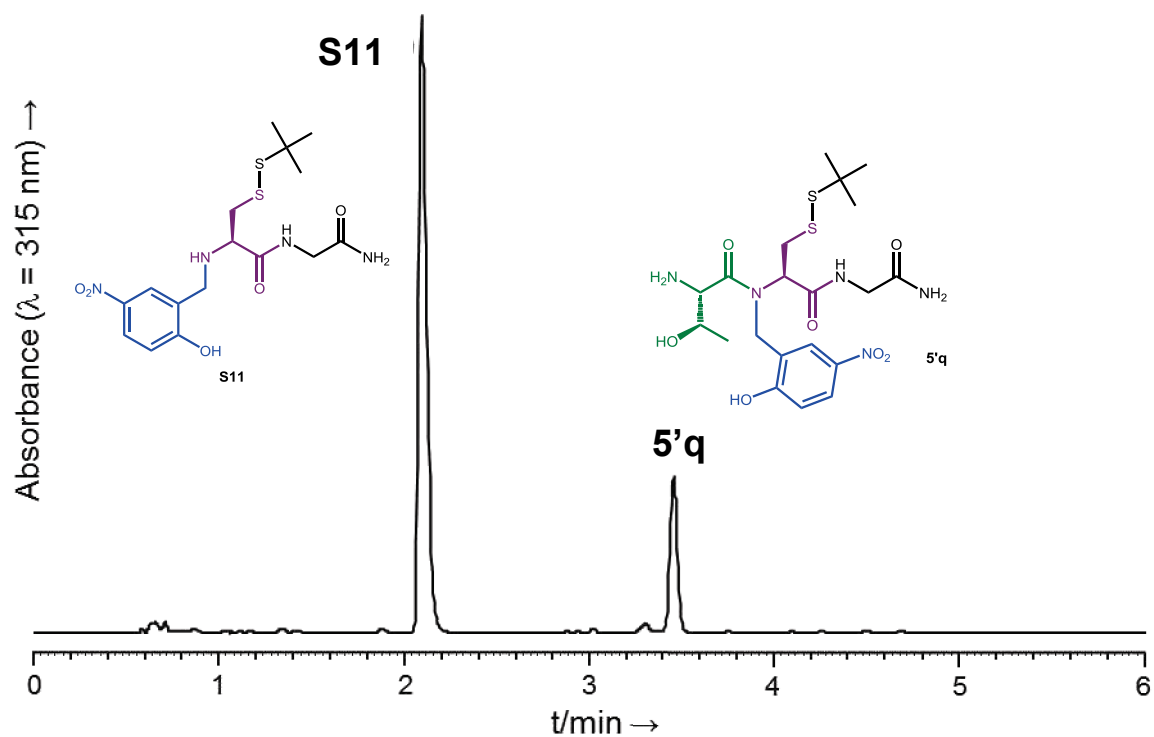
Supplementary figure S47: HPLC trace of crude **5'o**.



**Thr - 5'q:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{20}H_{32}N_5O_7S_2$ : 518.1743, found: 518.1736.

**HPLC**: retention time: 3.46 min (Chromolith, gradient: 20-50 % B/A over 5 min).

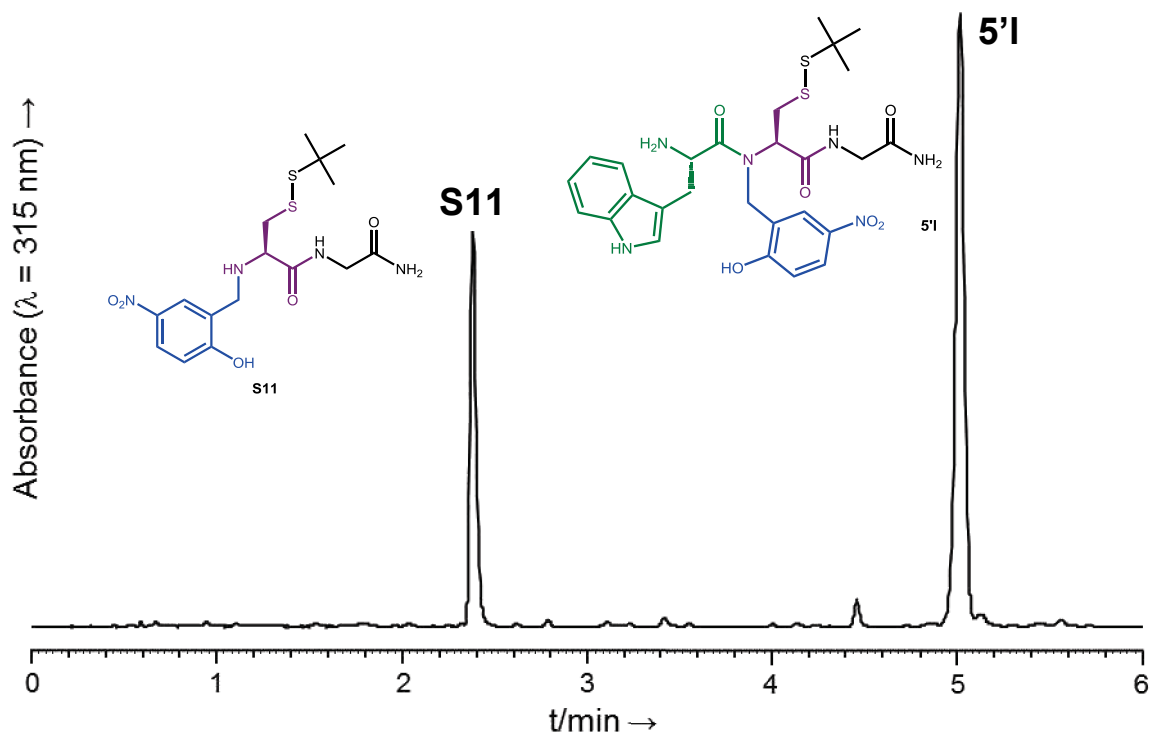


Supplementary figure S48: HPLC trace of crude **5'q**.

**Trp - 5'I:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{27}H_{35}N_6O_6S_2$ : 603.2059, found: 603.2052.

**HPLC**: retention time: 5.00 min (Chromolith, gradient: 20-50 % B/A over 5 min).

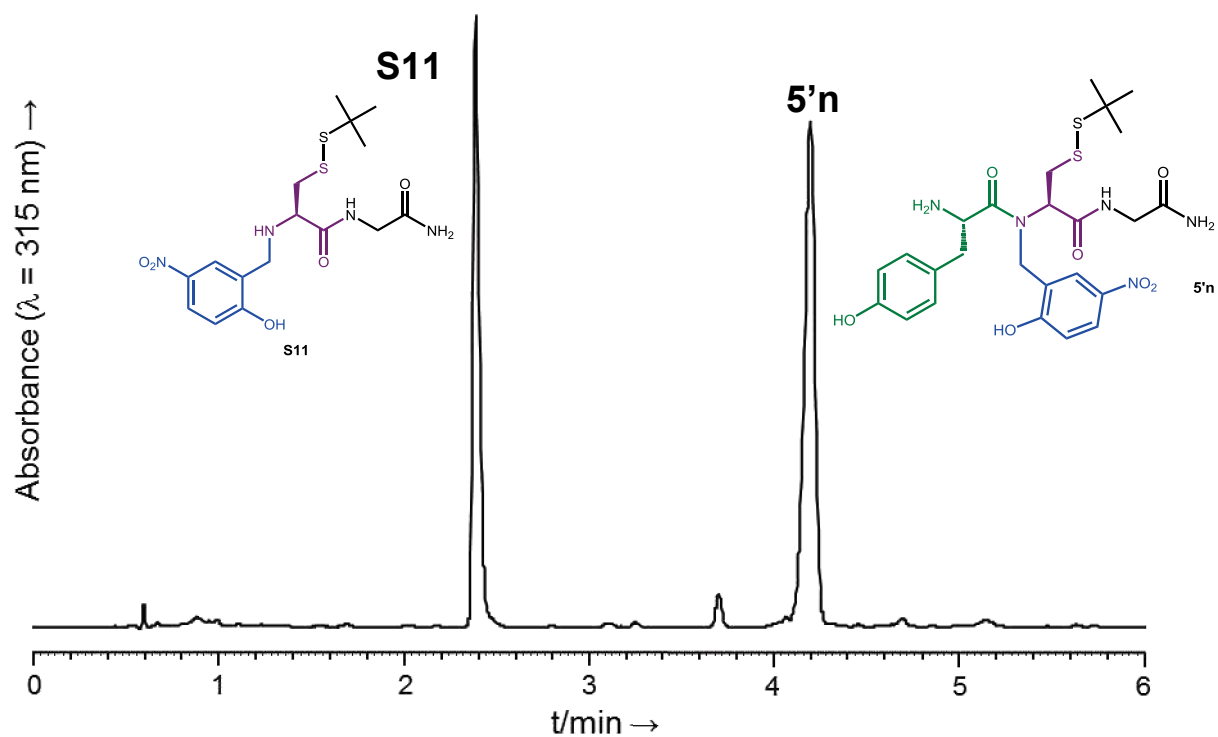


Supplementary figure S49: HPLC trace of crude **5'I**.

**Tyr - 5'n:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{25}H_{34}N_5O_7S_2$ : 580.1900, found: 580.1892.

**HPLC**: retention time: 4.18 min (Chromolith, gradient: 20-50 % B/A over 5 min).

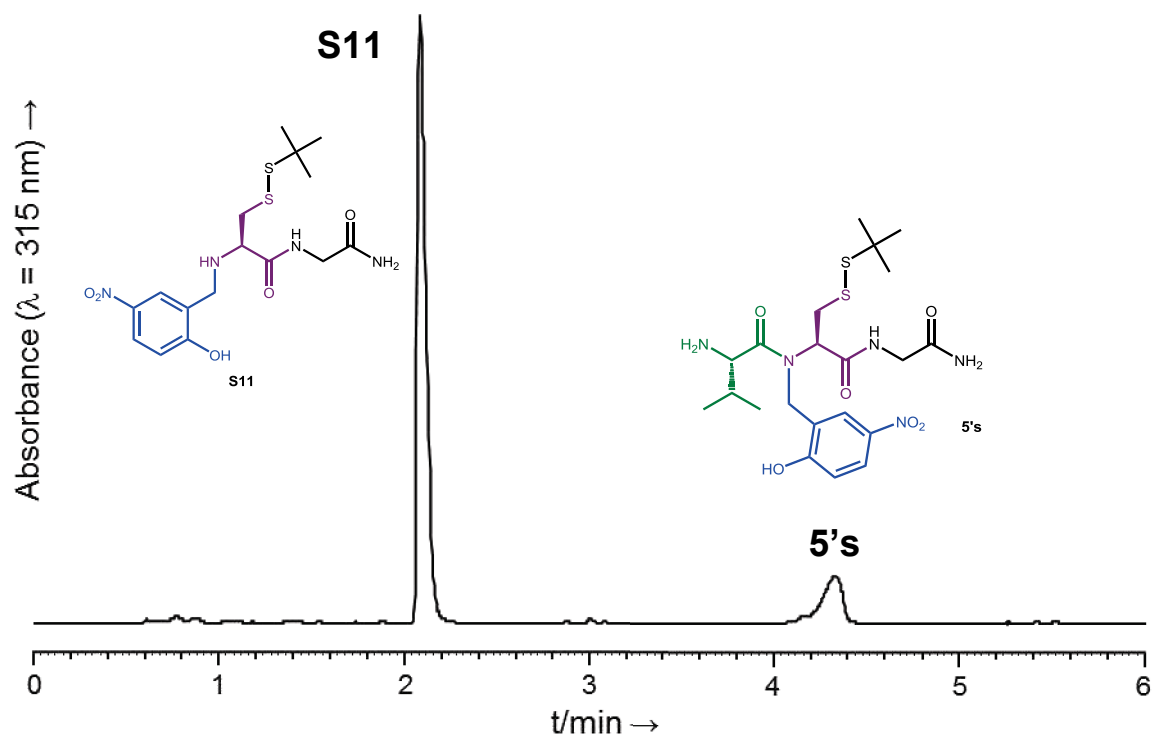


Supplementary figure S50: HPLC trace of crude **5'n**.

**Val - 5's:**

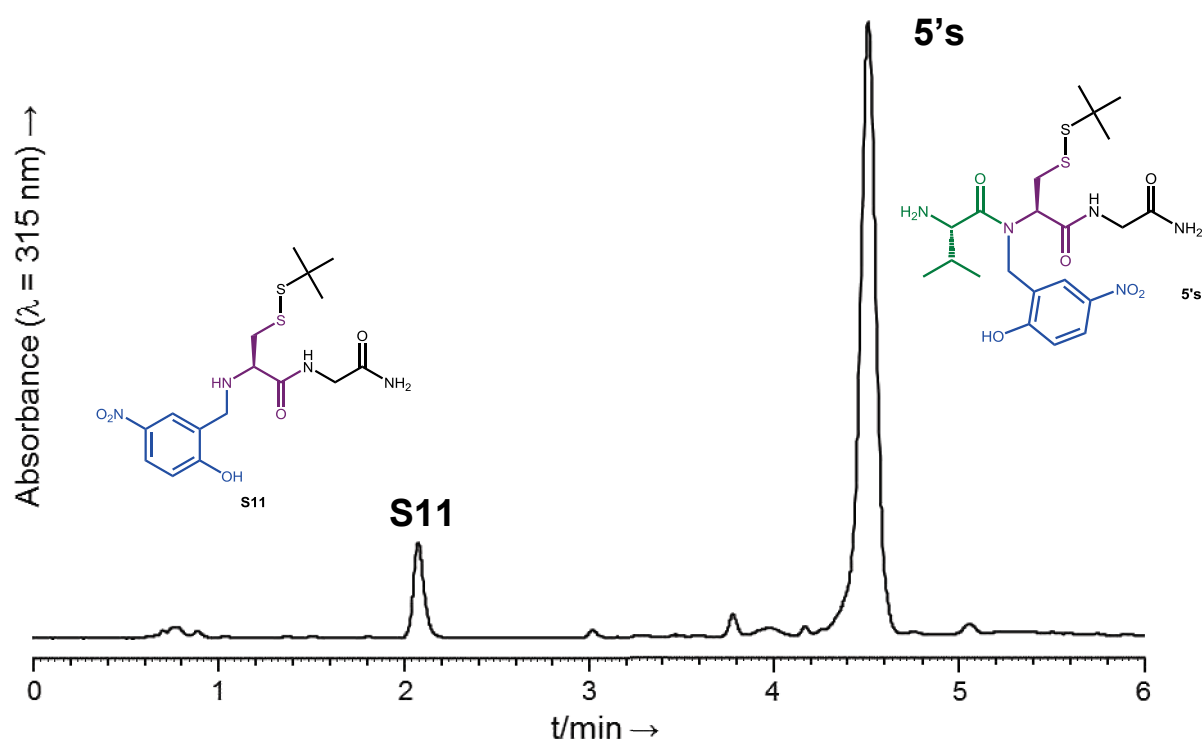
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{21}H_{34}N_5O_6S_2$ : 516.1950, found: 516.1947.

**HPLC**: retention time: 4.33 min (Chromolith, gradient: 20-50 % B/A over 5 min).



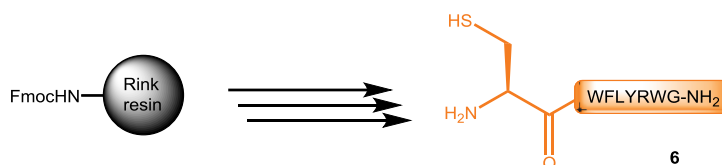
Supplementary figure S51: HPLC trace of crude **5's**.

Fmoc-Val-OH (0.05 mmol, 10 equiv.) was coupled using a Biotage Initiator+ SP Wave, on 5  $\mu$ mol of resin **4** using HCTU (19.6 mg, 0.0476 mmol, 9.5 equiv.) and *i*Pr<sub>2</sub>NEt (17.4  $\mu$ L, 0.1 mmol, 20 equiv.) in NMP (0.6 mL) under microwave heating at 70°C for 30 min. After that the peptide resin was treated with 20 % piperidine in NMP (3  $\times$  3 min); finally, resin was cleaved using 1mL of a TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH (93:5:2) solution for 2 h. After evaporation of TFA, the sample was analyzed by HPLC; the *N*-acylation yield was quantified by integration at 315 nm, not taking into account eventual differences in molar extinction coefficient of the products.



Supplementary figure S52: HPLC trace of crude **5's** after microwave-assisted coupling.

## 9- Kinetics studies of the NCL with model Ac-LYRAA(Hnb)C peptides

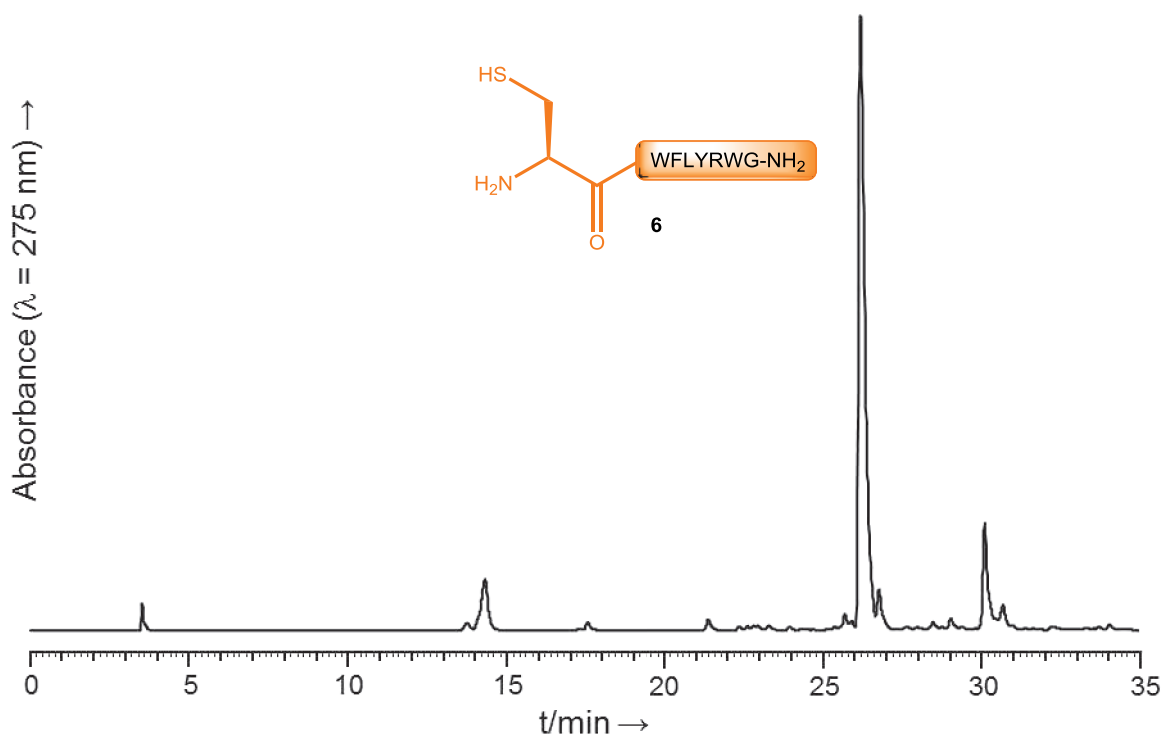
9a- Synthesis of the model cysteinyl peptide 6Supplementary scheme S16: Synthesis of the model cysteinyl peptide 6.

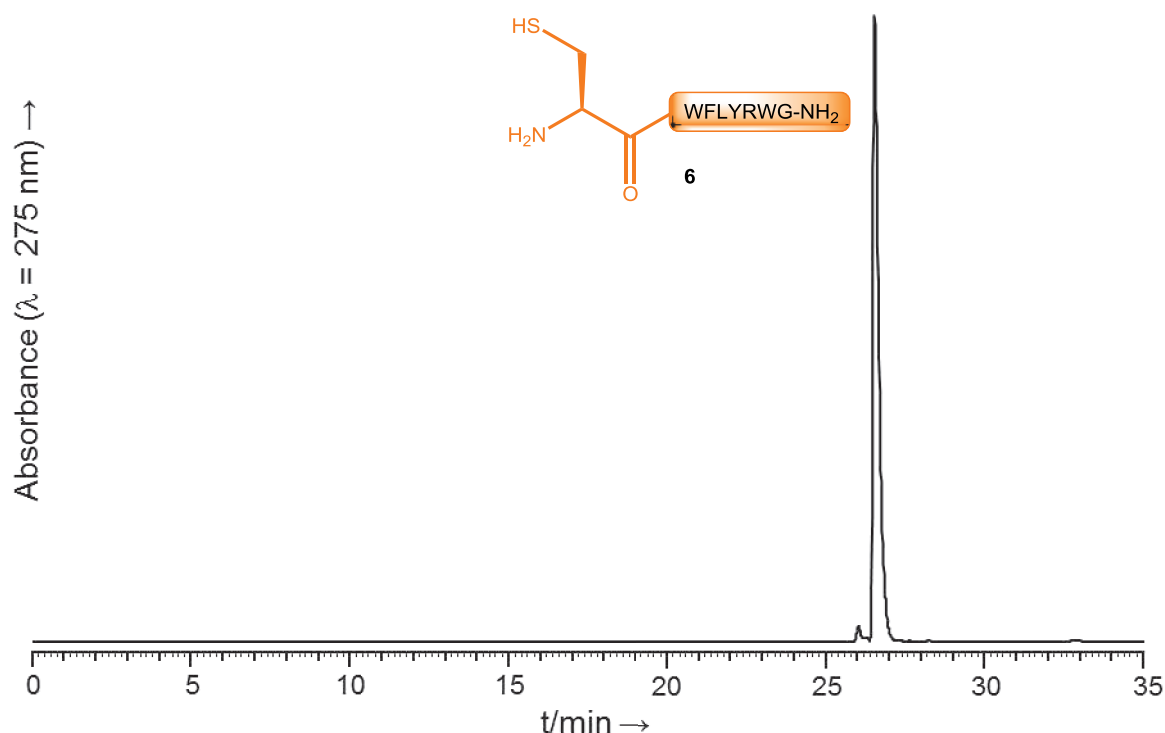
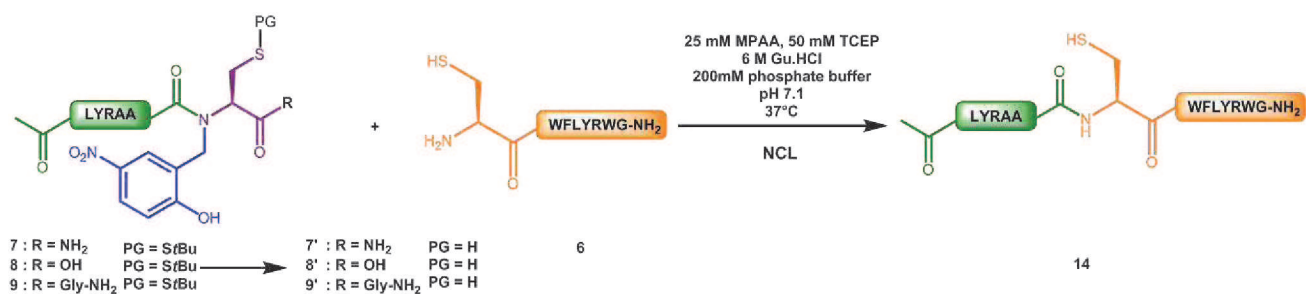
Peptide **6** was obtained through automated SPPS (protocol p S3) starting from Fmoc-Rink polystyrene resin (130 mg, 0.79 mmol/g, 0.1 mmol). The peptide-resin was cleaved (protocol p S3) and **6** was purified by HPLC (Nucleosil, gradient: 32-39 % B/A over 10 min).

**6:**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for C<sub>57</sub>H<sub>73</sub>N<sub>14</sub>O<sub>9</sub>S: 1129.53, found: 1129.53.

**HPLC**: retention time: 26.24 min (Nucleosil, gradient: 5-50 % B/A over 30 min).

Supplementary figure S53: HPLC trace of crude 6.

Supplementary figure S54: HPLC trace of pure **6**.9b- Influence of the nature of the cysteine C-terminusSupplementary scheme S17: Native chemical ligations with **7**, **8** or **9**.Ligation with **7**:

500  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.2 mg of the peptide **7** (final concentration 2 mM) and 0.7 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.

Ligation with **8**:

500  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.05 mg of the

peptide **8** (final concentration 2 mM) and 0.69 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.

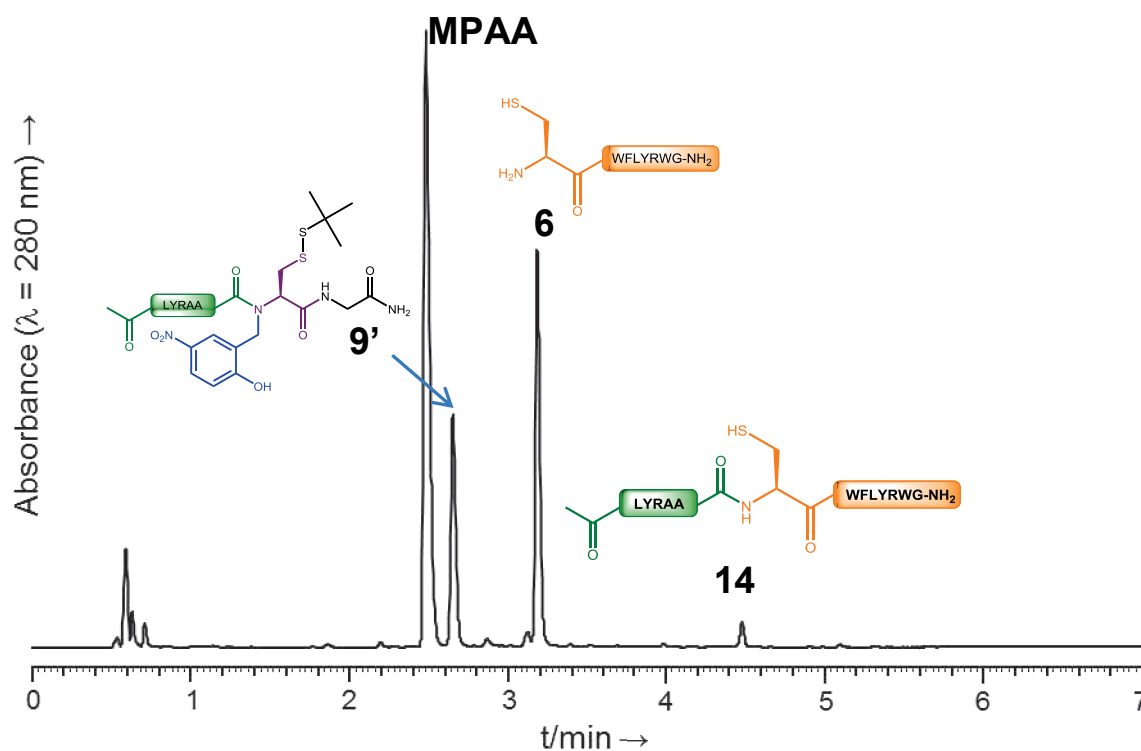
Ligation with **9**:

500  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.3 mg of the peptide **9** (final concentration 2 mM) and 0.7 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.

#### Ligation product **14**:

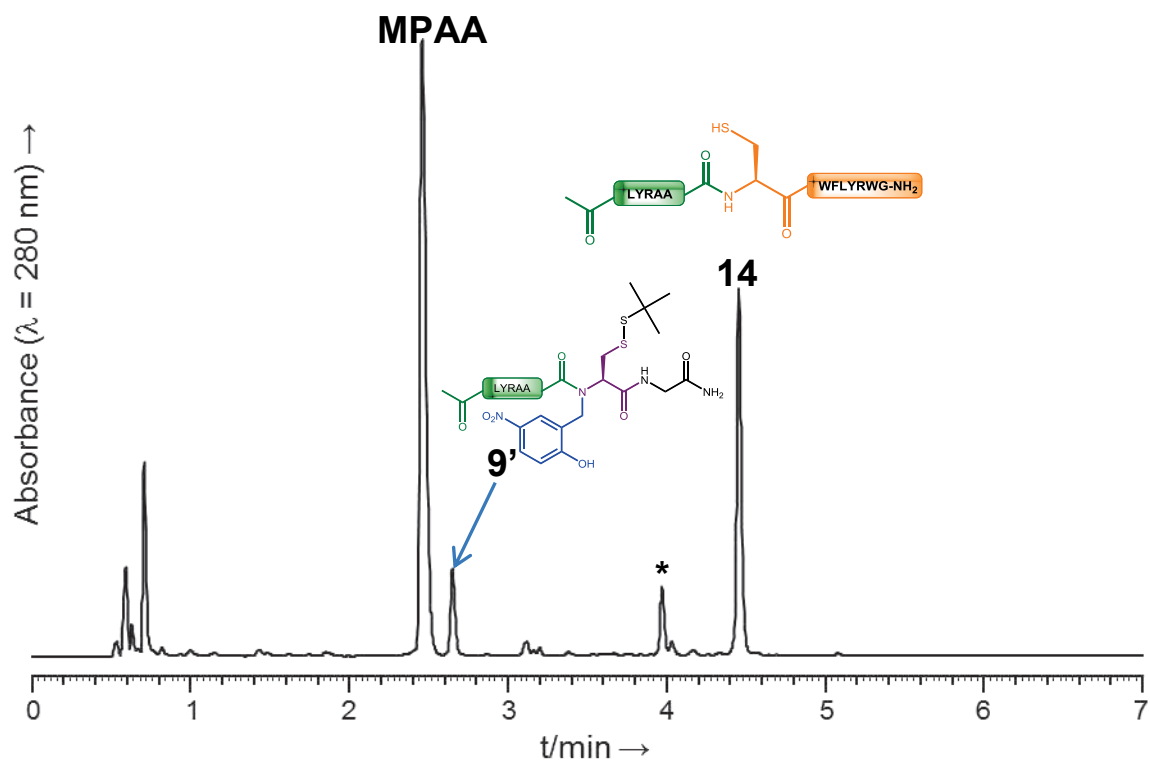
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{86}H_{117}N_{22}O_{16}S$ : 1745.8739, found: 1745.8773.

**HPLC**: retention time: 4.46 min (Chromolith, gradient: 20-70 % B/A over 6 min).

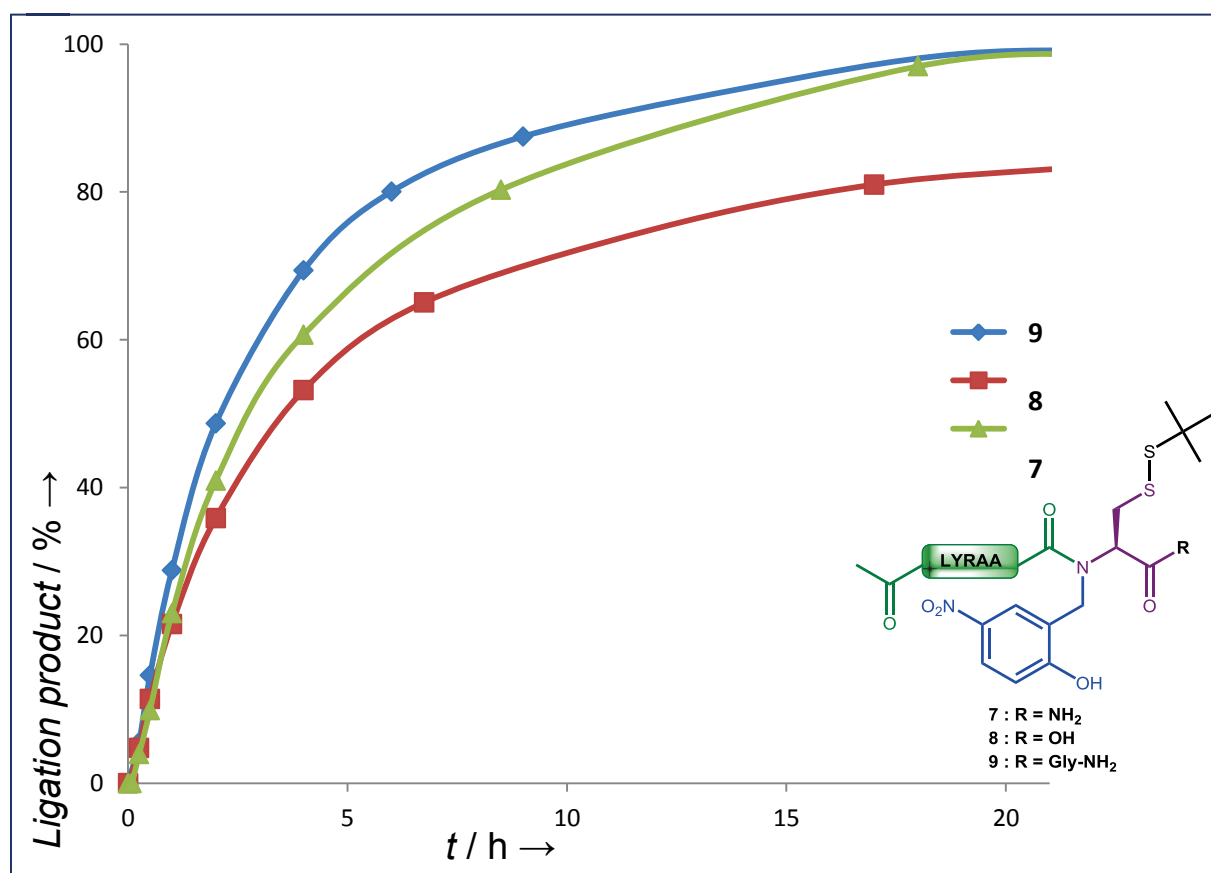


Supplementary figure S55: NCL at 15 min using **9**.





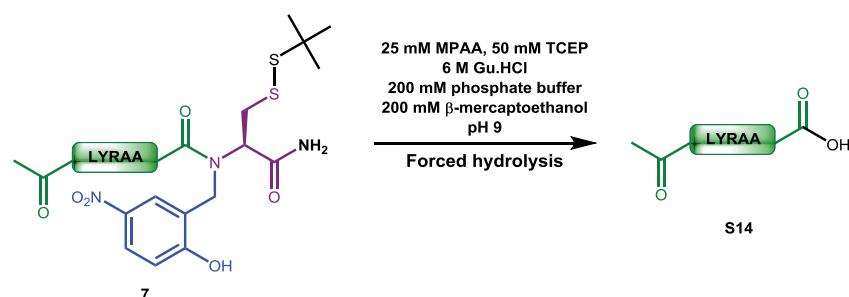
Supplementary figure S56: NCL after 24h using **9**. \*: non peptidic compound



Supplementary figure S57: Ligation kinetics observed using **7**, **8** or **9**.

9c- Study of the hydrolysis of the thioester precursor **9** and epimerization at the ligation site during a typical NCL with **9**

In order to quantify the hydrolysis of the thioester and the epimerization at the ligation site during a classical NCL with **6** and **9** (page S60), HPLC standards **S14** and **S15** were synthesized.



Supplementary scheme S18: Hydrolysis of peptide **7** to give HPLC standard **S14**.<sup>3</sup>

Reaction conditions inspired by Kent *et al.*<sup>4</sup>

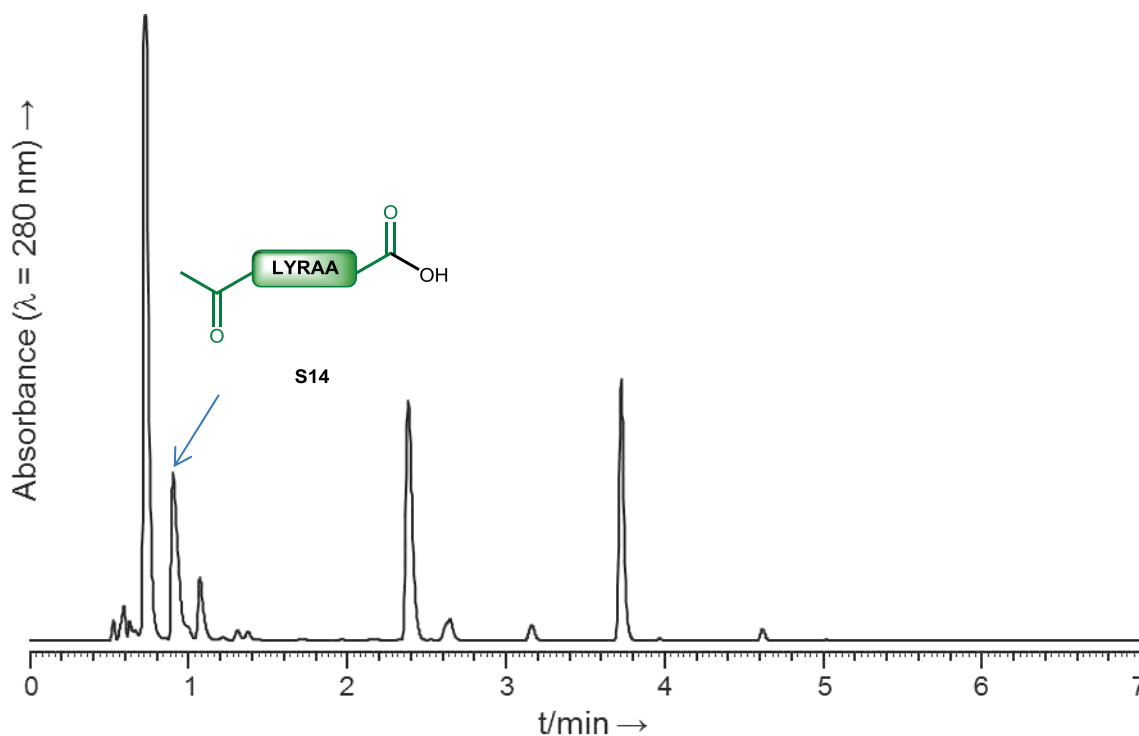
Peptide **7** (1 mM) was incubated under argon in 100  $\mu$ L of a deoxygenated 200 mM pH 9 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP, 200 mM  $\beta$ -mercaptoethanol and 6 M guanidine hydrochloride. The reaction was monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). After 16h at room temperature, the starting material was consumed and the product **S14** was characterized and purified to serve as a HPLC standard.

**S14:**

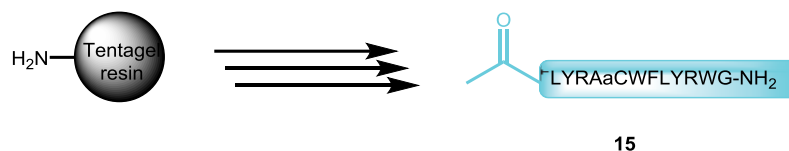
**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{29}H_{47}N_8O_8$ : 635.35, found: 635.28.

**HPLC**: retention time: 1.04 min (Chromolith, gradient: 20-70 % B/A over 6 min).

<sup>3</sup> Gates Z. P., Stephan J. R., Lee D. J., Kent S. B. H., *Chem. Comm.*, 2013, **49**, 786-788



Supplementary figure S58: HPLC trace of crude **S14**.



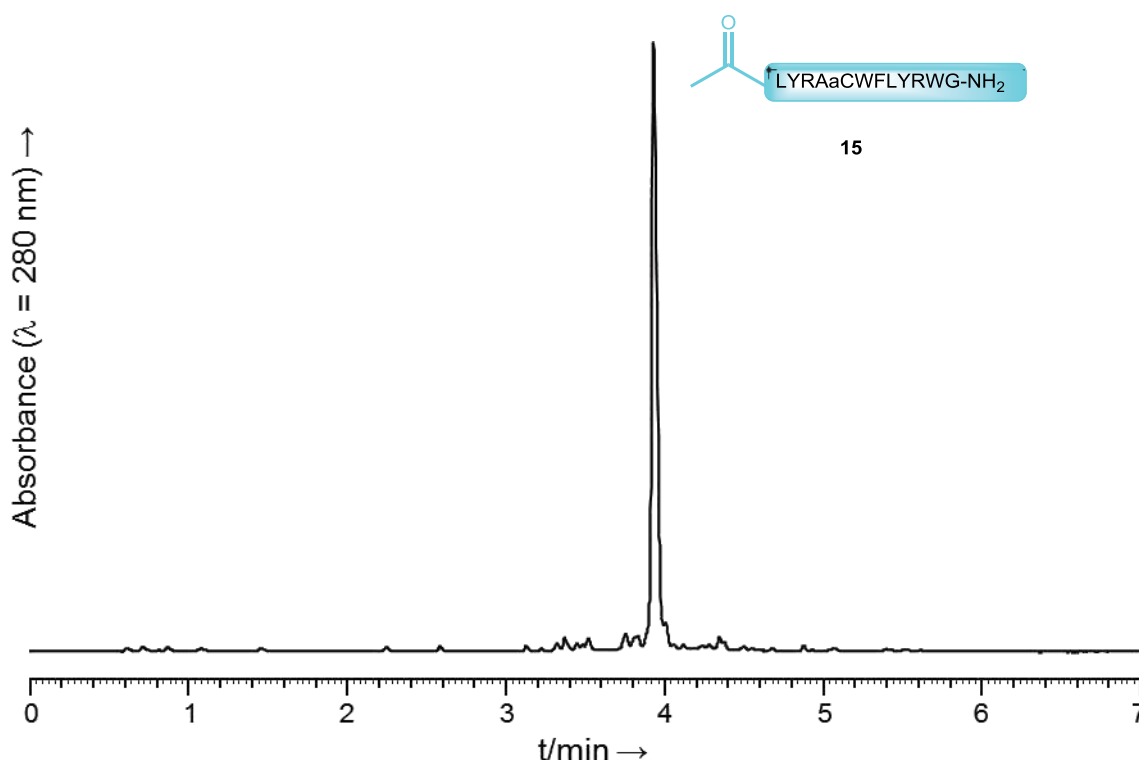
Supplementary scheme S19: Synthesis of HPLC standard **15**.

Peptide **15** was obtained through automated SPPS (protocol p S3) starting from Tentagel resin (476 mg, 0.21 mmol/g, 0.1 mmol). The peptide-resin was cleaved (protocol p S3) in order to characterize the corresponding peptide **15**.

**15:**

**MALDI-TOF** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{86}\text{H}_{117}\text{N}_{22}\text{O}_{16}\text{S}$ : 1745.87, found: 1745.84.

**HPLC**: retention time: 3.94 min (Chromolith, gradient: 20-70 % B/A over 6 min).



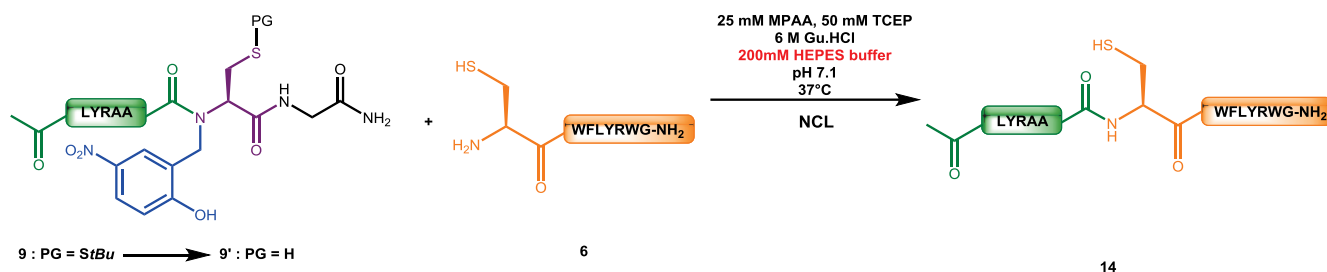
Supplementary figure S59: HPLC trace of crude **15**.

An HPLC co-injection of the NCL mixture after 24h using **9** with HPLC standards **S14** and **15** allowed determining the amount of hydrolysis of the thioester moiety and epimerization at the ligation site after a day of reaction.

<b>15</b> formed (%)	< 0.4
<b>S14</b> formed (%)	< 4

Supplementary table S4: Hydrolysis and epimerization during a typical NCL with **9**.

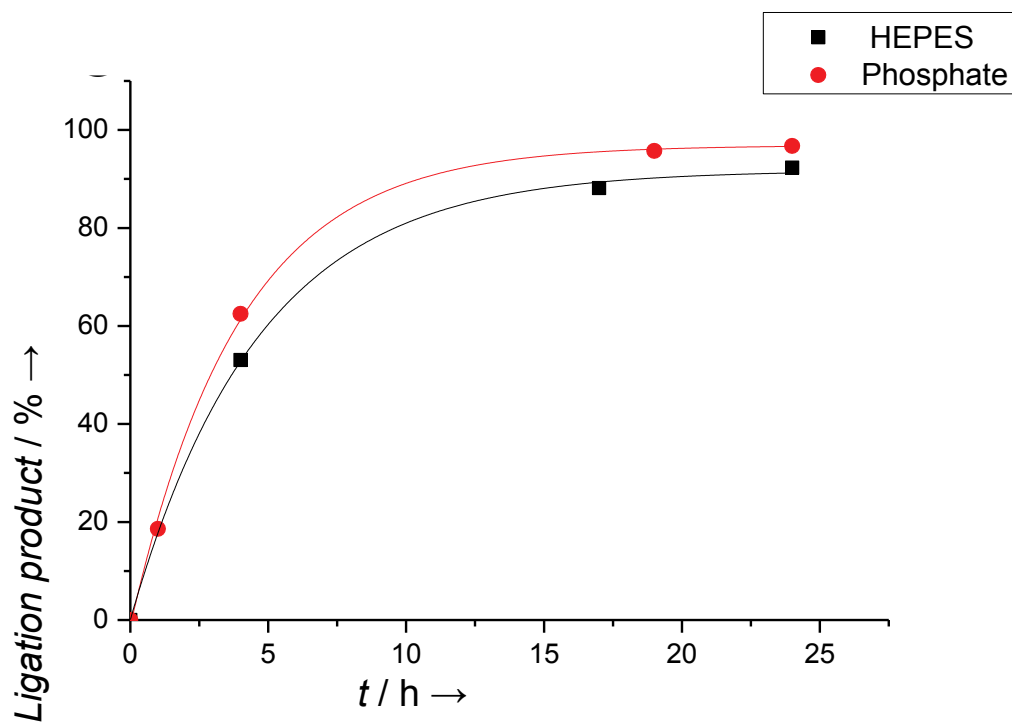
#### 9d- Influence of the buffer



Supplementary scheme S20: NCL in HEPES buffer with **9**.

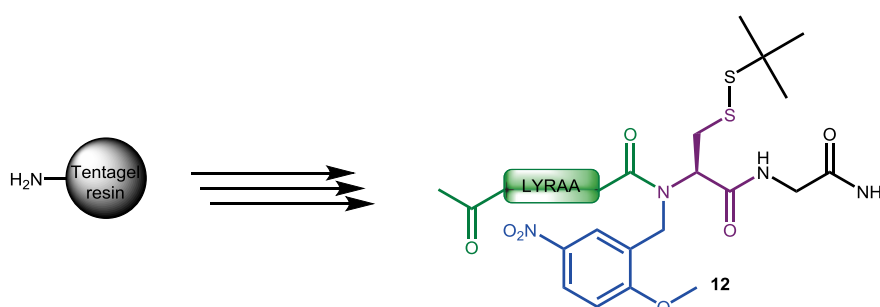
500  $\mu\text{L}$  of a degassed 0.2 M pH 7.1 HEPES buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 0.8 mg peptide **9** (final concentration 1.5 mM) and 0.7 mg peptide **6** (final concentration 1 mM) under argon.

The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.



Supplementary figure S60: Comparison of reaction kinetics using phosphate or HEPES buffer using **9**.

### 9e- Influence of masking phenol or thiol groups



Supplementary scheme S21: Synthesis of O-methylated peptide **12**.

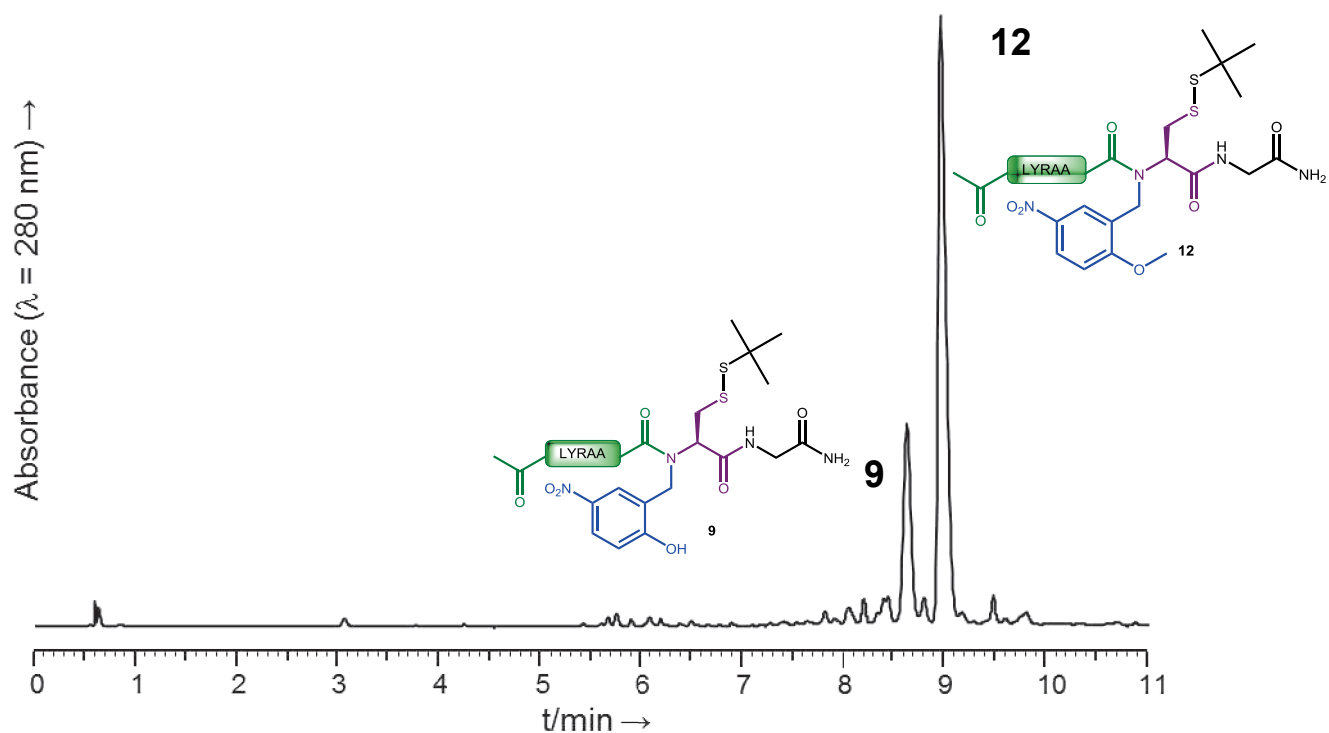
Tentagel R resin (476 mg, 0.21 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Rink-OH (270 mg, 0.5 mmol, 5 equiv.) and HATU (190 mg, 0.5 mmol, 5 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175  $\mu$ L, 1 mmol, 10 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring,

then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Gly-OH (297 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348 μL, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Then, Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Cys(*S*tBu)-OH (432 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (10 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348 μL, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). The Ac-LYRAA sequence was installed by standard SPPS procedure (protocol p S3), an extended double coupling (2 × 2h) was performed for the first alanine and a normal double coupling (2 × 30 min) was performed for the arginine. Peptide resin was then treated with 20% piperidine in NMP (10 mL, 3 × 3 min), washed with NMP then CH<sub>2</sub>Cl<sub>2</sub>. Methylation of the phenol group was performed using a solution of MeI (50 equiv.), *i*Pr<sub>2</sub>NEt (50 equiv.) in DMF for 4 h and the resin was washed with NMP and CH<sub>2</sub>Cl<sub>2</sub>. Finally, peptide-resin was cleaved following the general procedure (p S3) to give compound **12**, which was purified by semi-preparative HPLC (Nucleosil, gradient: 40-45 % B/A over 10 min).

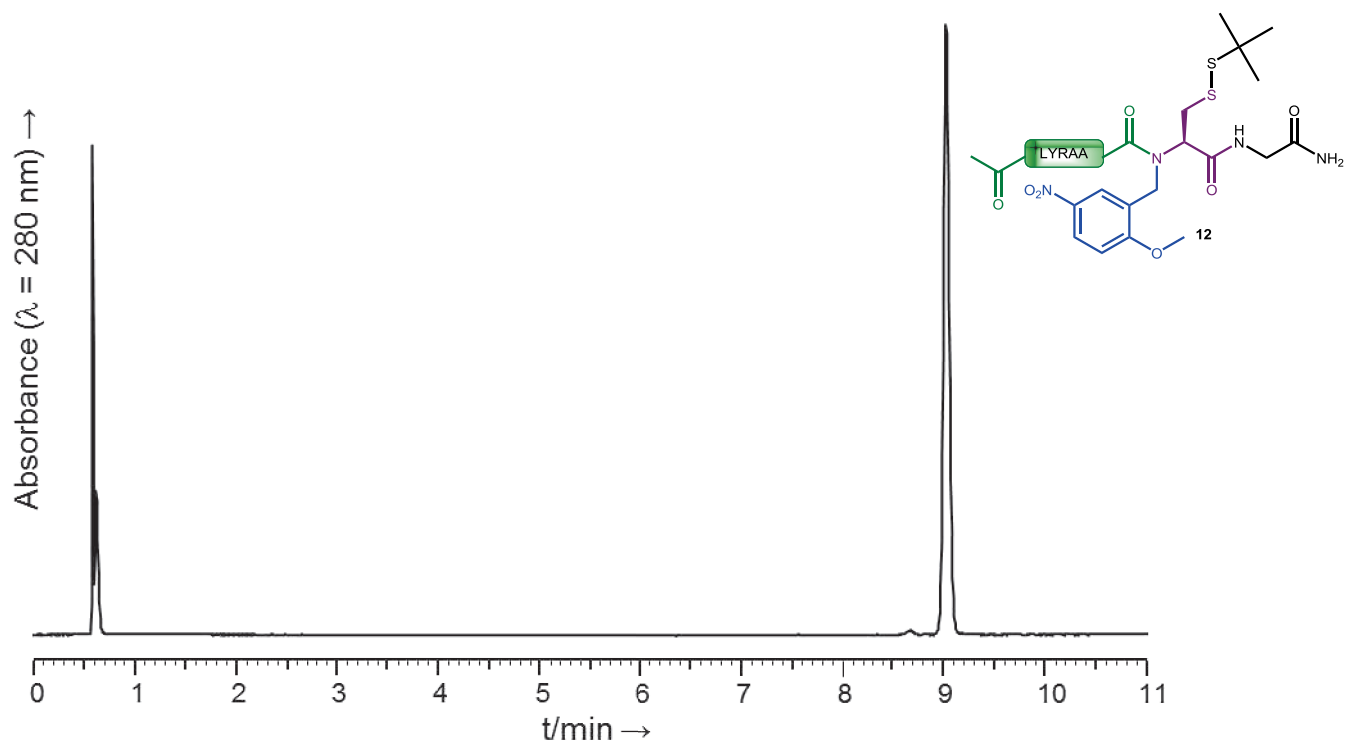
**12:**

**MALDI-TOF** (*m/z*): [MH]<sup>+</sup> calcd for C<sub>46</sub>H<sub>71</sub>N<sub>12</sub>O<sub>12</sub>S<sub>2</sub>: 1047.48, found: 1047.46.

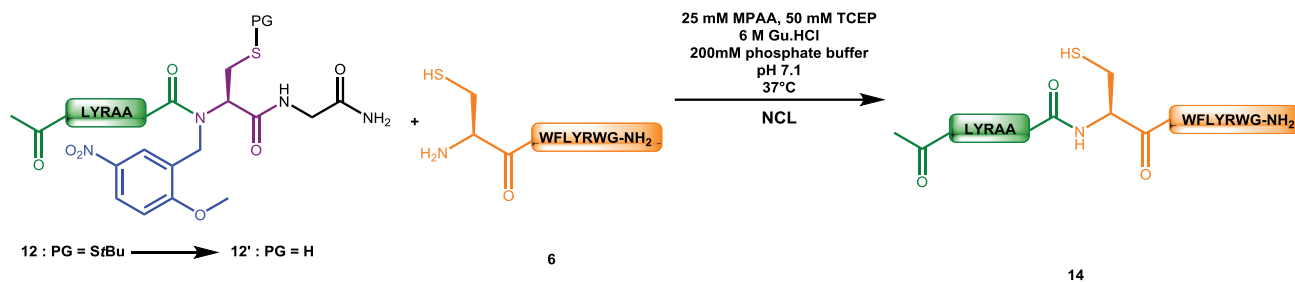
**HPLC**: retention time: 8.94 min (Chromolith, gradient: 5-50 % B/A over 10 min).



Supplementary figure S61: HPLC trace of crude **12**.



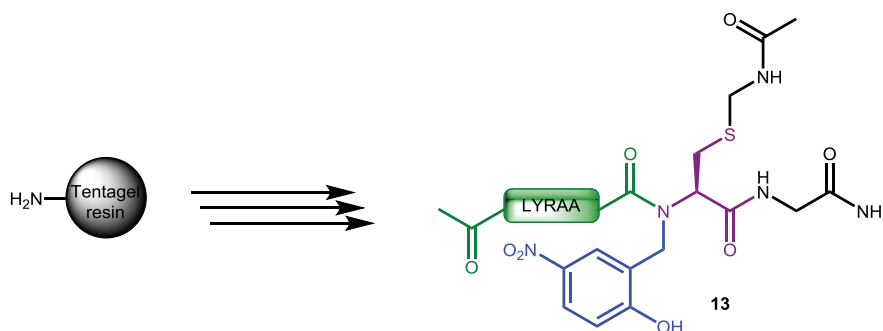
Supplementary figure S62: HPLC trace of pure **12**.



Supplementary scheme S22: Native chemical ligation with O-methylated 12.

### Ligation with 12:

500  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.05 mg of the peptide **12** (final concentration 2 mM) and 0.69 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.



Supplementary scheme S23: Synthesis of the S-Acm peptide 13.

Tentagel R resin (476 mg, 0.21 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Rink-OH (270 mg, 0.5 mmol, 5 equiv.) and HATU (190 mg, 0.5 mmol, 5 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175  $\mu$ L, 1 mmol, 10 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Gly-OH (297 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Then, Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Cys(Acm)-OH (414.5 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (10 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution

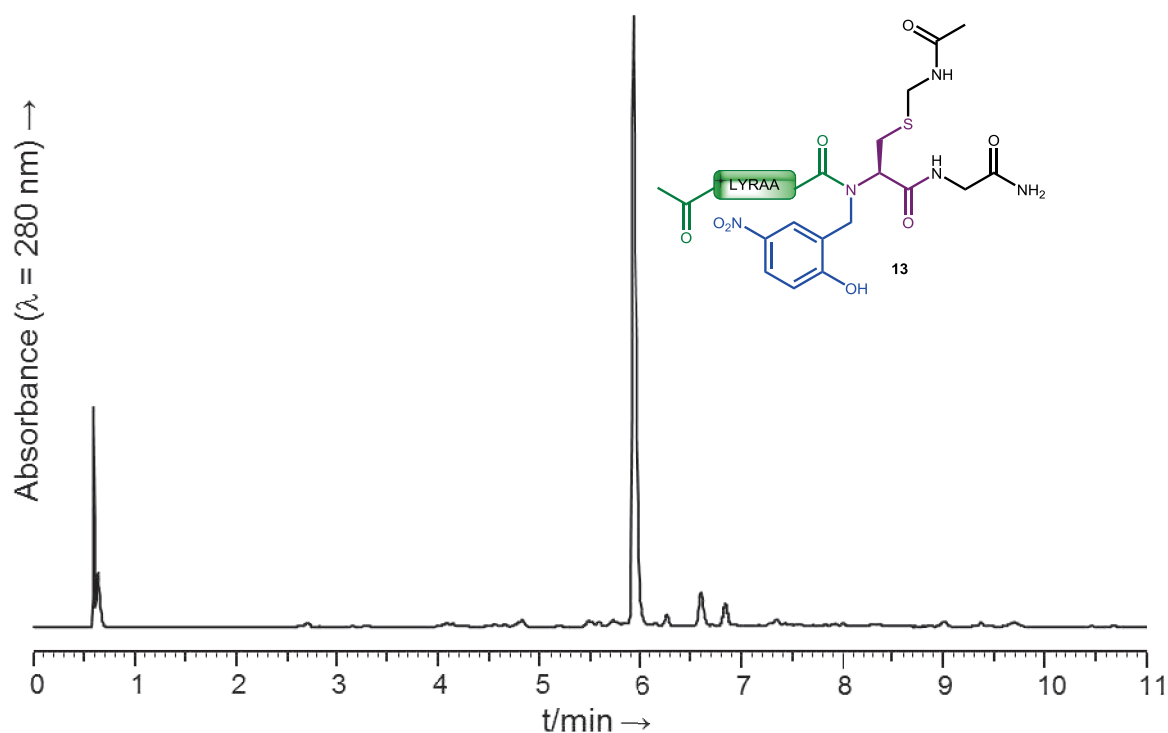


was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The 2-hydroxy-5-nitrobenzyl linker was then introduced following the general procedure (p S5). The Ac-LYRAA sequence was installed by standard SPPS procedure (protocol p S3), an extended double coupling ( $2 \times 2$  h) was performed for the first alanine and a normal double coupling ( $2 \times 30$  min) was performed for the arginine. Peptide resin was then treated with 20% piperidine in NMP (10 mL,  $3 \times 3$  min), washed with NMP then  $\text{CH}_2\text{Cl}_2$ . Finally, the peptide was cleaved following the general procedure p S3 to give compound **13**. **13** has been purified by semi-preparative HPLC (Nucleosil, gradient: 20-35 % B/A over 5 min).

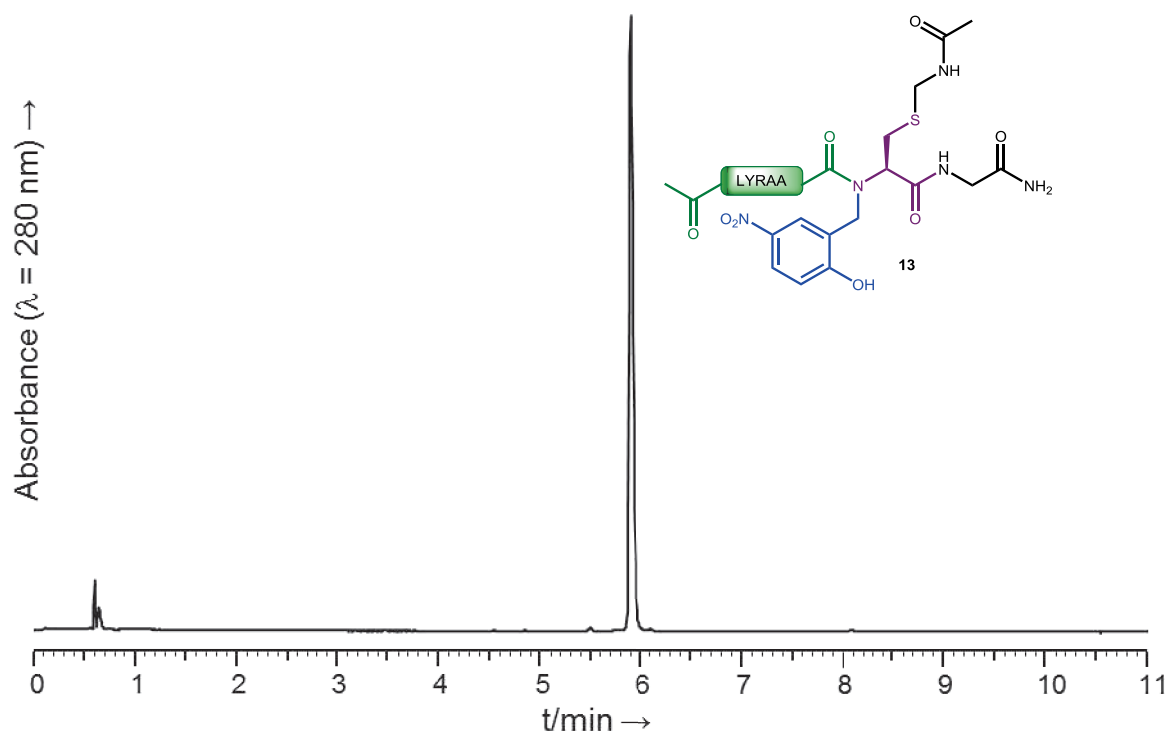
**13:**

**MALDI-TOF** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{44}\text{H}_{66}\text{N}_{13}\text{O}_{13}\text{S}$ : 1016.46, found: 1016.46.

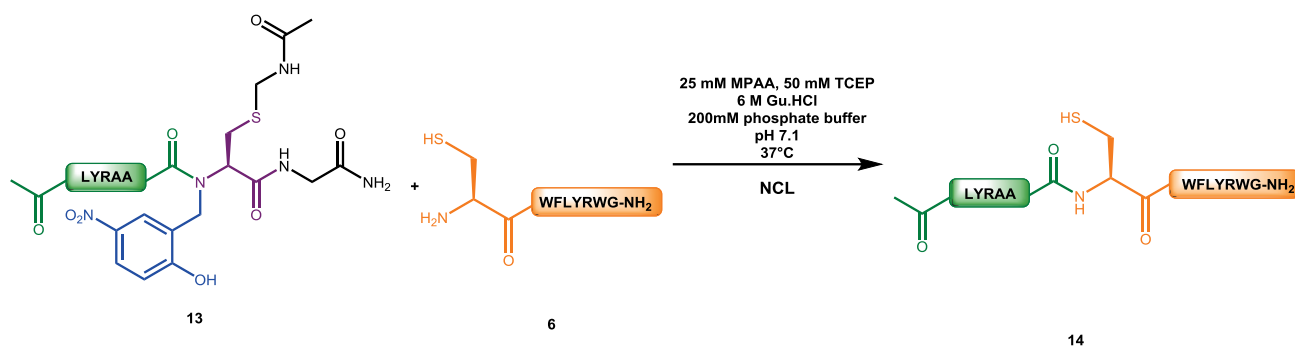
**HPLC**: retention time: 5.88 min (Chromolith, gradient: 5-50 % B/A over 10 min).



Supplementary figure S63: HPLC trace of crude **13**.



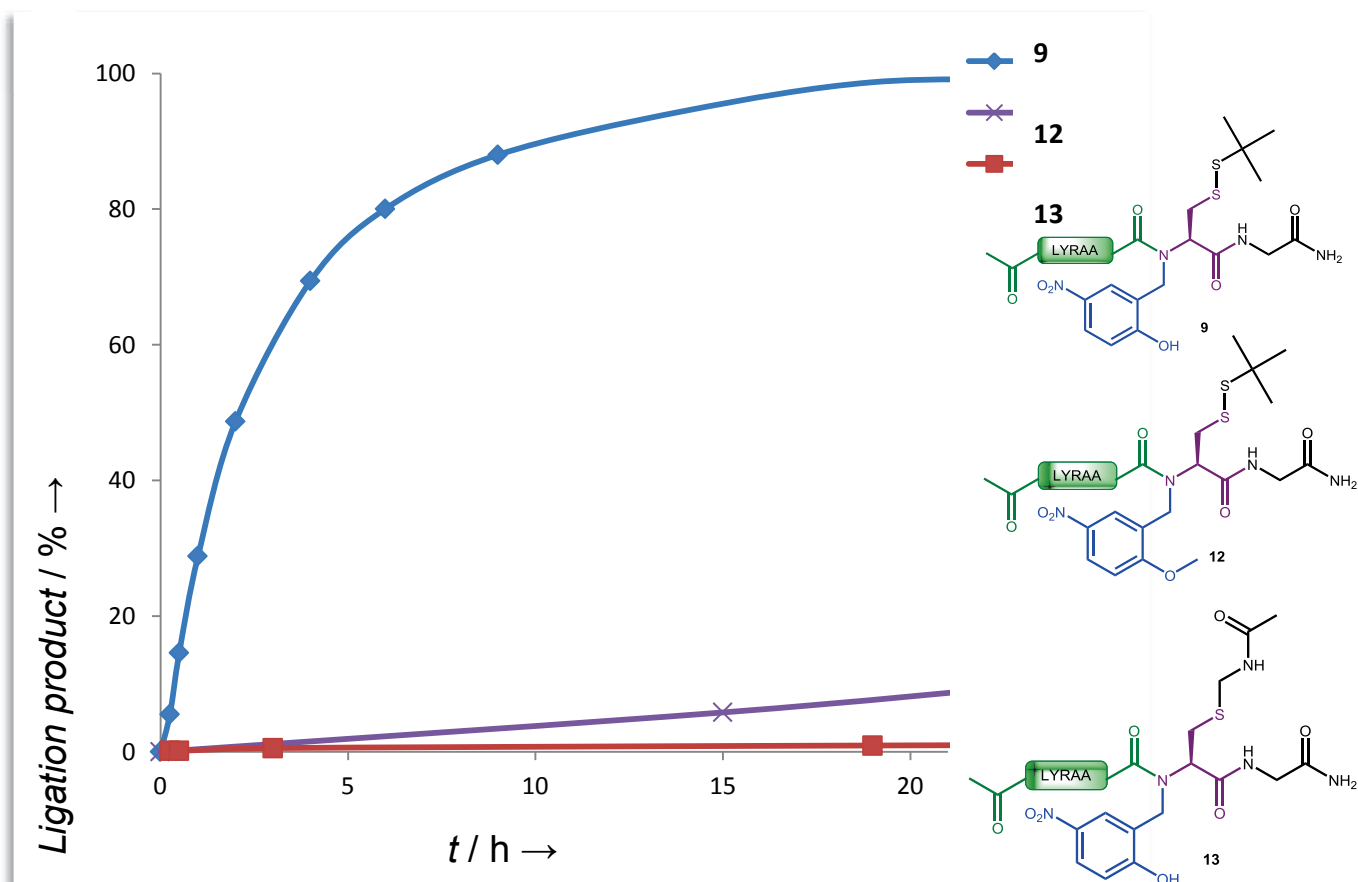
Supplementary figure S64: HPLC trace of pure **13**.



Supplementary scheme S24: Native chemical ligation with S-Acm peptide **13**.

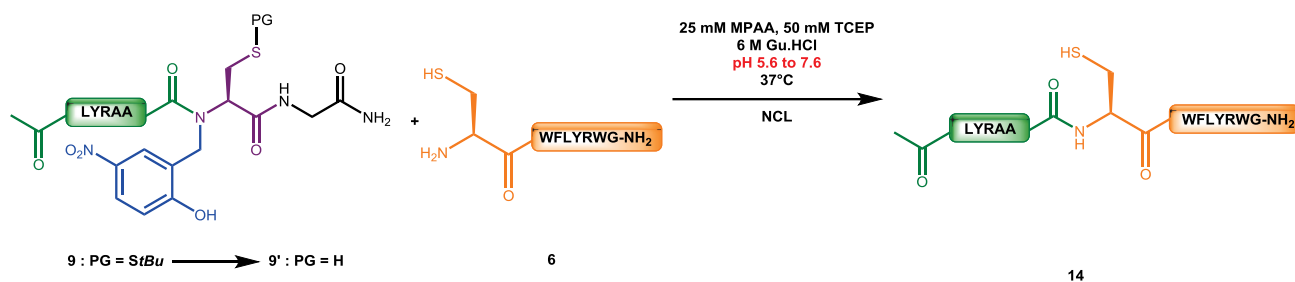
#### Ligation with **13**:

500  $\mu\text{L}$  of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.2 mg of the peptide **13** (final concentration 2 mM) and 0.69 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu\text{L}$  were diluted in 108  $\mu\text{L}$  of 1.5 % TFA in water and 100  $\mu\text{L}$  were injected in HPLC.



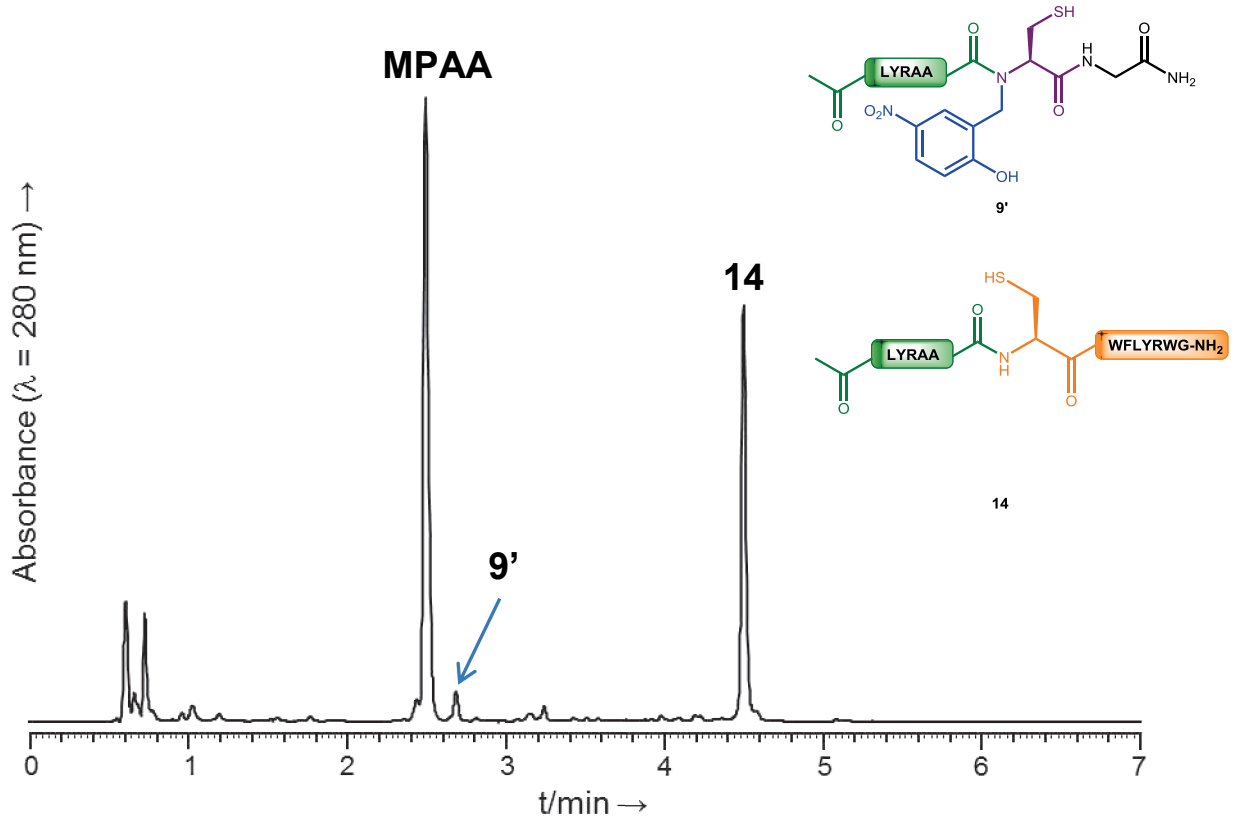
Supplementary figure S65: Ligation kinetics observed using **12** or **13** compared to ligation using **9**.

### 9f- Influence of the pH

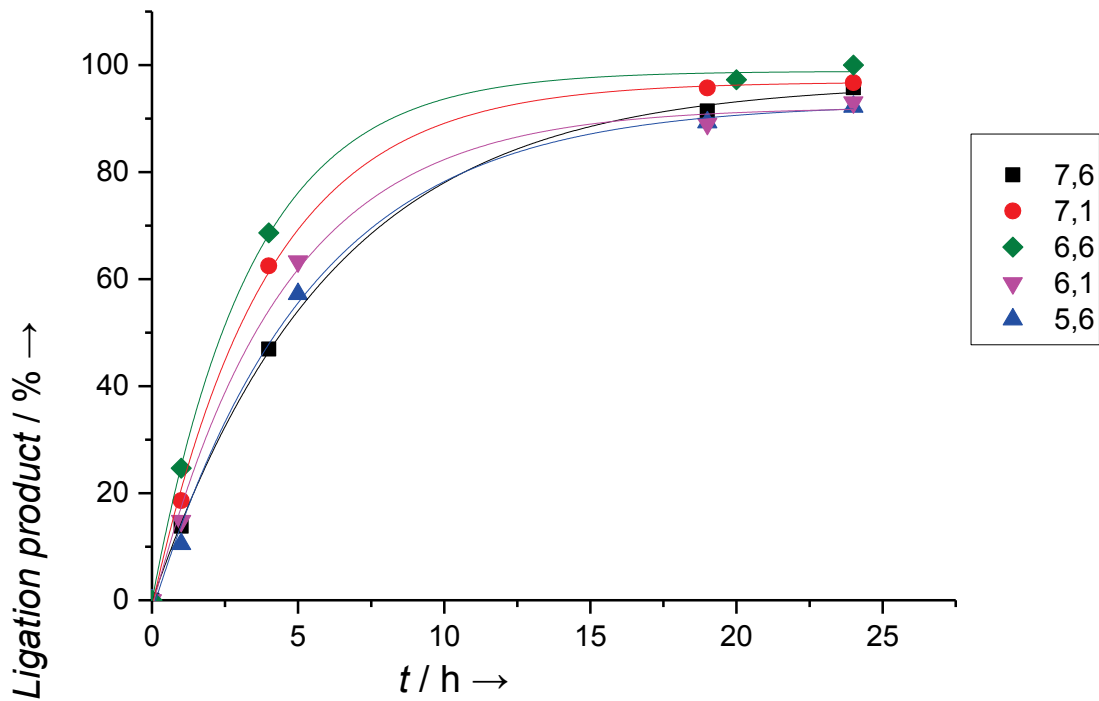


Supplementary scheme S25: NCL at different pHs with **9**.

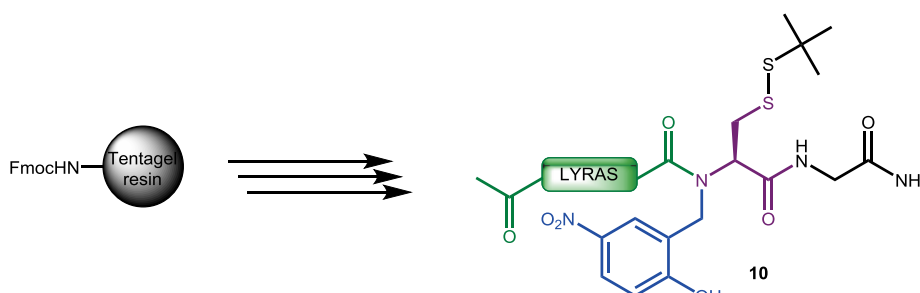
500  $\mu\text{L}$  of different degassed 0.2 M sodium phosphate buffers (pH 5.6, 6.1, 6.6, 7.1, 7.6) containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 0.8 mg of the peptide **9** (final concentration 1.5 mM) and 0.69 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu\text{L}$  were diluted in 108  $\mu\text{L}$  of 1.5 % TFA in water and 100  $\mu\text{L}$  were injected in HPLC.



Supplementary figure S66: NCL after 24h using **9** at pH 6.6.



Supplementary figure S67: Ligation kinetics observed at different pHs using **9**.

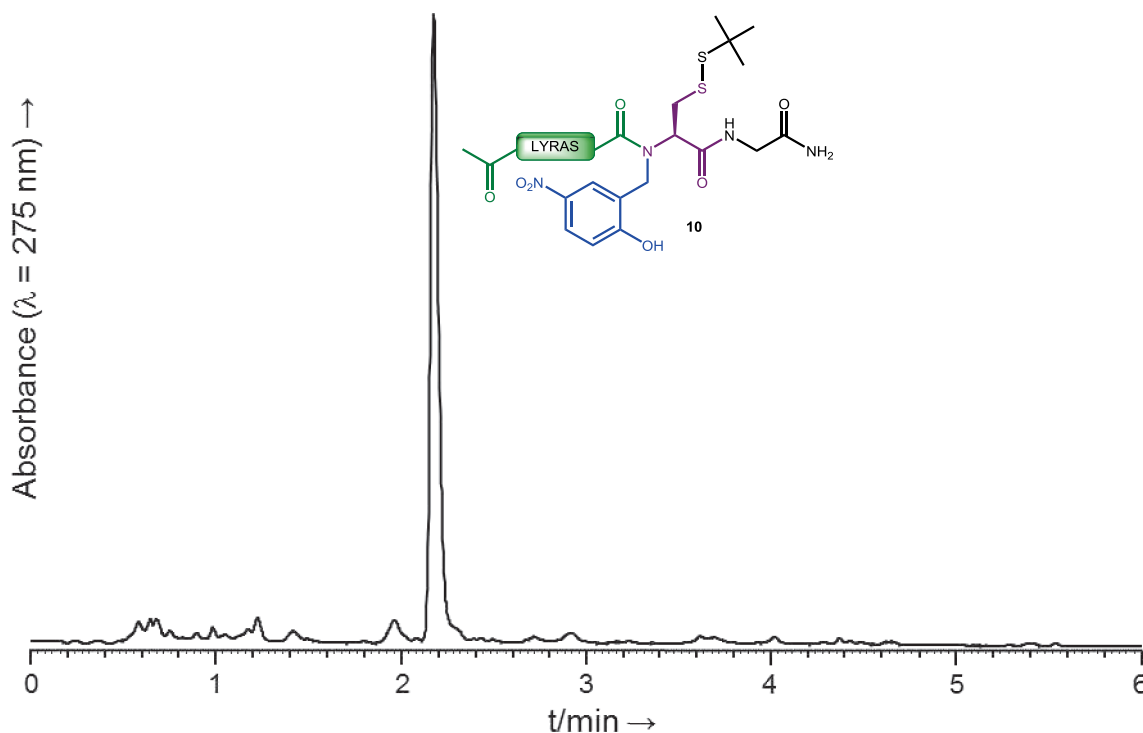
**10-Influence of the C-terminal aa: NCL with model Ac-LYRAX(Hnb)C peptides****10a- Synthesis of peptide crypto-thioester Ac-LYRAS-(Hnb)C(*t*Bu)G-NH<sub>2</sub> (10)****Supplementary scheme S26: Synthesis of model peptide 10.**

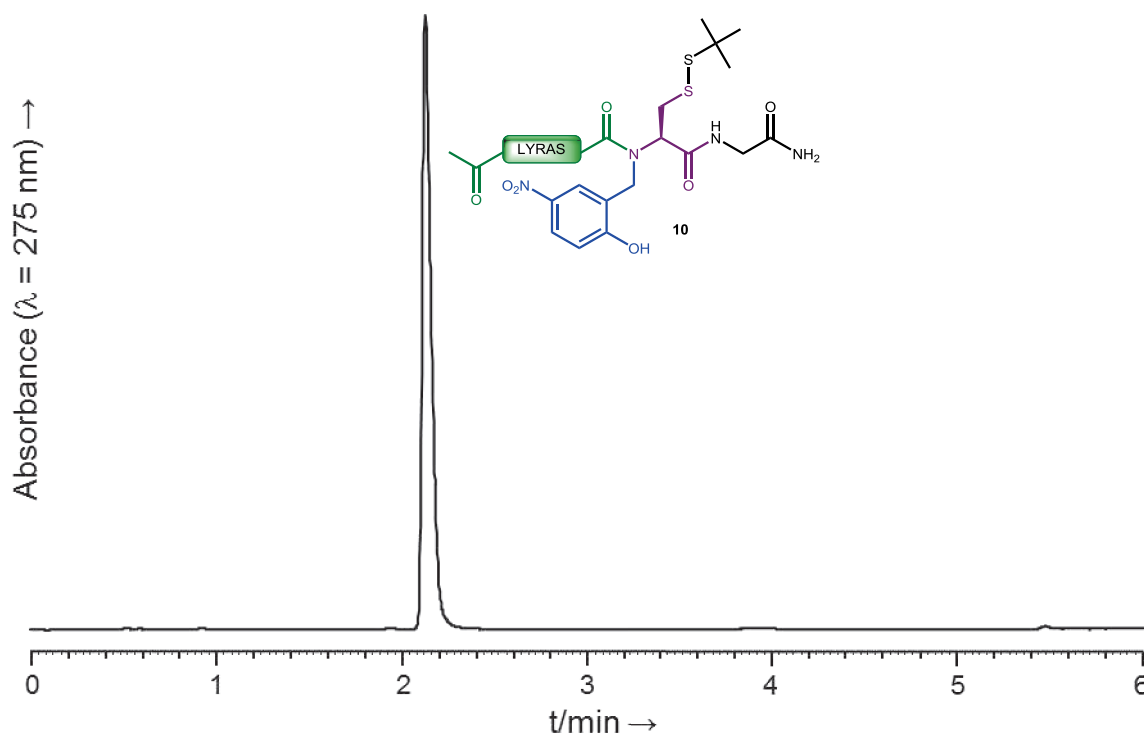
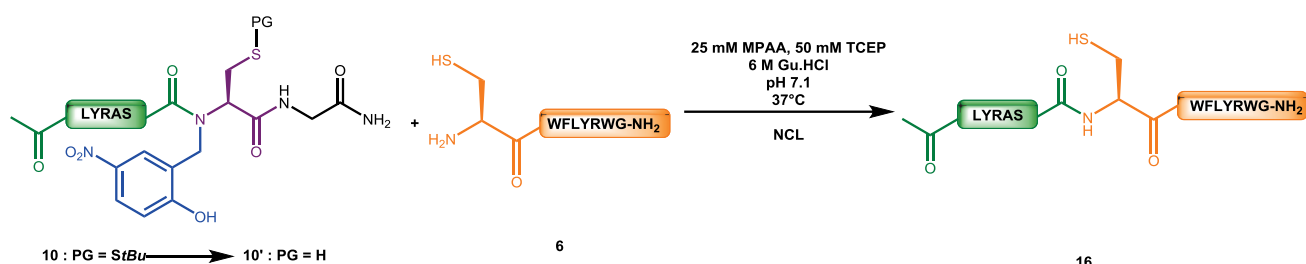
Peptide **10** was obtained through automated SPPS (protocol p S3) and using the automated reductive amination (protocol p S34) starting from Fmoc-Rink tentagel resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol). Fmoc-Ser(*t*Bu)-OH was coupled for 3  $\times$  30 min on the (Hnb)C(*t*Bu)G-Rink Tentagel resin. The complete sequence was installed through standard Fmoc-SPPS (protocol p S3), Fmoc-Arg(Pbf)-OH was coupled twice (2  $\times$  30 min) and a final piperidine treatment (20 % in NMP, 3 mL, 3 min,  $\times$ 3) was performed. Peptide-resin was cleaved (protocol p S3) and peptide **10** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 40-55 % B/A over 12 min).

**10:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for C<sub>45</sub>H<sub>69</sub>N<sub>12</sub>O<sub>13</sub>S<sub>2</sub>: 1049.4548, found: 1049.4532.

**HPLC**: retention time: 2.14 min (Chromolith, gradient: 30-60 % B/A over 5 min).

**Supplementary figure S68: HPLC trace of crude 10.**

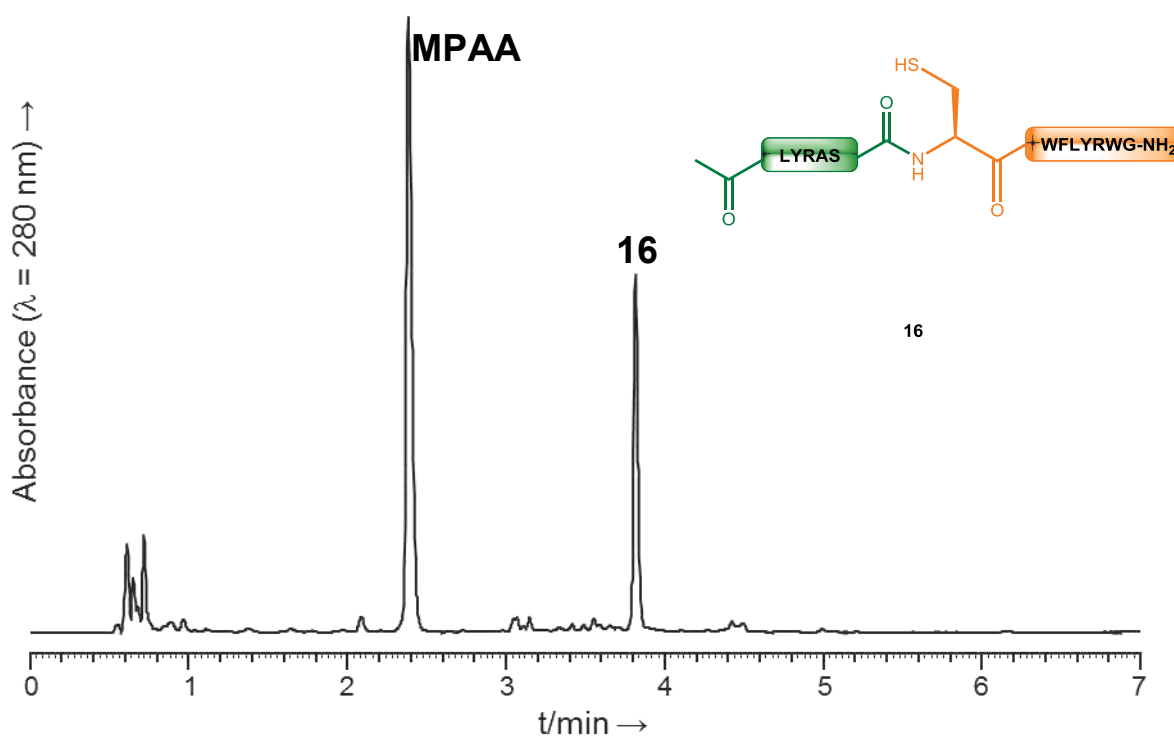
Supplementary figure S69: HPLC trace of pure **10**.10b- NCL with peptide crypto-thioester Ac-LYRAS-(Hnb)C(*StBu*)G-NH<sub>2</sub> (**10**)Supplementary scheme S27: Native chemical ligation with **10**.

250  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 0.53 mg of the peptide **10** (final concentration 2 mM) and 0.30 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.

**Ligation product 16:**

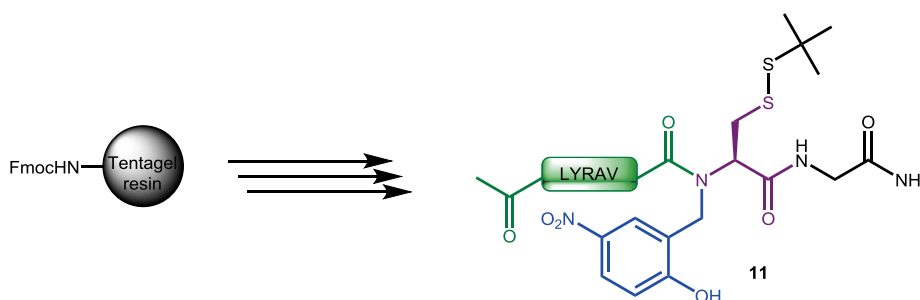
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{86}H_{117}N_{22}O_{17}S$ : 1761.8688, found: 1761.8687.

**HPLC**: retention time: 3.81 min (Chromolith, gradient: 20-70 % B/A over 6 min).



Supplementary figure S70: NCL after 25h using 10.

10c-Synthesis of peptide crypto-thioester Ac-LYRAV-(Hnb)C(S*t*Bu)G-NH<sub>2</sub> (11)



Supplementary scheme S28: Synthesis of model peptide 11.

Peptide **11** was obtained through automated SPPS (protocol p S3) and using the automated reductive amination (protocol p S34) starting from Fmoc-Rink tentagel resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol). Fmoc-Val-OH has been coupled for 5  $\times$  30 min on the (Hnb)C(S*t*Bu)G-Rink Tentagel resin. The complete sequence was installed through standard Fmoc-SPPS (protocol p S3), Fmoc-Arg(Pbf)-OH was coupled for 2  $\times$  30 min and a final piperidine (20 % in NMP, 3mL, 3 min,  $\times$ 3) was

performed. Peptide-resin was cleaved (protocol p S3) and peptide **11** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 40-55 % B/A over 12 min).

**11:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{47}H_{73}N_{12}O_{12}S_2$ : 1061.4912, found: 1061.4891.

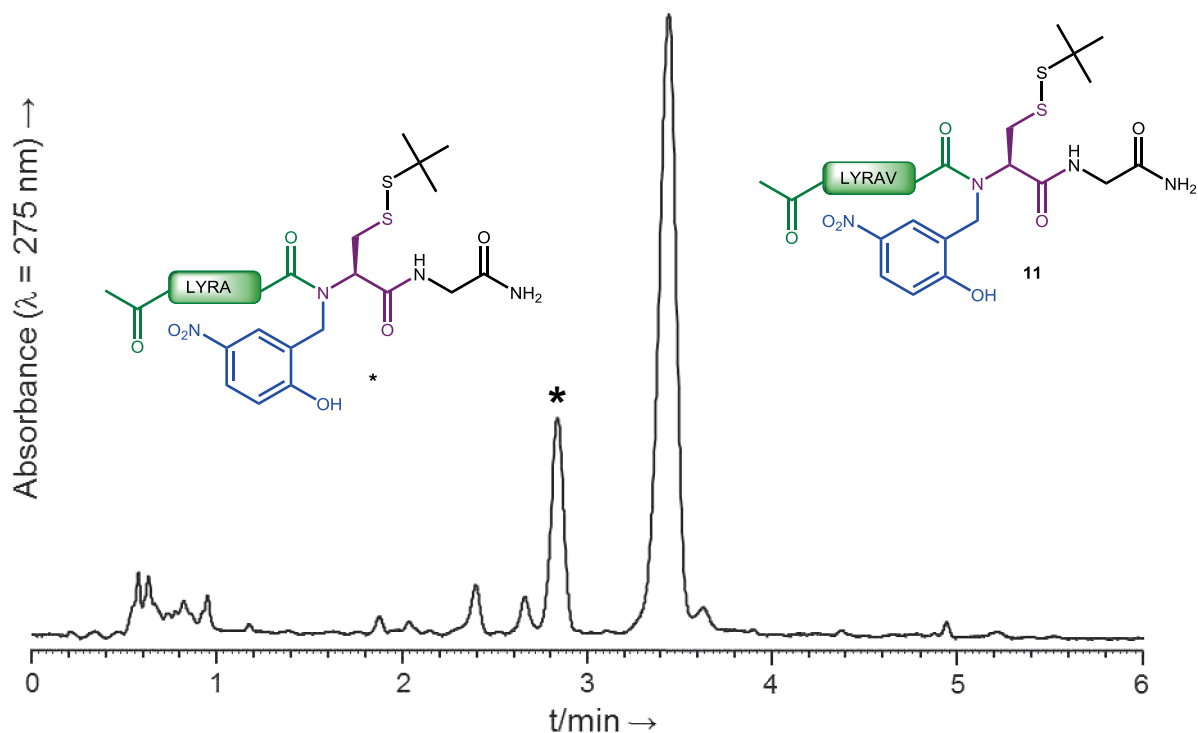
**HPLC**: retention time: 3.51 min (Chromolith, gradient: 30-60 % B/A over 5 min).

\*: minor compound observed (Val deletion) resulting from incomplete coupling of Fmoc-Val-OH because no acetylation-mediated capping step was performed.

**\*:**

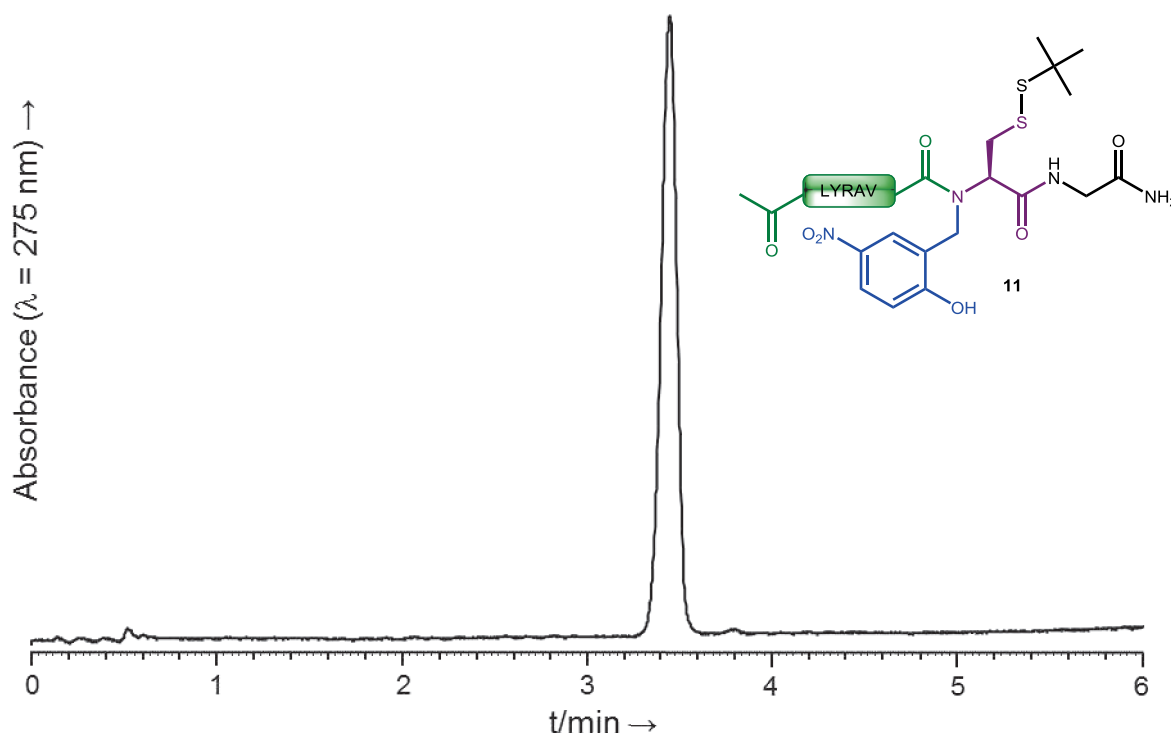
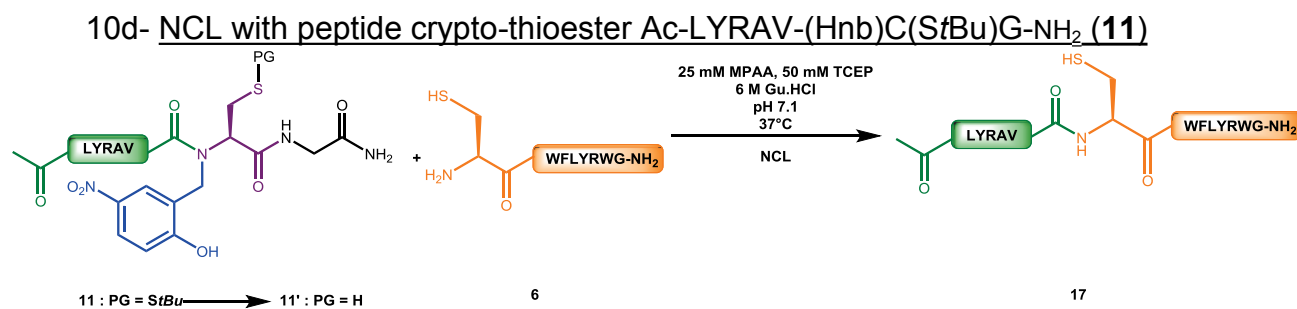
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{42}H_{64}N_{11}O_{11}S_2$ : 962.4228, found: 962.4224.

**HPLC**: retention time: 3.51 min (Chromolith, gradient: 30-60 % B/A over 5 min).



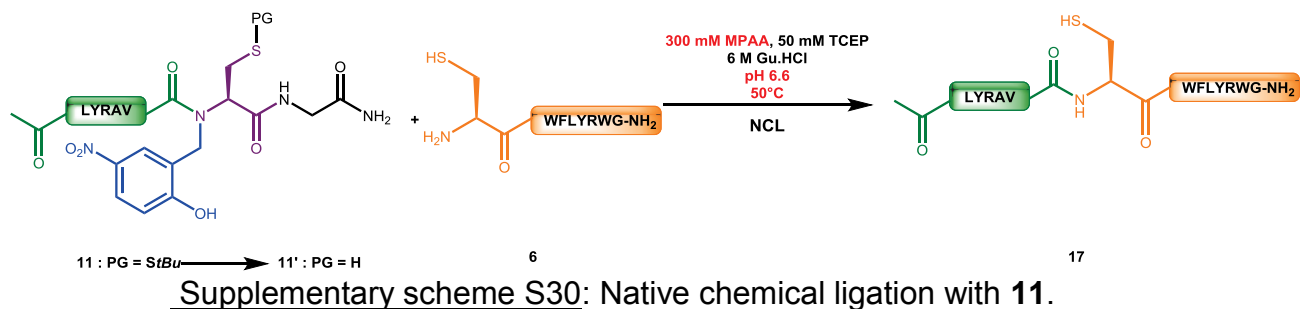
Supplementary figure S71: HPLC trace of crude **11**.



Supplementary figure S72: HPLC trace of pure **11**.Supplementary scheme S29: Native chemical ligation with **11**.

250  $\mu$ L of a degassed 0.2 M pH 7.1 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 0.53 mg of the peptide **11** (final concentration 2 mM) and 0.30 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu$ L were diluted in 108  $\mu$ L of 1.5 % TFA in water and 100  $\mu$ L were injected in HPLC.

10e- NCL with peptide crypto-thioester Ac-LYRAV-(Hnb)C(StBu)G-NH<sub>2</sub> (**11**) under optimized conditions

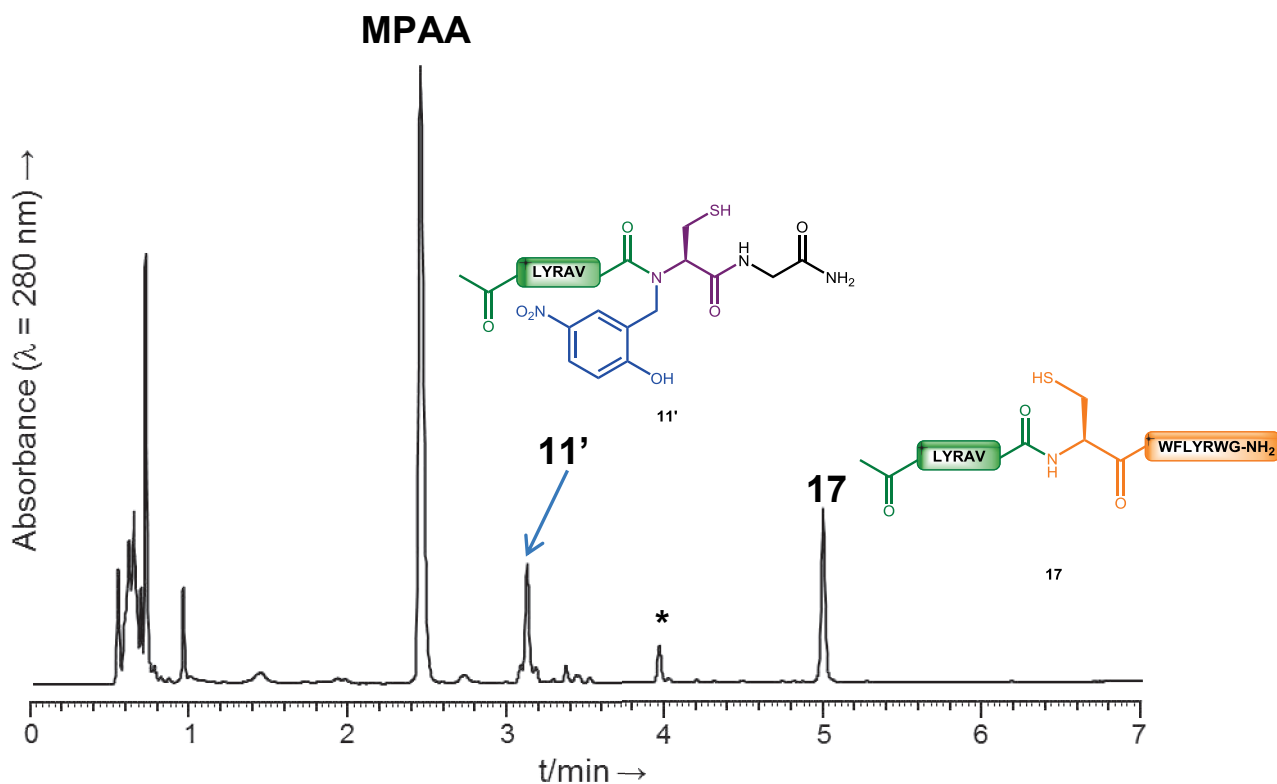


500  $\mu\text{L}$  of a degassed 0.2 M pH 6.6 sodium phosphate buffer containing 300 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.2 mg of the peptide **11** (final concentration 2 mM) and 0.68 mg peptide **6** (final concentration 1 mM) under argon. The ligation was carried out at 50°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu\text{L}$  were diluted in 108  $\mu\text{L}$  of 1.5 % TFA in water, this aqueous phase was washed with di-ethyl ether (500  $\mu\text{L}$   $\times$ 3) and 100  $\mu\text{L}$  were injected in HPLC.

**Ligation product **17**:**

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{88}\text{H}_{121}\text{N}_{22}\text{O}_{16}\text{S}$ : 1773.9052, found: 1773.9059.

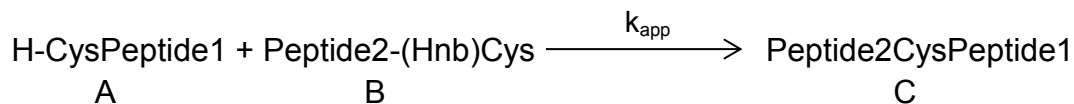
**HPLC**: retention time: 5.0 min (Chromolith, gradient: 20-70 % B/A over 6 min).



Supplementary figure S73: NCL after 24h using **11** under optimized conditions. \*: non peptidic compound

10f- Determination of apparent second order kinetic constants

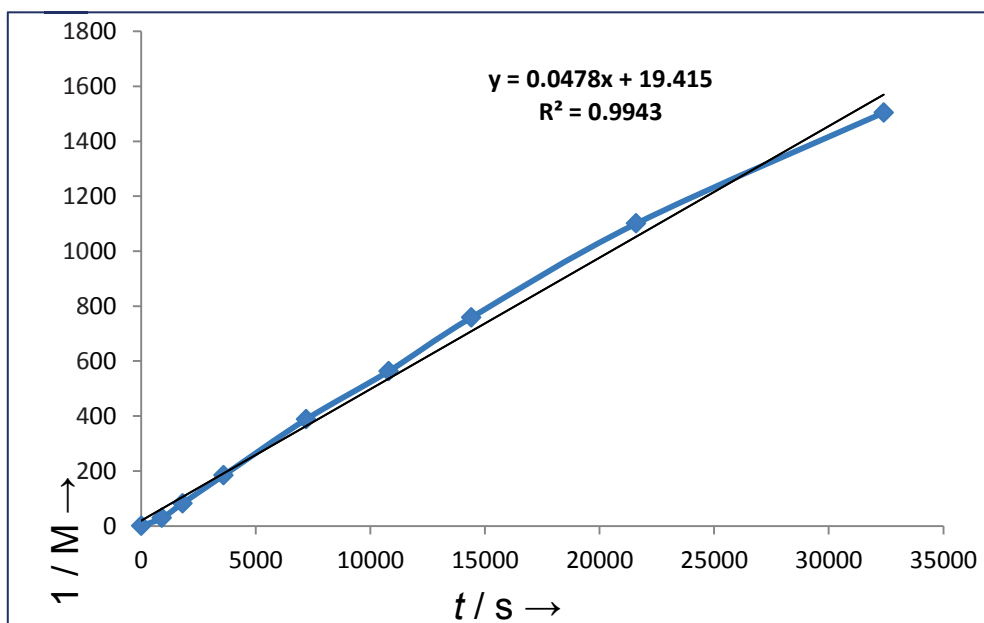
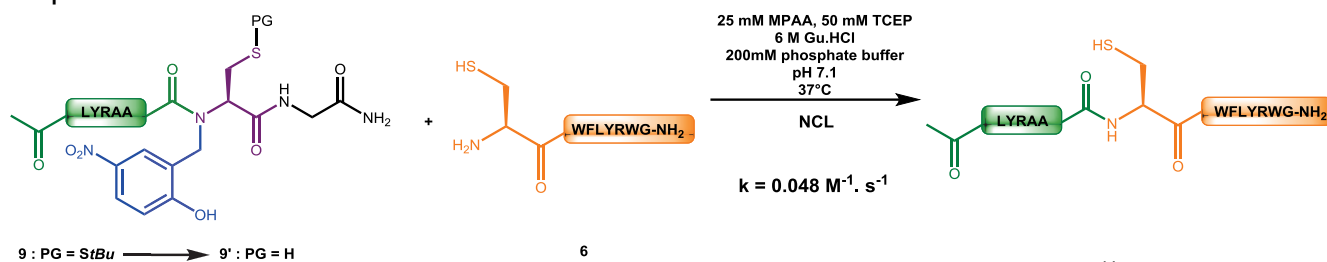
Kinetics of selected representative NCL reactions were modeled as an apparent second order reaction, considering the chemical equation:



following the general rate equation for  $[A]_0 \neq [B]_0$ :

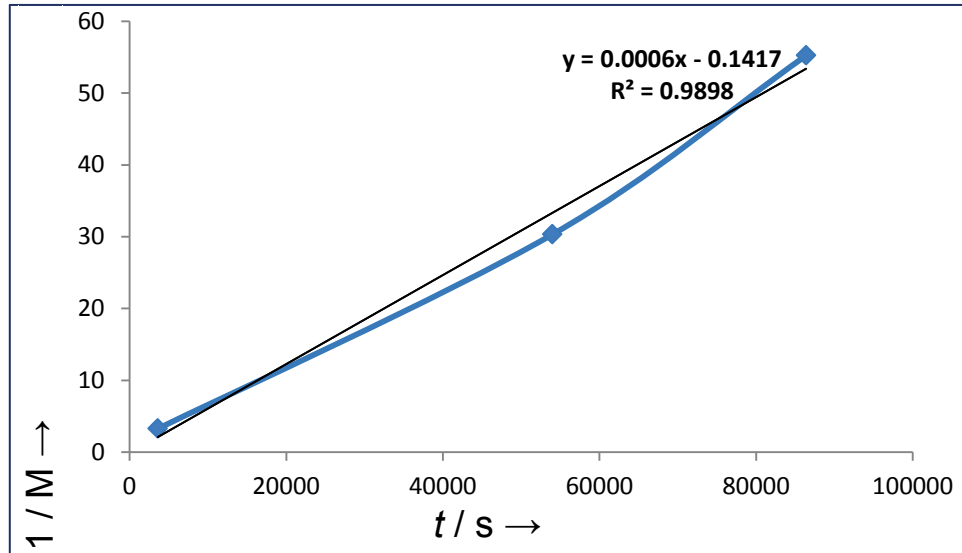
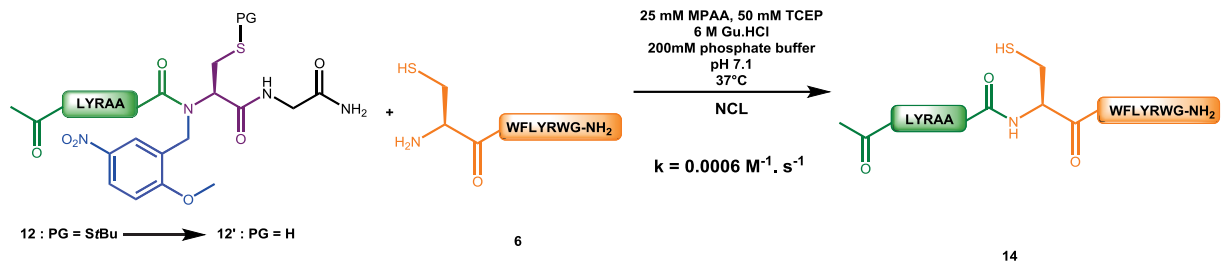
$$kt = \frac{1}{[B]_0 - [A]_0} \ln \left( \frac{[A]_0 \cdot ([B]_0 - [A])}{[B]_0 \cdot ([A]_0 - [A])} \right)$$

Values for  $t < 10\text{min}$  were not taken into account due to the early *StBu* deprotection step.

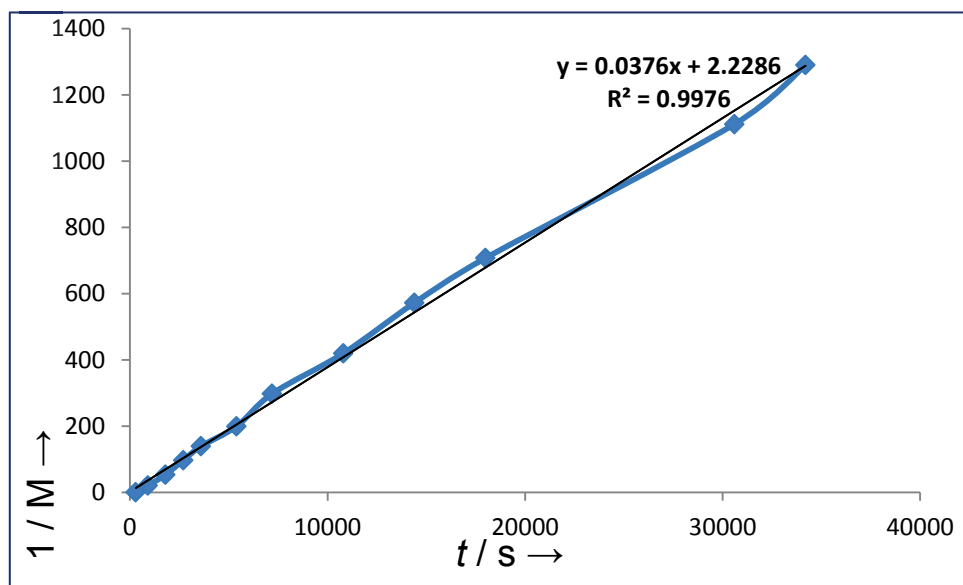
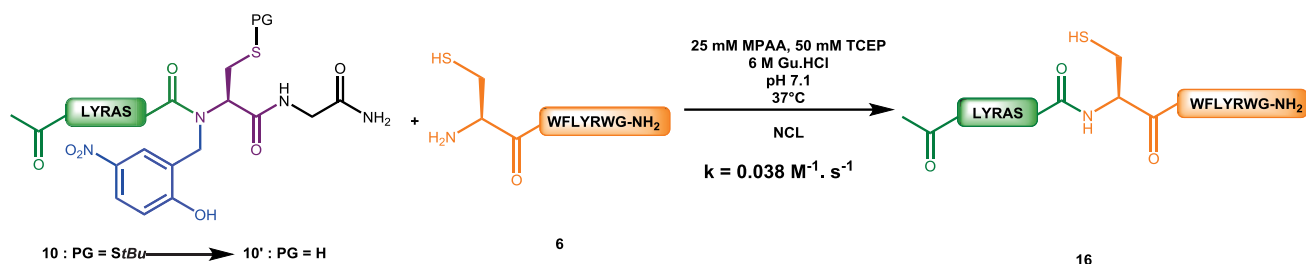


Supplementary figure S74: Apparent second order kinetic constant for NCL using **9**.  
(Conditions corresponding to Figure 3 of the article)

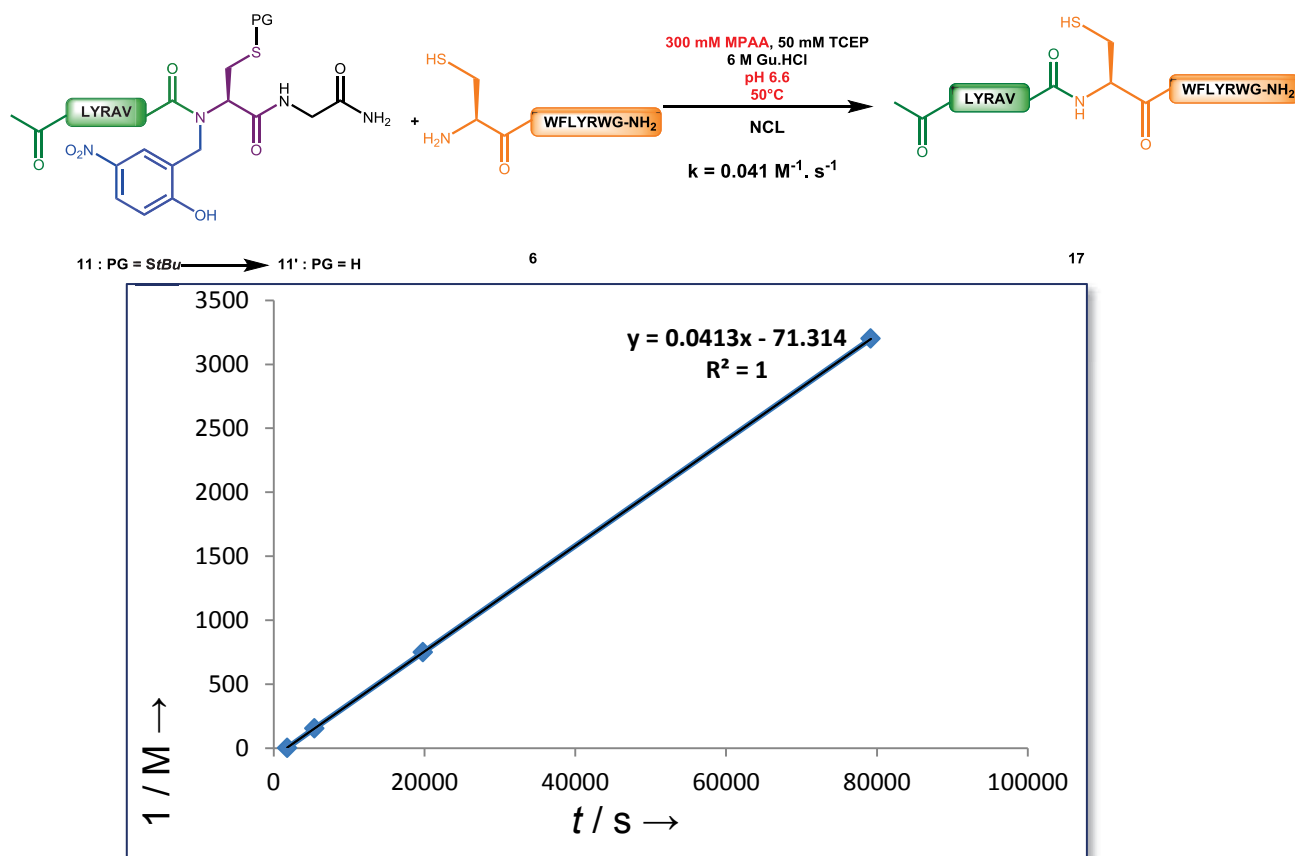
Partie expérimentale - article 1



Supplementary figure S75: Apparent second order kinetic constant for NCL using **12**.  
(Conditions corresponding to Figure 3 of the article)



**Supplementary figure S76:** Apparent second order kinetic constant for NCL using **10**.  
(Conditions corresponding to Figure 3 of the article)

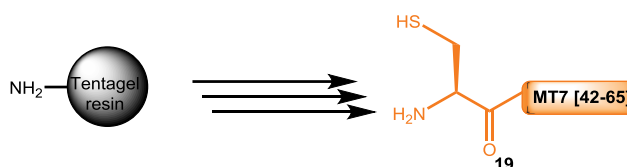


Supplementary figure S77: Apparent second order kinetic constant for NCL using **11** (Conditions corresponding to Figure 3 of the article)

### 11-Application to MT7

#### 11a- Synthesis of cysteinyl peptide MT7 [42-65] (**19**)

Amino acid sequence of the MT7 [42-65] segment:



Supplementary scheme S31: Synthesis of cysteinyl peptide MT7 [42-65] **19**.

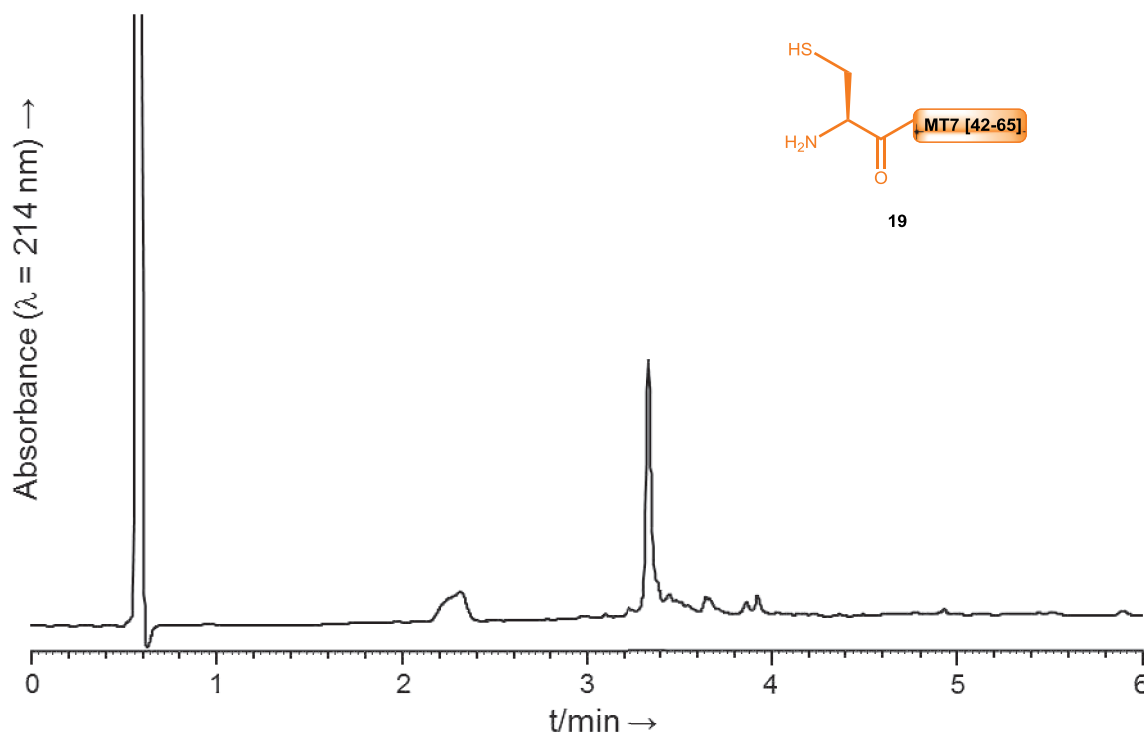
Peptide **19** was obtained through automated SPPS (protocol p S3) starting from Tentagel resin (476, 0.21 mmol/g, 0.1 mmol). Fmoc-K(Boc)-Mppa-OH (323.3 mg, 0.5 mmol, 5 equiv.) and HATU (190 mg, 0.5 mmol, 5 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175  $\mu$ L, 1 mmol, 10 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10

mL) for 3 min and the complete sequence was installed through standard Fmoc-SPPS (elongation yield = 71 %). After cleavage of a quarter of the resin (25  $\mu\text{mol}$ , protocol p S3), product **19** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 15-23 % B/A over 16 min), 4.06  $\mu\text{mol}$  were obtained (yield = 16.2 %, UV titration at 280 nm, molar extinction coefficient = 1290  $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ).

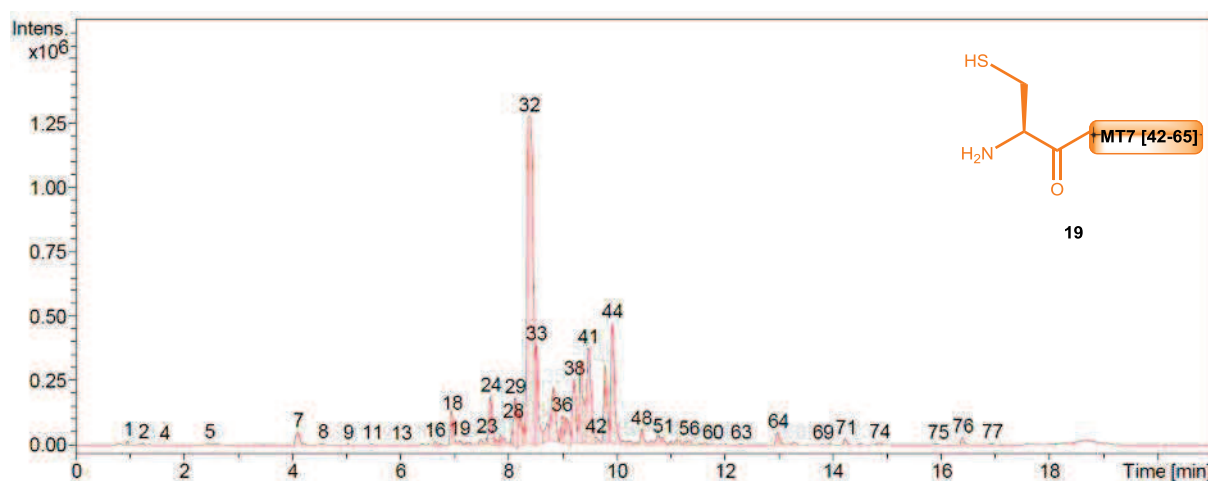
**19:**

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{104}\text{H}_{173}\text{N}_{32}\text{O}_{36}\text{S}_5$ : 2606.1293, found: 2606.1277.

**HPLC**: retention time: 3.33 min (Chromolith, gradient: 10-40 % B/A over 5 min).



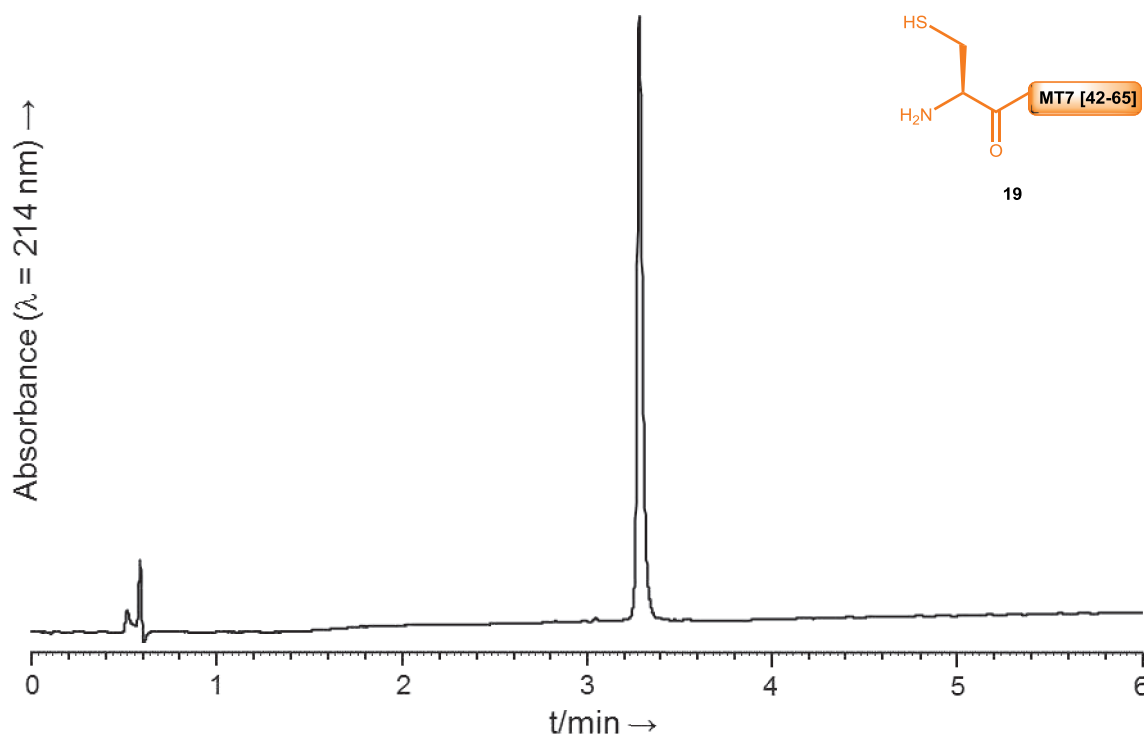
Supplementary figure S78: HPLC trace of crude **19**.



Supplementary figure S79: LC/MS analysis of crude **19**.

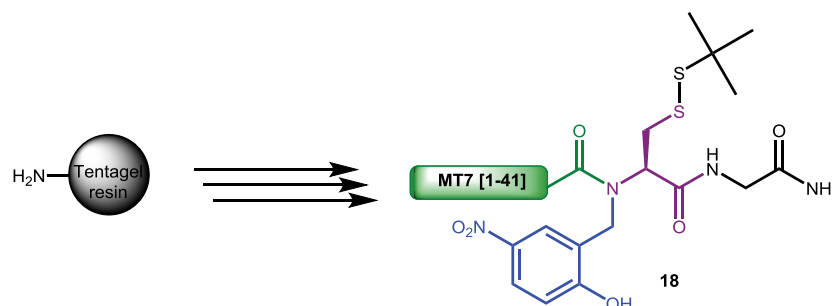
Peak (tr (min))	[MH] <sup>+</sup> (m/z) Calc.	[MH] <sup>+</sup> (m/z) found	Attributed to
7 (4.11)	1013.3855	1013.3823	Ac-[57-65]
18 (6.96)	1240.5125	1240.5095	Ac-[55-65]
24 (7.67)	1610.7089	1610.7057	Ac-[52-65]
28 (8.09)	1339.5809	1339.5755	Ac-[54-65]
32 (8.38)	2606.1293	2606.1277	<b>19</b>
33 (8.51)	-	2588.1143	Not attributed
35 (8.83)	2198.9997	2198.9964	Ac-[47-65]
38 (9.20)	-	1886.7800	Not attributed
39 (9.34)	2662.1919	2662.1880	<b>19</b> + 56 Da (tBu)
41 (9.41)	2662.1919	2662.1884	<b>19</b> + 56 Da (tBu)
43 (9.79)	2662.1919	2662.1869	<b>19</b> + 56 Da (tBu)
44 (9.91)	2662.1919	2662.1872	<b>19</b> + 56 Da (tBu)

Supplementary table S5: Attribution of selected peaks observed in LC/MS analysis of crude **19**.



Supplementary figure S80: HPLC trace of pure **19**.



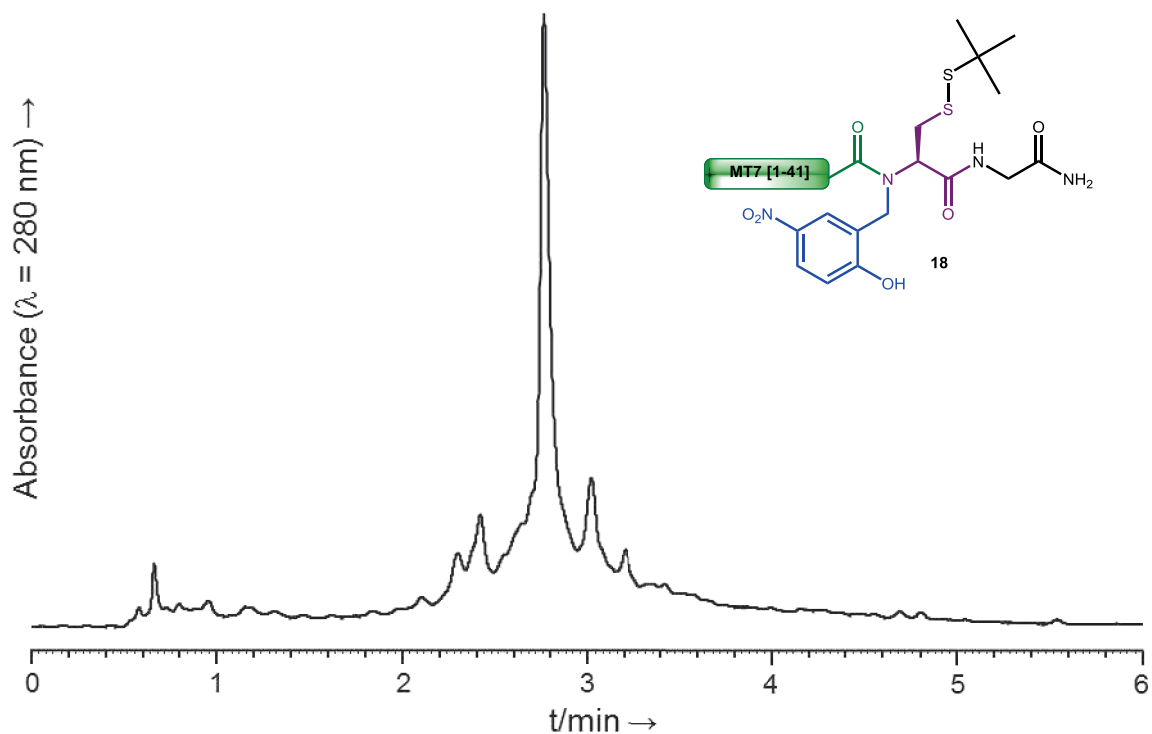
11b- Synthesis of crypto-thioester peptide MT7 [1-41]-(Hnb)C(S*t*Bu)G-NH<sub>2</sub> (**18**)Amino acid sequence of the MT7 [1-41]-(Hnb)C(S*t*Bu)G-NH<sub>2</sub> segment:H-LTCVKNSNSIWFPTSEDCPDGQNLCKRWQYISPRMYDFTRG(Hnb)C(S*t*Bu)G-NH<sub>2</sub>Supplementary scheme S32: Synthesis of crypto-thioester peptide MT7 [1-41]-(Hnb)C(S*t*Bu)G-NH<sub>2</sub> **18**.

Tentagel R resin (476 mg, 0.21 mmol/g, 0.1 mmol) was introduced into a polypropylene syringe fitted with a polypropylene frit and a PTFE stopcock and swollen with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Then, Fmoc-Rink-OH (270 mg, 0.5 mmol, 5 equiv.) and HATU (190 mg, 0.5 mmol, 5 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (175  $\mu$ L, 1 mmol, 10 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Gly-OH (297 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (2 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Then, Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. Fmoc-Cys(S*t*Bu)-OH (432 mg, 1 mmol, 10 equiv.) and HCTU (372 mg, 0.90 mmol, 9 equiv.) were dissolved in NMP (10 mL) prior to addition of *i*Pr<sub>2</sub>NEt (348  $\mu$ L, 2 mmol, 20 equiv.). The resulting solution was immediately added to the resin. The syringe was left for 1.5 h under gentle stirring, then the syringe was drained and the resin was washed with NMP. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (10 mL) for 3 min. The *N*-2-hydroxy-5-nitrobenzyl group was then introduced following the general procedure (p S5). Then, the complete MT7 [1-41] sequence was installed on a quarter of the resin trough standard Fmoc-SPPS (25  $\mu$ mol, protocol S3). After cleavage of the resin (protocol p S3), the crude product **18** was pre-purified using a Sep-Pak cartridge (Waters Vac 12cc, C<sub>18</sub> - 2g, washing: 50 mL 0 % B/A then 50 mL 20 % B/A and elution: 50 mL 40 % B/A) and then purified by semi-preparative RP-HPLC (Nucleosil, gradient: 35-43 % B/A over 8 min) 1.62  $\mu$ mol were obtained (yield of 12.96 %, UV titration at 280 nm, molar extinction coefficient = 15940 L.mol<sup>-1</sup>.cm<sup>-1</sup>).

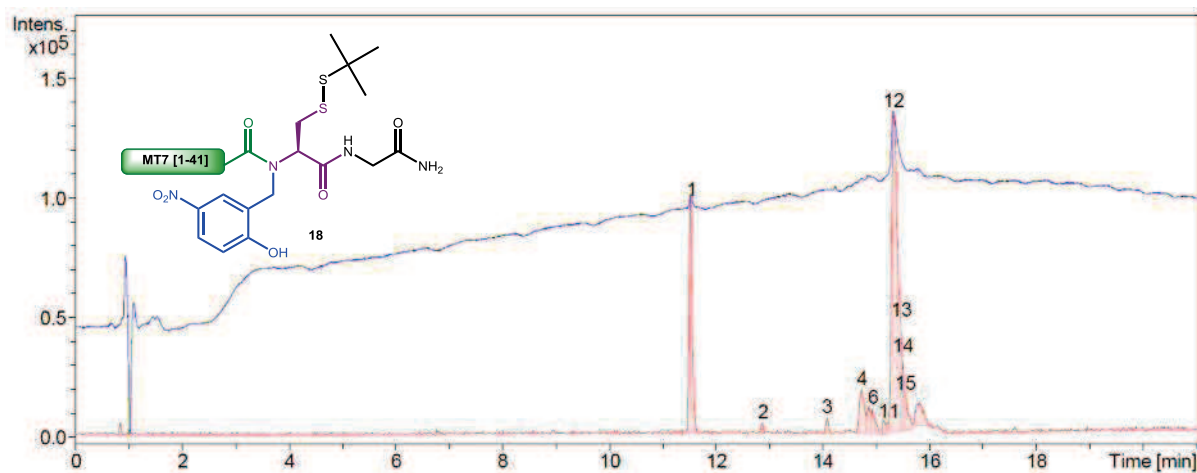
**18:**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd for  $C_{234}H_{345}N_{62}O_{67}S_6$ : 5287.3819, found: 5287.3818.

**HPLC**: retention time: 2.82min (Chromolith, gradient: 30-50 % B/A over 5 min).



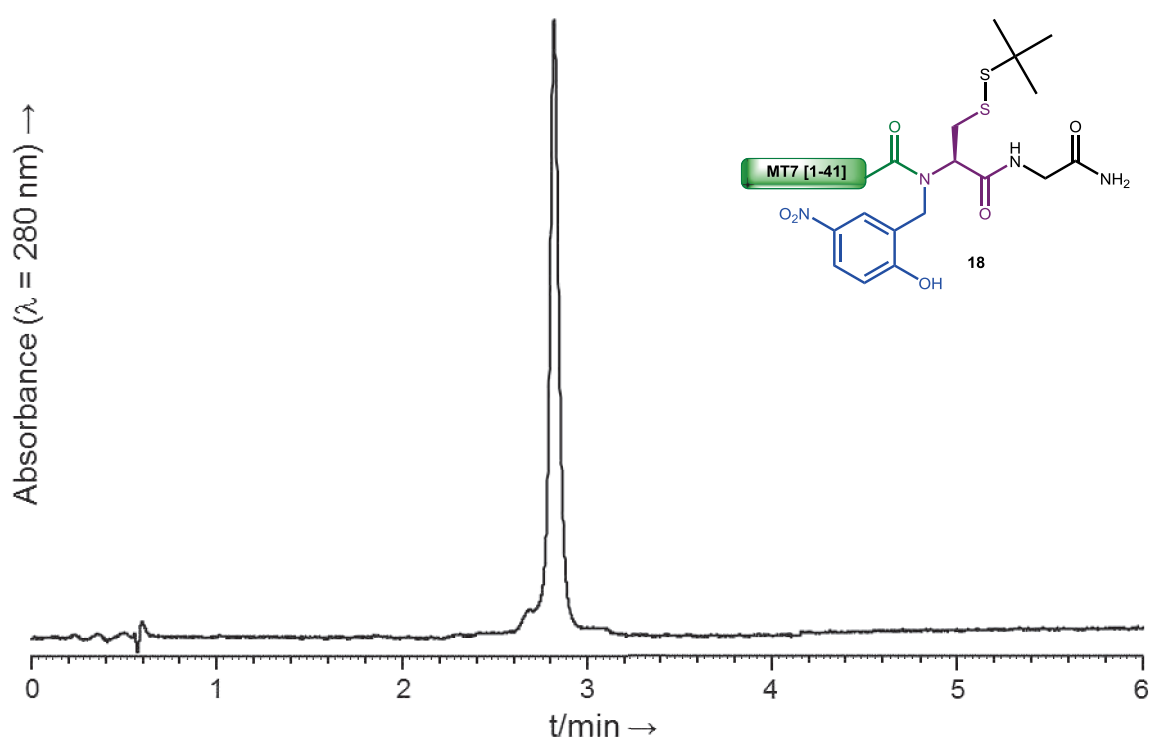
Supplementary figure S81: HPLC trace of crude 18.



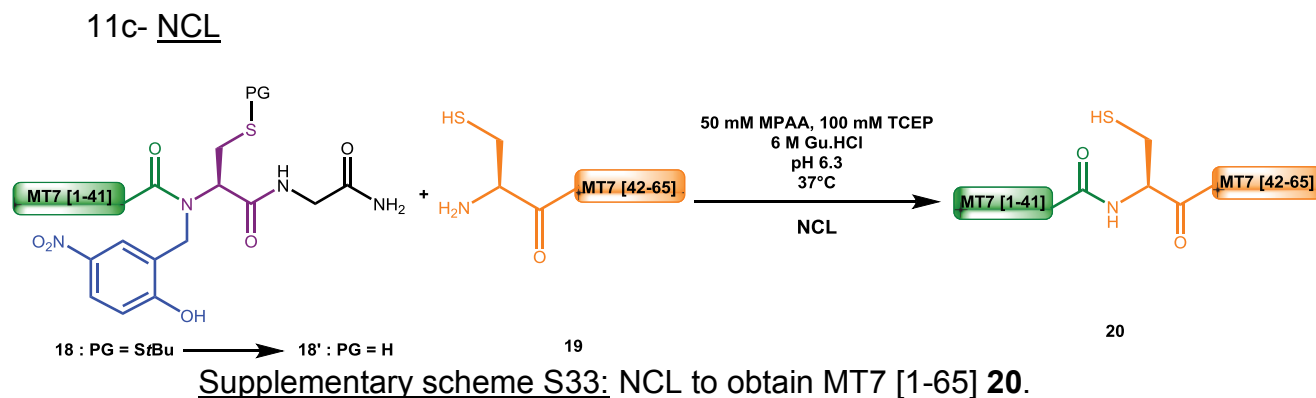
Supplementary figure S82: LC/MS analysis of crude 18.

Peak (tr (min))	[MH] <sup>+</sup> (m/z) Calc.	[MH] <sup>+</sup> (m/z) found	Attributed to
1 (11.53)	-	1826.8617	Not attributed
2 (12.87)	-	1882.9243	Not attributed
3 (14.06)	-	614.2409	Not attributed
4 (14.72)	-	3479.5302	Not attributed
5 (14.86)	5303,3768	5303.3519	<b>18</b> +16 Da (Oxidized Met)
10 (15.14)	2691,2423	2691.2184	Ac-[25-41]-(Hnb)C( <i>St</i> Bu)-G-NH <sub>2</sub>
12 (15.31)	5287.3819	5287.3818	<b>18</b>
13 (15.45)	-	5287.3423	Not attributed
14 (15.48)	-	5290.3463	Not attributed
16 (15.80)	5343.4444	5343.4466	<b>18</b> + 56 Da ( <i>t</i> Bu)

Supplementary table S6: Attribution of selected peaks observed in LC/MS analysis of crude **18**.



Supplementary figure S83: HPLC trace of pure **18**.



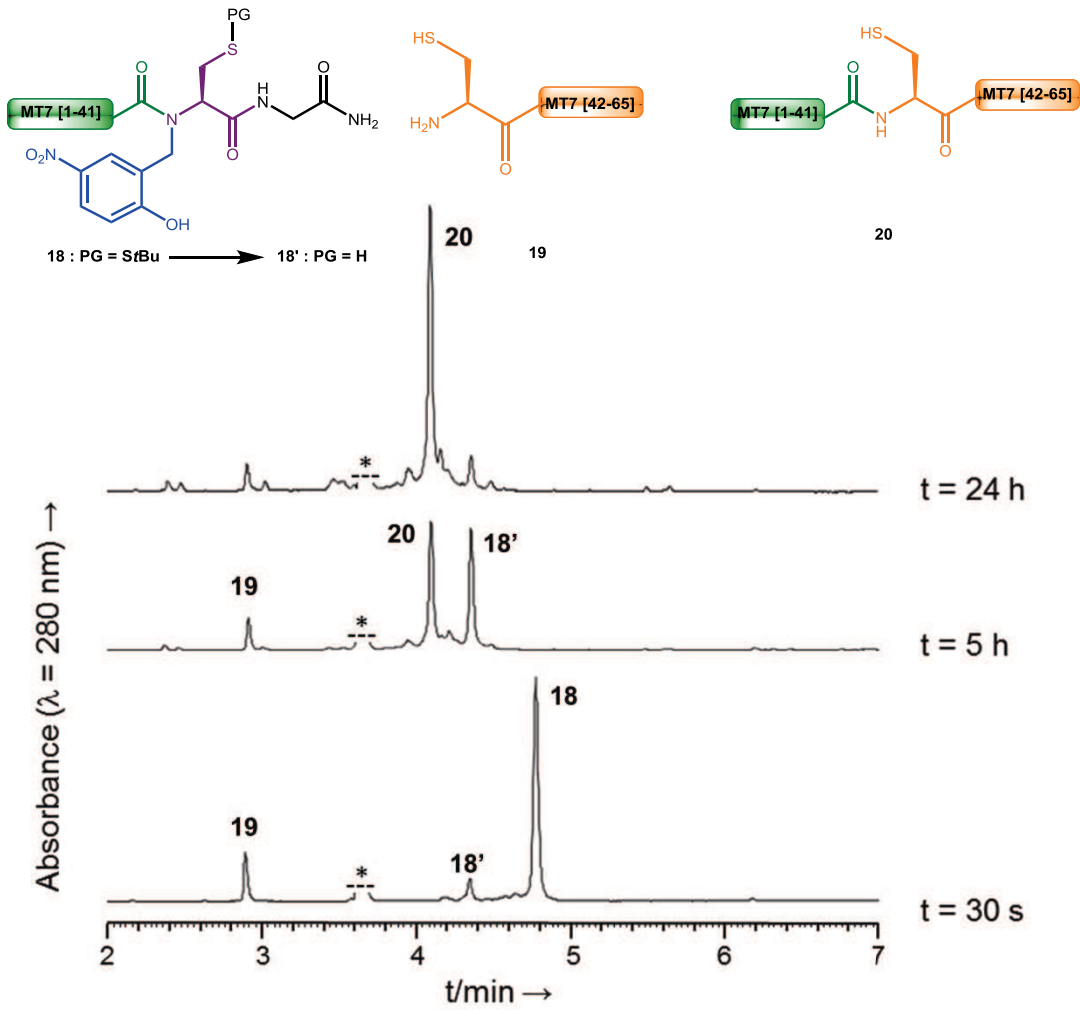
170  $\mu\text{L}$  of a degassed 0.2 M pH 6.3 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride were added to 1.05 mg of the MT7 [1-41]-(Hnb)C(SfBu)G-NH<sub>2</sub> peptide **18** (0.168  $\mu\text{mol}$ , final concentration 1 mM) and 0.66 mg of the MT7 [42-65] peptide **19** (0.218  $\mu\text{mol}$ , 1.3 equiv., final concentration 1.33 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 10-60% B/A over 6 min). For this, aliquots of 2  $\mu\text{L}$  were diluted in 108  $\mu\text{L}$  of 1.5% TFA in water and 100  $\mu\text{L}$  were injected to the HPLC. After 24h, the reaction mixture was acidified with TFA (10  $\mu\text{L}$ ), 1 mL solvent A and 200  $\mu\text{L}$  solvent B. this solution was extracted with diethylether (4  $\times$  10 mL) to remove the MPAA. TCEP was added (final 100 mM) and the pH was adjusted to 5.0; after 20 min the pH was adjusted to 1 and the ligation product **20** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 25-43 % B/A over 10 min). 23.4 nmol were obtained (yield = 14 %, UV titration at 280 nm, molar extinction coefficient = 14870 L.mol<sup>-1</sup>.cm<sup>-1</sup>).

**20:**

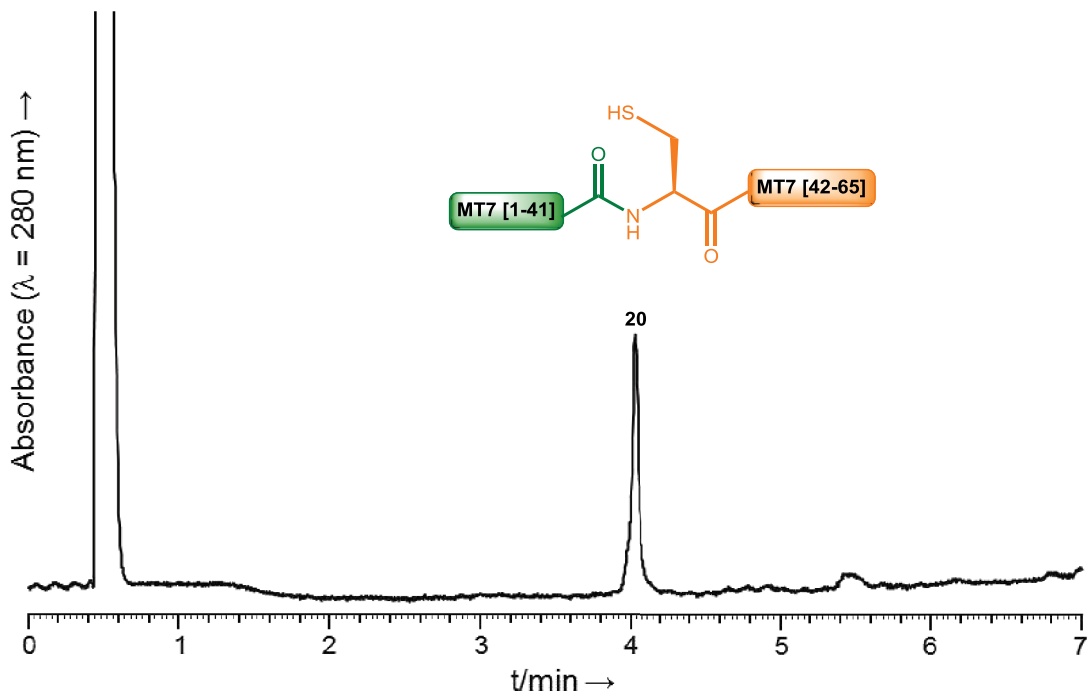
**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> calcd for C<sub>322</sub>H<sub>493</sub>N<sub>90</sub>O<sub>98</sub>S<sub>9</sub>: 7476.3847, found: 7476.4022.

**HPLC**: retention time: 4.34min (Chromolith, gradient: 10-60% B/A over 6 min).

Partie expérimentale - article 1



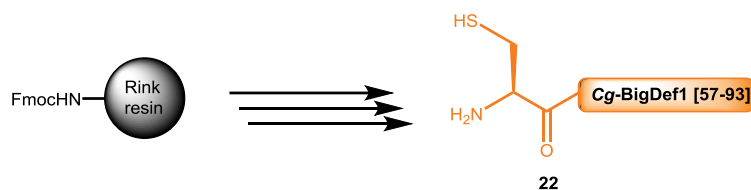
Supplementary figure S84: NCL of MT7 [1-65] **20**. \*: MPA



Supplementary figure S85: HPLC trace of pure MT7 [1-65] **20**.

**12-Application to Cg-BigDef1****12a- Synthesis of cysteinyl peptide Cg-BigDef1 [57-93] (**22**)**

Amino acid sequence of the Cg-BigDef1 [57-93] segment:



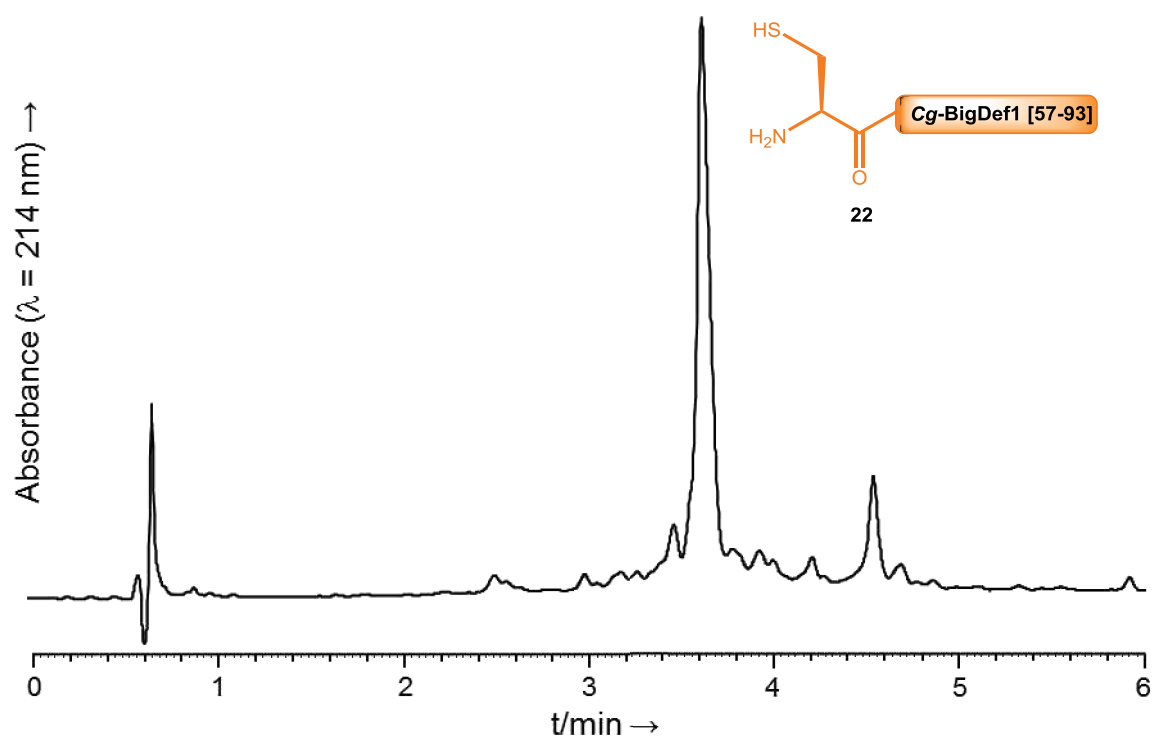
Supplementary scheme S34: Synthesis of cysteinyl peptide Cg-BigDef1 [57-93] peptide **22**.

Peptide **22** was obtained through automated SPPS (protocol p S3) starting from H-Rink amide Chemmatrix resin (200 mg, 0.52 mmol/g, 0.1 mmol); the elongation yield was 64% (determined by UV titration). After cleavage of a part of the resin (21  $\mu\text{mol}$ ) (protocol p S3), the product **22** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 25-38 % B/A over 13 min) 3.03  $\mu\text{mol}$  were obtained (yield = 14.4%, UV titration at 280 nm, molar extinction coefficient = 13580  $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ).

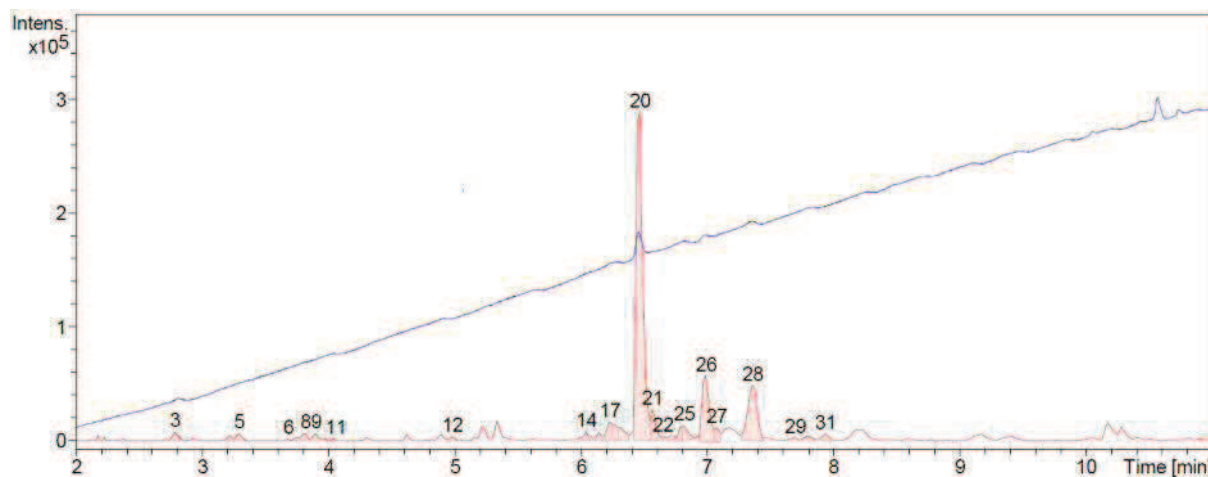
**22:**

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd for  $\text{C}_{185}\text{H}_{271}\text{N}_{64}\text{O}_{53}\text{S}_6$ : 4428.8802, found: 4428.8855.

**HPLC**: retention time: 3.57 min (Chromolith, gradient: 23-34 % B/A over 5 min).

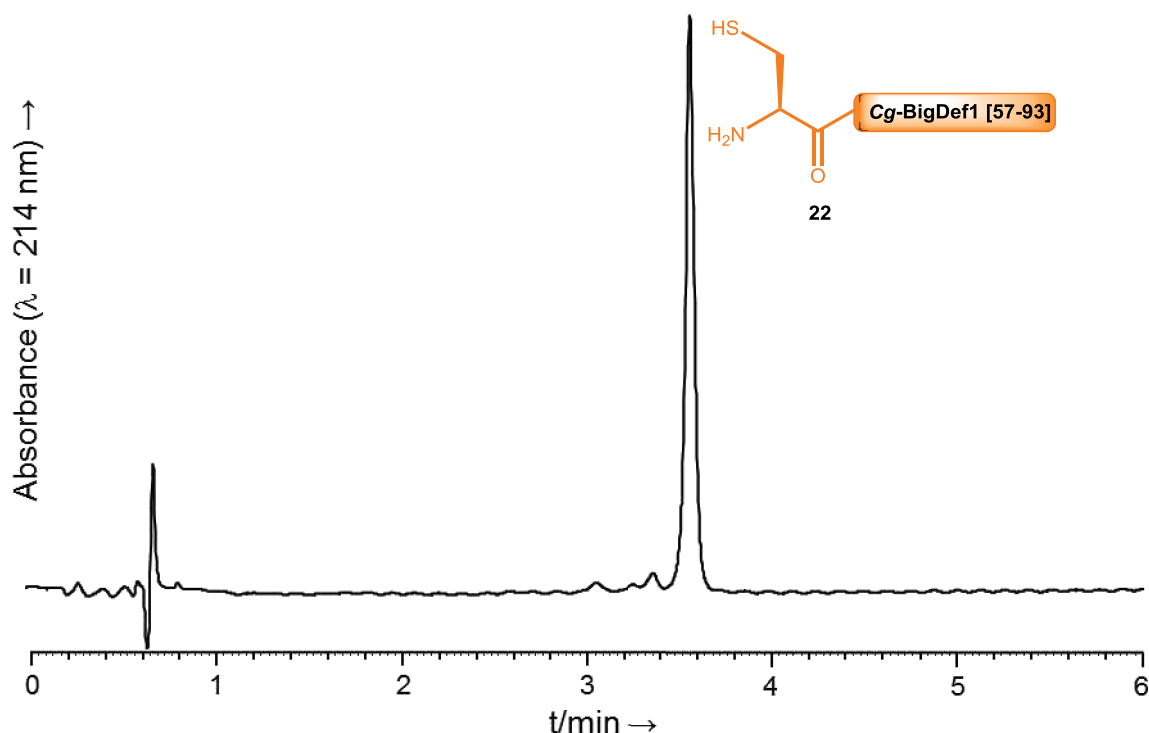


Supplementary figure S86: HPLC trace of crude **22**.

Supplementary figure S87: LC/MS analysis of crude **22**.

Peak (tr (min))	$[MH]^+$ ( $m/z$ ) Calc.	$[MH]^+$ ( $m/z$ ) found	Attributed to
1 (1.79)	888.4409	888.4405	Ac-[87-93]
2 (1.82)	888.4409	888.4401	Ac-[87-93]
3 (2.79)	1351.6588	1351.6580	Ac-[83-93]
4 (3.22)	1454.6680	1454.6670	Ac-[82-93]
7 (3.75)	1896.8492	1896.8480	Ac-[78-93]
8 (3.81)	1782.8063	1782.8043	Ac-[79-93]
9 (3.90)	1668.7634	1668.7612	Ac-[80-93]
16 (6.14)	2229.9969	2229.9911	Ac-[76-93]
20 (6.46)	4428.8802	4428.8855	<b>22</b>
23 (6.72)	2738.1775	2738.1721	Ac-[72-93]
26 (6.99)	4484.9481	4484.9377	<b>22</b> + 56 Da (tBu)
28 (7.36)	3410.4465	3410.4426	Ac-[66-93]

Supplementary table S7: Attribution of selected peaks observed in LC/MS analysis of crude **22**.



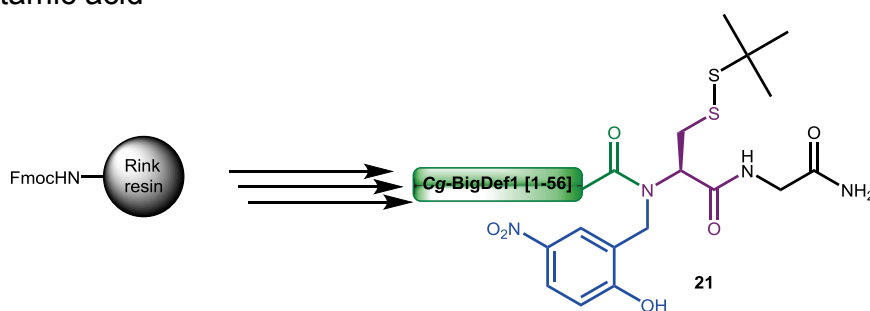
Supplementary figure S88: HPLC trace of pure **22**.

12b- Synthesis of crypto-thioester peptide Cg-BigDef1 [1-56]-(Hnb)C(StBu)G-NH<sub>2</sub> (**21**)

Amino acid sequence of the Cg-BigDef1 [1-56]-(Hnb)C(StBu)G-NH<sub>2</sub> segment:

XAQALLPIASYAGLTVSAPVFAALVTVYGAYALYRYNIRRENSYQRIRSDHDSHS(Hnb)C(StBu)G-NH<sub>2</sub>

X = pyroglutamic acid



Supplementary scheme S35: Synthesis of crypto-thioester peptide Cg-BigDef1 [1-56]-(Hnb)C(StBu)G-NH<sub>2</sub> peptide **21**.

Peptide **21** was obtained through automated SPPS (protocol p S3) and using the automated reductive amination (protocol p S34) starting from Fmoc-Rink tentagel resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol) and for some amino acids double or triple coupling have been applied (see below) as determined as difficult couplings from a previous synthesis of the sequence not incorporating an (Hnb)Cys transthioesterification device. After cleavage of the resin (protocol p S3), the product **21** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 48-58 % B/A over



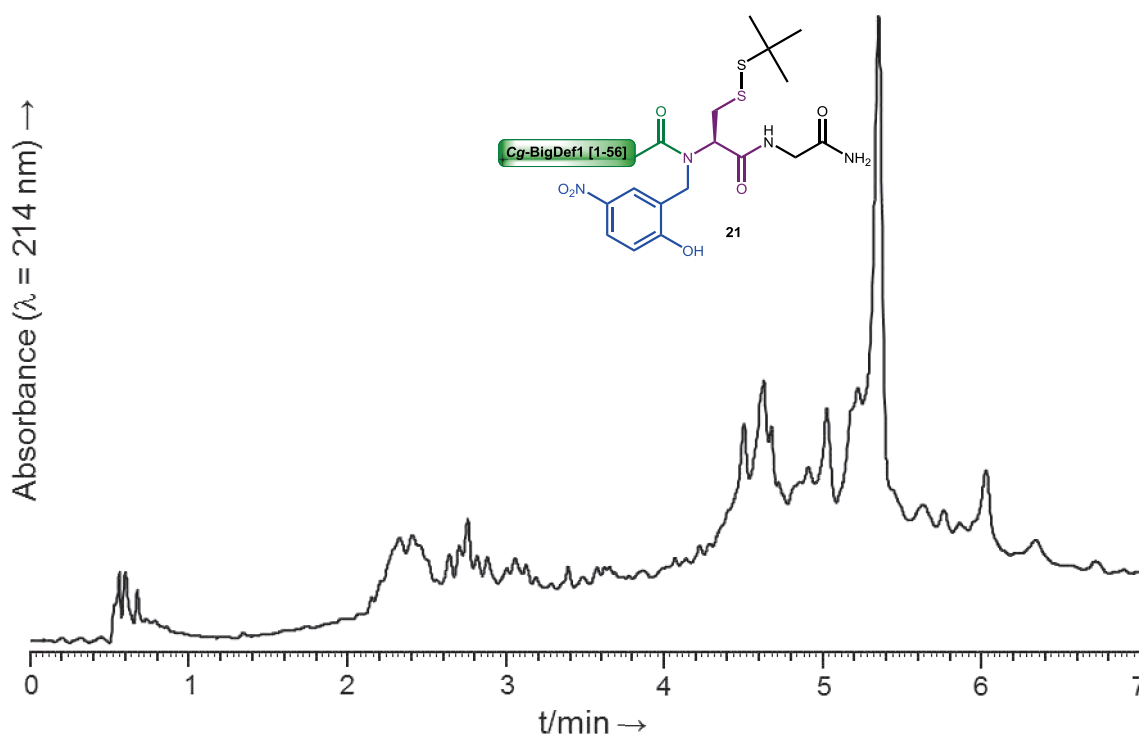
30 min) 1.69  $\mu\text{mol}$  were obtained (yield = 6.8 %, UV titration at 280 nm, molar extinction coefficient =  $10100 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ).

XAQALLPIASYAGLTVSAPVFAALVTVYGAYALRYNIRRENSYQRRIRSDHDSHS(Hnb)C(StBu)G-NH<sub>2</sub>  
 X = pyroglutamic acid; green = double coupling, blue = triple coupling

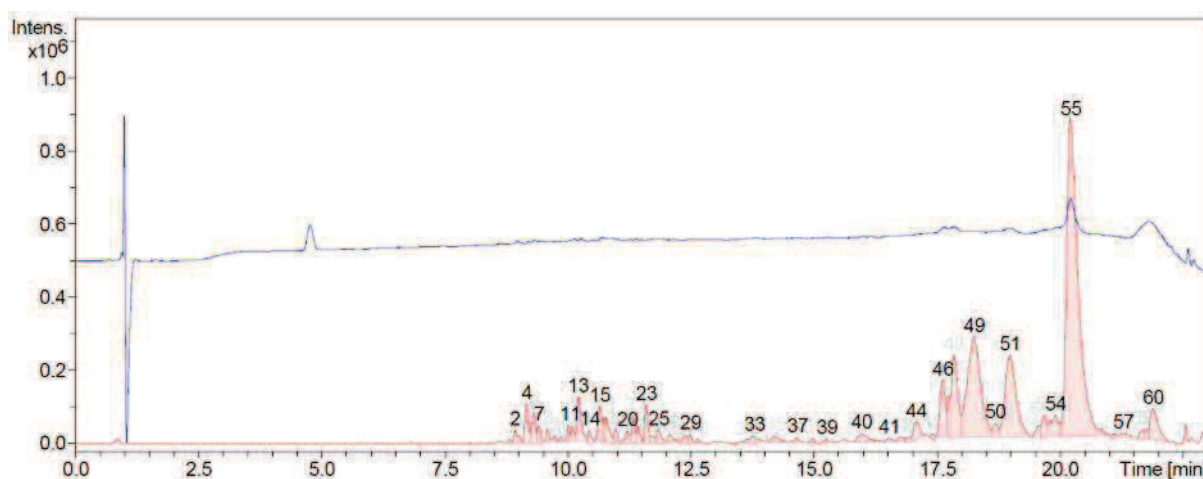
**21:**

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd for C<sub>298</sub>H<sub>457</sub>N<sub>86</sub>O<sub>86</sub>S<sub>2</sub>: 6680.3472, found: 6680.3554

**HPLC**: retention time: 5.40 min (Chromolith, gradient: 20-70 % B/A over 6 min).



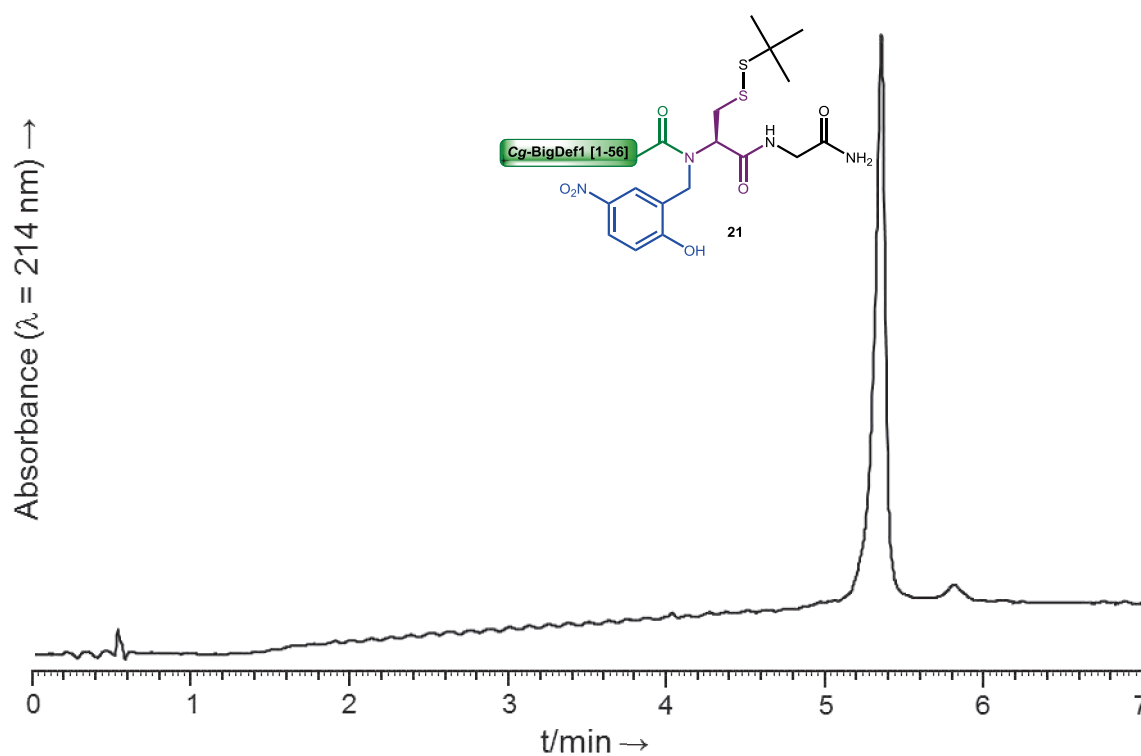
Supplementary figure S89: HPLC trace of crude **21**.



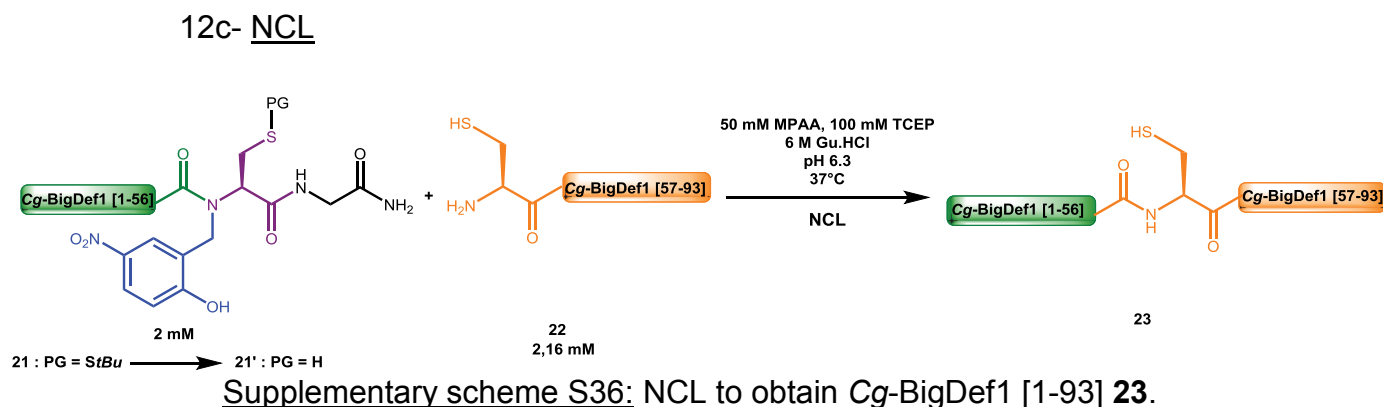
Supplementary figure S90: LC/MS analysis of crude **21**.

Peak (tr (min))	[MH] <sup>+</sup> (m/z) Calc.	[MH] <sup>+</sup> (m/z) found	Attributed to
4 (9.16)	2966.3611	2966.3649	Ac-[37-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
13 (10.21)	-	3545.6465	Not attributed
15 (10.65)	3923.8319	3923.8386	Ac-[29-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
23 (11.59)	1493.5902	1493.5893	Ac-[48-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
46 (17.61)	5583.7555	5583.7657	Ac-[12-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
47 (17.75)	5904.8880	5904.8903	Ac-[9-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
48 (17.83)	5746.8189	5746.8267	Ac-[11-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
49 (18.24)	4984.4277	4984.4364	Ac-[19-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
51 (18.97)	6115.0248	6115.0354	Ac-[7-56]-(Hnb)C(StBu)-G-NH <sub>2</sub>
55 (20.20)	6680.3472	6680.3554	<b>21</b>
60 (21.88)	-	6776.3467	Not attributed

Supplementary table S8: Attribution of selected peaks observed in LC/MS analysis of crude **21**.



Supplementary figure S91: HPLC trace of pure **21**.



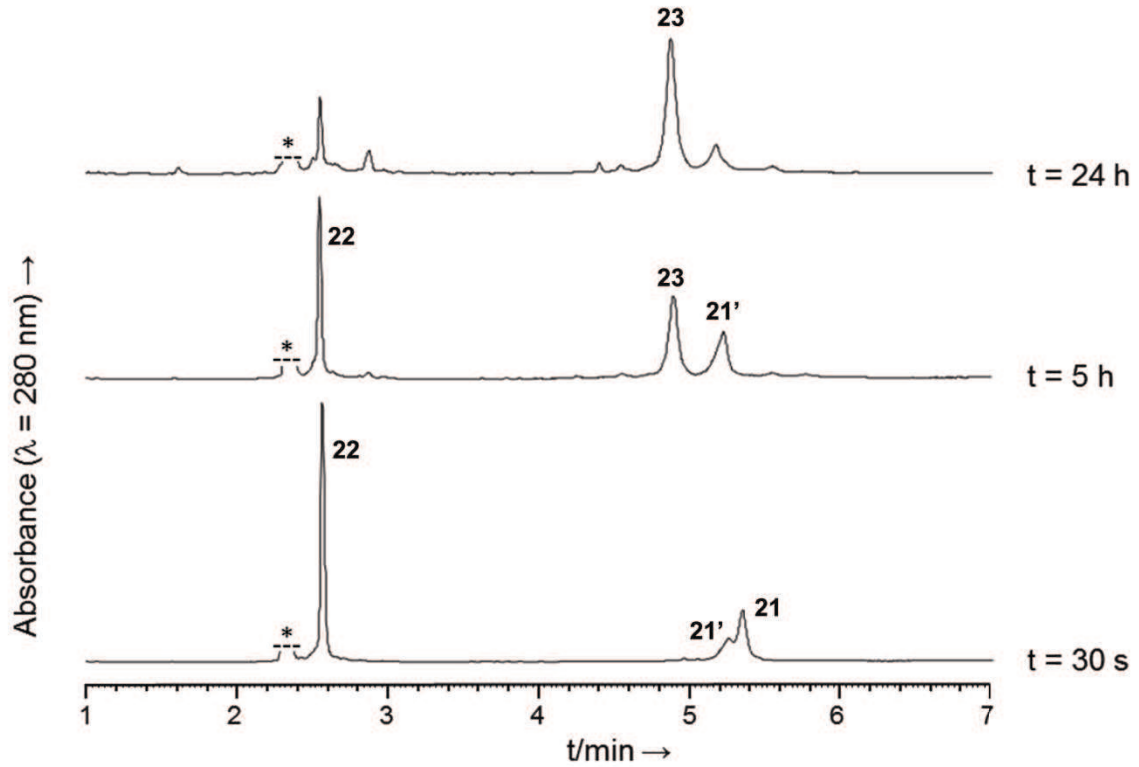
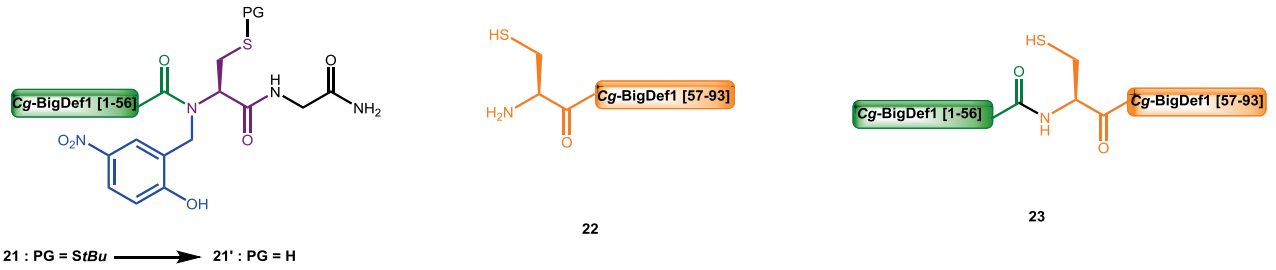
352.5  $\mu\text{L}$  of a degassed 0.2 M pH 6.3 sodium phosphate buffer containing 50 mM MPAA, 100 mM TCEP and 6 M guanidine hydrochloride were added to 5.38 mg of the Cg-BigDef1 [1-56]-(Hnb)C(StBu)G-NH<sub>2</sub> peptide **21** (0.705  $\mu\text{mol}$ , final concentration 2 mM) and 4 mg of the Cg-BigDef1 [57-93] peptide **22** (0.763  $\mu\text{mol}$ , 1.08 equiv., final concentration 2.16 mM) under argon. The ligation was carried out at 37°C and monitored by RP-HPLC (Chromolith, gradient: 20-70 % B/A over 6 min). For this, aliquots of 2  $\mu\text{L}$  were diluted in 108  $\mu\text{L}$  of 1.5% TFA in water and 100  $\mu\text{L}$  were injected to the HPLC. After 24h, the reaction mixture was acidified with TFA (30  $\mu\text{L}$ ) and 2 mL of 1.5% TFA in water; this solution was extracted with diethylether (4  $\times$  10 mL) to remove the MPAA. The ligation product precipitated; after centrifugation and removal of the supernatant, the solid was diluted in 400  $\mu\text{L}$  of solvent B and 1.2 mL of solvent A (V total = 1.6 mL); 45 mg of TCEP were added (final 100 mM) and the pH was adjusted to 4.5, after 20 min the pH was adjusted to 1 and the ligation product **23** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 50-60 % B/A over 10 min). 130 nmol were obtained (yield = 18.4 %, UV titration at 280 nm, molar extinction coefficient = 21320 L.mol<sup>-1</sup>.cm<sup>-1</sup>).

**23:**

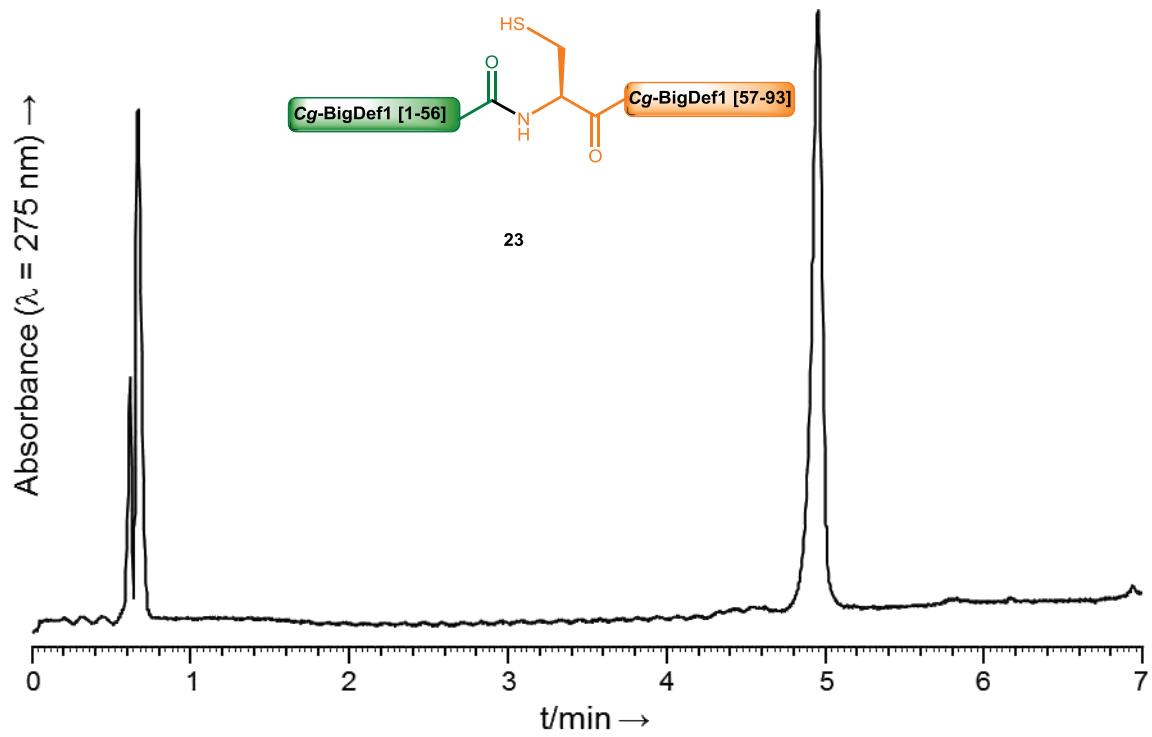
**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> average mass calcd for C<sub>467</sub>H<sub>703</sub>N<sub>146</sub>O<sub>134</sub>S<sub>6</sub> : 10699.0090, found: 10699.1541

**HPLC**: retention time: 4.96 min (Chromolith, gradient: 20-70 % B/A over 6 min).

Partie expérimentale - article 1



Supplementary figure S92: NCL reaction to obtain Cg-BigDef1 [1-93] **23**. \*: MPA



Supplementary figure S93: HPLC trace of pure Cg-BigDef1 [1-93] **23**.

## Partie expérimentale article 2 :

### A native chemical ligation strategy to overcome side reactions during Fmoc-based synthesis of C-terminal cysteine-containing peptides

Lelièvre D.; Terrier V. P.; Delmas A. F.; Aucagne V., *Org. Lett.*, **2016**, article soumis

#### Supporting Information

##### Table of contents:

1) General information	263
2) General procedures for solid phase peptide synthesis	264
3) Synthesis of [1-51] AhPDF1.1b ( <b>1</b> ) via standard Fmoc-SPPS	265
4) Studies on the <i>N</i> -acylation of the <i>N</i> -Hnb-Cys device with Fmoc-Pro-OH	270
a- Preliminary results	270
b- Synthesis of <i>N</i> -Hnb-Cys resin <b>S5</b>	272
c- Study of the intramolecular <i>N</i> → <i>O</i> shift kinetics	273
d- Selective cleavage of the Hnb esters	274
5) Synthesis of [1-50] AhPDF1.1b-(Hnb)C( <i>St</i> Bu)G-NH <sub>2</sub> ( <b>2</b> )	277
6) Synthesis of [1-51] AhPDF1.1b ( <b>1</b> ) via NCL	281
a- Optimization of the NCL conditions on purified peptide <b>2</b>	281
b- Application to the crude peptide <b>2</b>	284
7) Application to the synthesis of [1-51] AhPDF1.5 ( <b>6</b> )	286
a- Synthesis of [1-50] AhPDF1.5-(Hnb)C( <i>St</i> Bu)G-NH <sub>2</sub> ( <b>5</b> )	286
b- Synthesis of <b>6</b> via NCL	288

## 1- General information

All reagents and solvents were used without further purification. Protected amino acids, Rink's amide linker and HCTU were purchased from Merck Biosciences (Nottingham, UK). H-Cys(Trt)-Trityl-ChemMatrix resin was purchased from PCAS BioMatrix (St-Jean-sur-Richelieu, Québec). Aminomethyl TentaGel R resin was purchased from Rapp polymers (Tuebingen, Germany). Peptide synthesis grade DMF and HATU were purchased 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).

High resolution ESI-MS analyses were performed on a maXis<sup>TM</sup> ultra-high-resolution Q-TOF 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. In the latter case, the sample was co-crystallized with a solution of  $\alpha$ -cyano-4-hydroxy-cinnamic acid (HCCA) as a matrix. The reported  $m/z$  values correspond to the monoisotopic ions if not specified otherwise. Peptides incorporating an *N*-(2-hydroxy-5-nitrobenzyl) group displayed a typical MALDI fragmentation pattern, consisting in -16 Da and -34 Da minor peaks in addition to the  $[MH]^+$  peak.

HPLC analyses and 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 autosampler. Nucleosil C18 (300 Å, 5  $\mu$ m, 250  $\times$  4.6 mm, 1 mL/min flow rate) or Chromolith HighResolution RP-18e (150 Å, 100  $\times$  4.6 mm, 3 mL/min flow rate) columns were used for analysis and Nucleosil C18 (300 Å, 5  $\mu$ m, 250  $\times$  10 mm, 3 mL/min flow rate) for purification. Chromatography was conducted at room temperature unless otherwise mentioned. Solvents A and B were 0.1% TFA in H<sub>2</sub>O and 0.1% TFA in MeCN, respectively. Each gradient was followed by a washing step (up to 95% B/A over 0.5 min for the HR Chromolith; over 1 min for the Nucleosil C18 column) to identify eventual co-products not eluted during the gradient.

LC/HRMS analyses were carried out on a Ultimate<sup>®</sup> 3000 RSLC HPLC system (Dionex, Germering, Germany), coupled with the maXis<sup>TM</sup> mass spectrometer and fitted with a Zorbax 300 SB-C18 RRHD (300 Å, 1.8  $\mu$ m, 100  $\times$  2.1 mm, 0.5 mL/min flow rate, 40 °C) column. Solvents A and B were 0.1% formic acid in H<sub>2</sub>O 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.

## 2- General procedures for solid phase peptide synthesis

Fmoc-based solid phase peptide syntheses (SPPS) were carried out either on a Prelude synthesizer from Protein Technologies or an ABI433 synthesizer from Applied Biosystems. Standard side-chain protecting groups were used for peptide elongation: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Glu(OtBu), Gln(Trt), His(Trt), Lys(Boc), Ser(*t*Bu), Thr(*t*Bu), Trp(Boc) and Tyr(*t*Bu), and Cys(*S**t*Bu) for the synthesis of the *N*-Hnb-Cys thioesterification device.

Syntheses on the Prelude synthesizer were performed at a 25  $\mu$ mol scale. Protected amino acids (0.25 mmol, 10 equiv.) were coupled using HCTU (98 mg, 0.238 mmol, 9.5 equiv.) and *i*Pr<sub>2</sub>NEt (87  $\mu$ L, 0.5 mmol, 20 equiv.) in NMP (3 mL) for 30 min. Capping of eventual unreacted amine groups was achieved by treatment with acetic anhydride (143  $\mu$ L, 1.51 mmol, 60 equiv.), *i*Pr<sub>2</sub>NEt (68  $\mu$ L, 0.39 mmol, 15.5 equiv.) and HOBt (6 mg, 0.044 mmol, 1.8 equiv.) in DMF (3 mL) for 7 min. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (3 mL) for 3 min. Syntheses on the ABI433 synthesizer were performed at a 100  $\mu$ mol scale, employing the same protocols, with an adjustment of the amounts of reagents and solvents according to a 100  $\mu$ mol scale, in order to achieve the same reagents stoichiometries and reaction concentrations.

The crude peptides were deprotected and cleaved from the resin through a treatment with TFA/H<sub>2</sub>O/DODT<sup>1</sup>/*i*Pr<sub>3</sub>SiH/phenol (78/5/10/2/5) for 3 h, and the peptide was precipitated by dilution into an ice-cold 1:1 diethyl ether/petroleum ether mixture, recovered by centrifugation and further washed three times with diethyl ether.

---

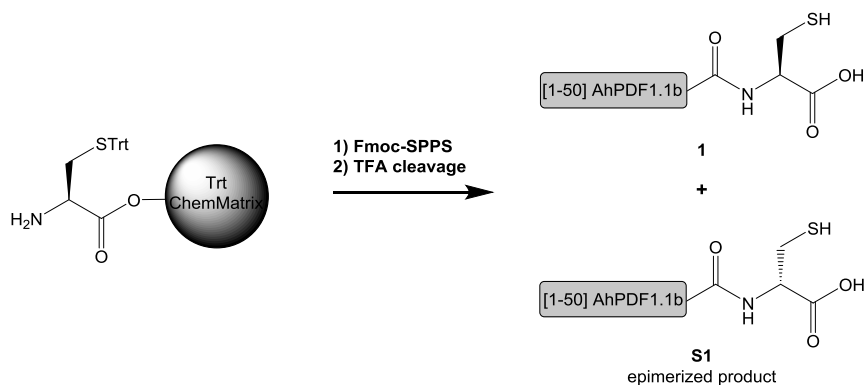
<sup>1</sup> DODT: 2,2'-(Ethylenedioxy)diethanethiol



### 3- Synthesis of [1-51] AhPDF1.1b (1) via standard Fmoc-SPPS

Amino acids sequence of [1-51] AhPDF1.1b (1):

$^1\text{ZRLCEKPSGTWSGVCGNNGACRNQCIRLEKARHGSCNYVFPAHKCICYFP}^{51}\text{C-OH}$   
 Z = pyroglutamic acid



Supplementary scheme S1: Synthesis of peptide **1** via standard Fmoc-SPPS. Synthesis was performed on a 433A synthesizer, starting from H-Cys(Trt)-Trityl-ChemMatrix resin (323 mg, loading: 0.31 mmol/g, 100  $\mu\text{mol}$ ).

During the SPPS elongation of [1-51] AhPDF1.1b, small aliquots of peptide-resin were withdrawn after 9 and 31 cycles. These aliquots were deprotected and cleaved (protocol p S3) to afford crude [42-51] AhPDF1.1b **S2** and [20-51] AhPDF1.1b **S3** that were analyzed by HPLC and MS in order to quantify the amount of epimerization along the SPPS process. In both cases, a minor, later eluted peak (8% and 27%, respectively) was observed showing the same  $m/z$  as the expected products **S2** and **S3**, respectively, attributed to the epimerization of the C-terminal L-Cys into D-Cys (co-products **S2'** and **S3'**, respectively). The L/D ratio was determined by integration of the respective HPLC peaks at 280 nm.

**S2**

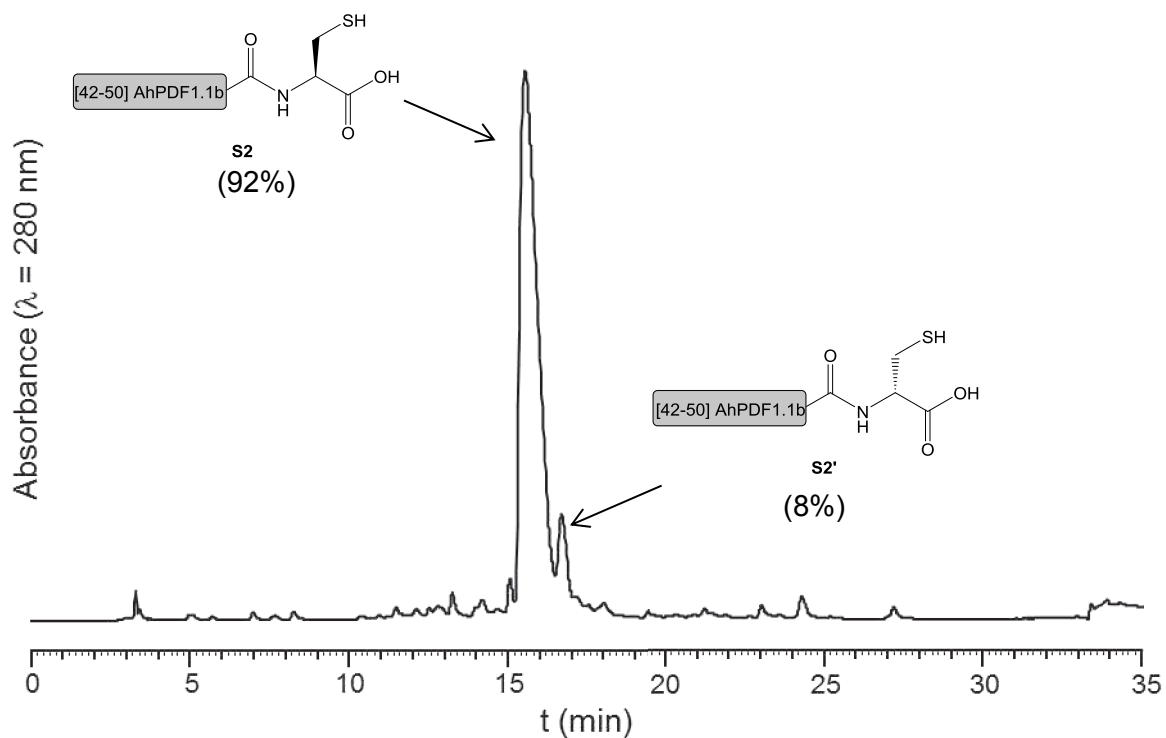
**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{53}H_{78}N_{13}O_{12}S_3$  : 1184.506, found: 1184.528.

**HPLC analysis**:  $t_R$  = 15.8 min (Nucleosil, gradient: 20-40% B/A over 30 min).

**S2'**

**MALDI-TOF** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{53}H_{78}N_{13}O_{12}S_3$  : 1184.506, found: 1184.518.

**HPLC analysis**:  $t_R$  = 16.7 min (Nucleosil, gradient: 20-40% B/A over 30 min).



Supplementary figure S1: HPLC trace of crude **S2** showing epimerized co-product **S2'**.

**S3**

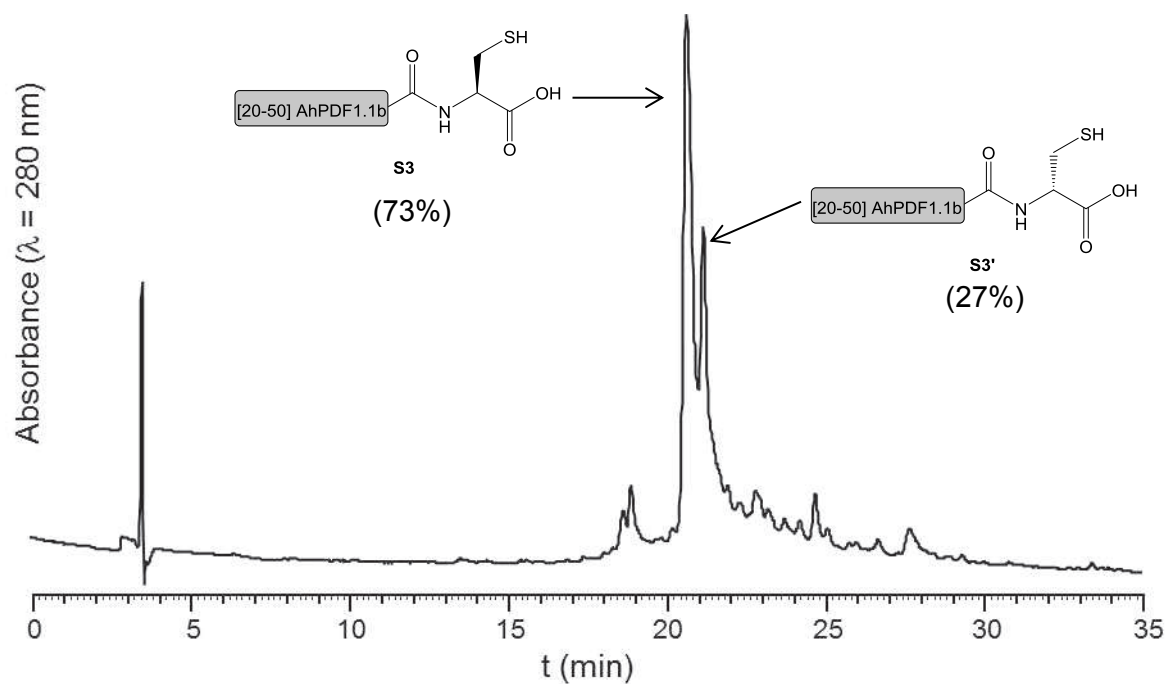
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{161}H_{249}N_{50}O_{41}S_6$ : 3730.7261, found: 3730.7292.

**HPLC analysis**:  $t_R$  = 20.8 min (Nucleosil, gradient: 20-40% B/A over 30 min).

**S3'**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{161}H_{249}N_{50}O_{41}S_6$ : 3730.7261, found: 3730.7289.

**HPLC analysis**:  $t_R$  = 21.2 min (Nucleosil, gradient: 20-40% B/A over 30 min).



Supplementary figure S2: HPLC trace of crude **S3** showing epimerized co-product **S3'**.

After completion of [1-51] AhPDF1.1b SPPS elongation, peptide-resin was deprotected and cleaved (protocol p S3) to afford crude peptide **1** together with 32% of an isomer attributed to the epimerized co-product **S1**. Traces of a product consistent with piperidinyllalanine formation were also observed. **1** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 28-38% B/A over 15 min) and was obtained with an overall yield of 7% (as compared to the initial loading of the H-Cys(Trt)-Trityl-ChemMatrix resin, and quantified by UV titration at 280 nm of pure **1**, based on an  $\epsilon^{280} = 8080 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$  in 8:2:0.01 H<sub>2</sub>O/MeCN/TFA).

**1**

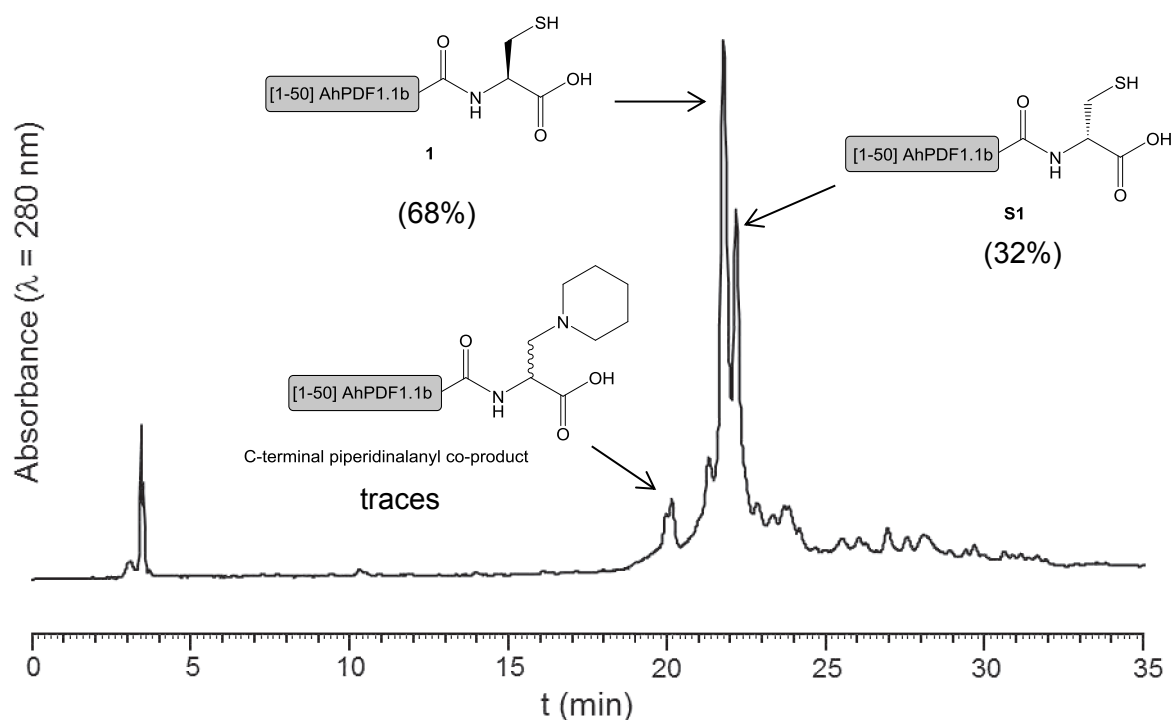
**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> calcd. for C<sub>242</sub>H<sub>373</sub>N<sub>76</sub>O<sub>68</sub>S<sub>8</sub>: 5687.5831, found: 5687.5851.

**HPLC analysis**:  $t_R = 21.8$  min (Nucleosil, gradient: 20-40% B/A over 30 min).

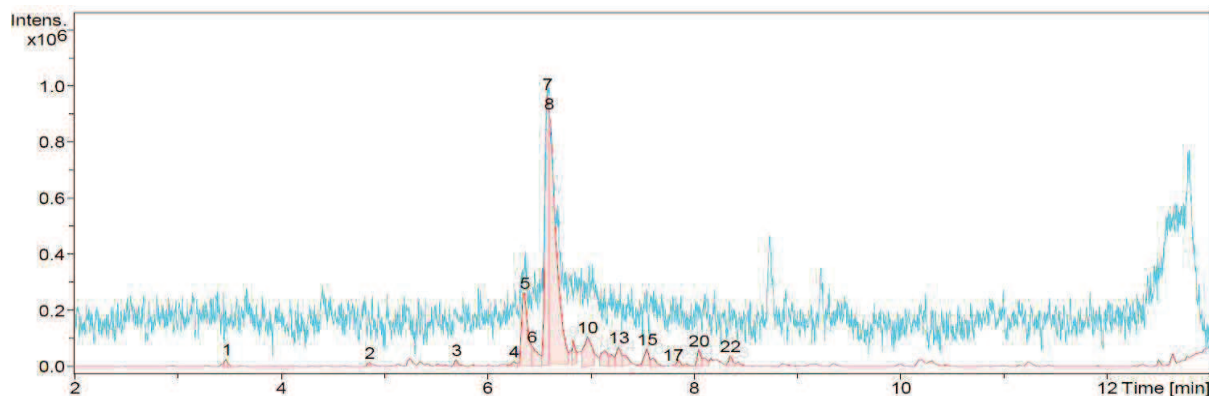
**S1**

**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> calcd. for C<sub>242</sub>H<sub>373</sub>N<sub>76</sub>O<sub>68</sub>S<sub>8</sub>: 5687.5831, found: 5687.5831.

**HPLC analysis**:  $t_R = 22.2$  min (Nucleosil, gradient: 20-40% B/A over 30 min).



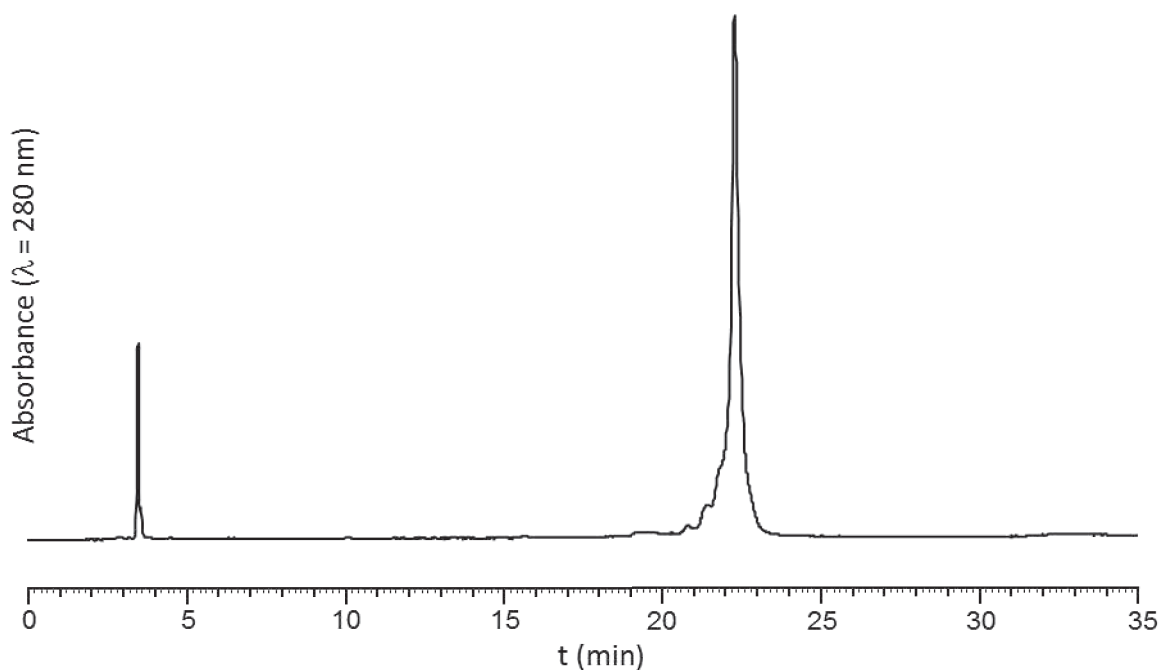
**Supplementary figure S3:** HPLC trace of crude **1** showing epimerized co-product **S1** and traces of a minor piperidinalanyl co-product.



**Supplementary figure S4:** LC/MS analysis of crude **1**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
5 (6.36)	5738.6845	5738.6878	<b>1</b> + 51 Da (C-terminal piperidinalanyl)
7 (6.58)	5687.5831	5687.5851	<b>1</b>
8 (6.60)	5687.5831	5687.5831	<b>S1</b> (C-terminally epimerized <b>1</b> )
10 (6.97)	5743.6457	5743.6471	<b>1</b> + 56 Da ( <i>t</i> Bu)
13 (7.27)	5743.6457	5743.6464	<b>1</b> + 56 Da ( <i>t</i> Bu)
15 (7.54)	2230.9335	2230.9317	Ac-[32-51] AhPDF1.1b
20 (8.05)	2093.8746	2093.8743	Ac-[33-51] AhPDF1.1b
22 (8.35)	1732.7690	1732.7676	Ac-[37-51] AhPDF1.1b

**Supplementary table S1:** Attribution of the main peaks observed during LC/MS analysis of crude [1-51] AhPDF1.1b **1**.

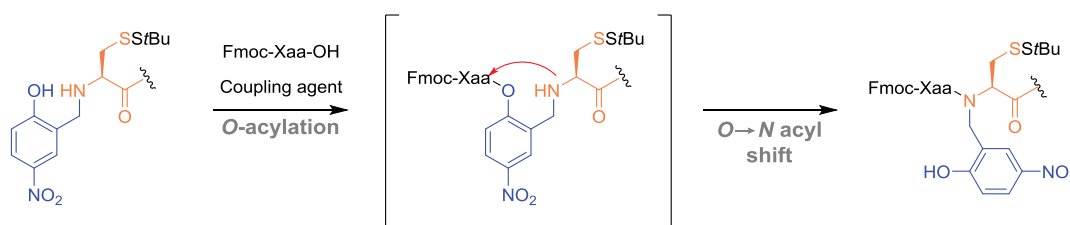


**Supplementary figure S5:** HPLC trace of purified **1** obtained *via* classical Fmoc SPPS.

## 4. Mechanistic studies on the *N*-acylation of the *N*-Hnb-Cys device with Fmoc-Pro-OH

### 4a- Preliminary results

Consistent with our previous observations,<sup>2</sup> the *N*-acylation of the *N*-Hnb-Cys device by the C-terminal Fmoc-Xaa-OH of the sequence is believed to be assisted by the hydroxyl group of Hnb, in a similar way as the mechanism commonly admitted for peptides containing an *N*-(2-hydroxy-5-methoxybenzyl) (Hmb)<sup>3</sup> or related 2-hydroxybenzyl groups.<sup>4</sup> The *N*-acylation thus probably proceeds in two steps, first *O*-acylation of the sterically non-hindered phenol or phenolate group of the device, followed by an intramolecular *O*→*N* acyl shift (supplementary scheme S2). In the case of sterically-hindered amino acids (Xaa = Val, Ile, Thr(O*t*Bu)) and Pro, the *N*-acylation yield obtained under standard SPPS coupling conditions are rather low, but can be dramatically enhanced by repeating the coupling step, or heating the reaction.



**Supplementary scheme S2:** Hypothetical mechanism for the *N*-acylation of the Hnb-Cys device by Fmoc-Xaa-OH.

Due to the inherent difficulties to directly monitor the solid-supported reaction, it is quite puzzling to determine whether the *O*-acylation or the *O*→*N* shift is the kinetically limiting step of the process. Considering previous studies by Alewood and co-workers on 2-hydroxy-5-nitrobenzyl or 2-hydroxy-6-nitrobenzyl groups as *N*-substituents of *N*-terminal amino acids<sup>4</sup> of peptide sequences, we initially speculated that the *O*→*N* shift should probably be fast, due to the electron-withdrawing properties of the nitro group, making Hnb a good leaving group.

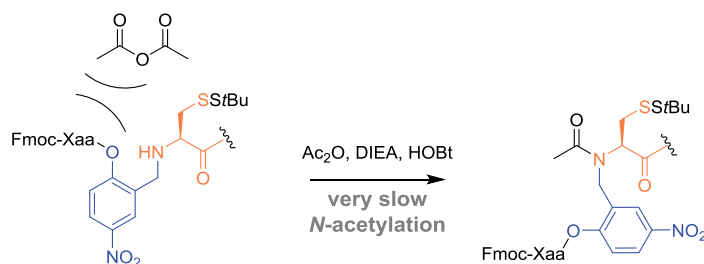
However during preliminary studies (not shown), we observed the acetic anhydride-mediated capping of unreacted secondary amine was unexpectedly difficult. Indeed, we found that very long (several hours) or repeated acetic anhydride treatments (up to tenfold) were necessary to achieve a quantitative capping with the Ac<sub>2</sub>O/HOBt/DIEA/DMF solution we standardly use for capping during SPPS. This finding is rather counterintuitive, considering the high reactivity and low steric hindrance of acetic anhydride.

<sup>2</sup> Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas A. F., Aucagne, V. *Chem. Sci.* **2016**, DOI: 10.1039/C5SC02630J.

<sup>3</sup> Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C. *J. Chem. Soc. Chem. Comm.* **1993**, 369–372.

<sup>4</sup> Miranda, L. P.; Meutermans, W. D.; Smythe, M. L.; Alewood, P. F. *J. Org. Chem.* **2000**, *65*, 5460–5468.

We hypothesized that the *N*-acetylation could be slowed down by the steric hindrance around the secondary amine arising from a non-rearranged *O*-acylated intermediary (supplementary scheme S3).



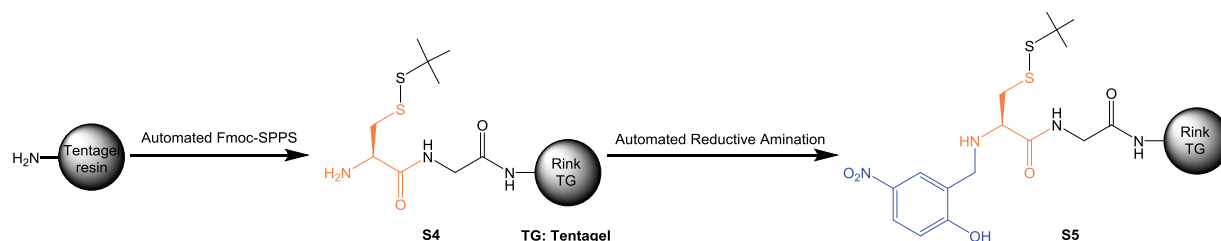
**Supplementary scheme S3:** Presence of non-rearranged *O*-acylated Hnb device as a possible explanation for the observed difficult *N*-acetylation of unreacted secondary amine of the Hnb-Cys device after coupling of a sterically-hindered Fmoc-Xaa-OH.

In order to gain further insights in the *N*-acylation mechanism, and help us to devise an efficient and automated capping protocol for the unreacted *N*-Hnb amine, we took the occasion of the herein described application to the very demanding Pro cryptothioester **2** for exploring this specific aspect.

Considering our hypotheses based on (1) slow *O*→*N* shift and (2) steric hindrance of the secondary amine by the non-rearranged *O*-acylated Hnb, we thought of performing a few simple indirect experiments to confirm our speculations: (1) washing the resin after a 30 min standard coupling, in order to remove the excess of Fmoc-Pro activated ester, and further incubate the peptide-resin under similar conditions as the one used for the coupling but omitting Fmoc-Pro and coupling reagent (DIEA, NMP), and (2) finding conditions to selectively cleave the Hnb ester, in order to probe the expected fast kinetics of *N*-acetylation of a non-*O*-acylated *N*-Hnb-Cys.

#### 4b- Synthesis of *N*-Hnb-Cys resin **S5**

To perform this mechanistic studies, resin **S5** equipped with our *N*-Hnb-Cys transesterification device was first synthesized under the conditions we earlier reported (supplementary scheme S4).<sup>2</sup>



Supplementary scheme S4: Synthesis of peptidyl resin **S5**.

The Rink's linker, Gly and Cys(SfBu) were successively coupled by automated SPPS (general procedure p S3) on a Tentagel R resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol) in order to obtain peptide-resin **S4**.

#### Reductive amination:

Peptide-resin **S4** (25  $\mu$ mol) was washed two times with 3 mL of a 1:1 DMF/MeOH mixture for 30 s, then swollen in 3 mL of a 9:9:2 DMF/MeOH/AcOH mixture for 5 min. The reactor was drained off and the resin was washed four times with 3 mL of a 1:1 DMF/MeOH mixture for 30 s. This process forms the acetic acid salt of the amine group of cysteine.

2-Hydroxy-5-nitrobenzaldehyde (HNBA) in 1:1 DMF/MeOH (125 mM, 10 equiv., 2 mL) was then added, and the reactor was left for 1 h under stirring by nitrogen bubbling. The reactor was drained off and the resin was washed four times with 3 mL of 1:1 DMF/MeOH for 15 s.

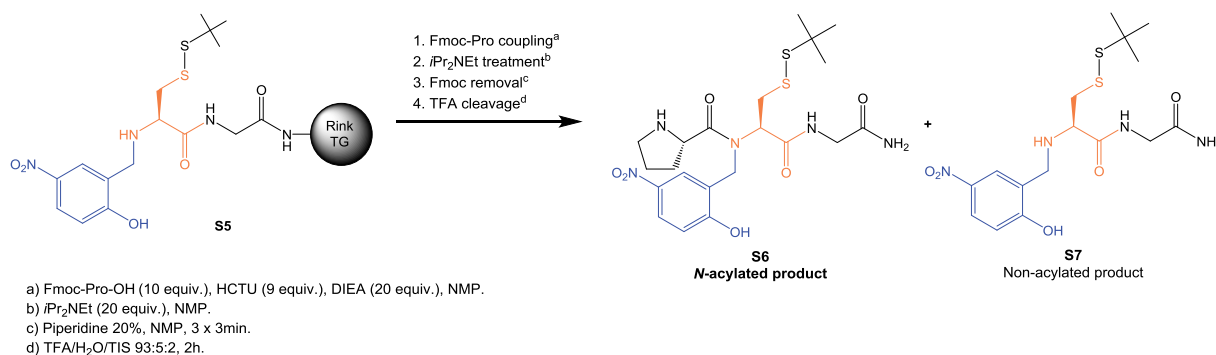
Without delay, a fresh solution of sodium cyanoborohydride in 9:9:2 DMF/MeOH/AcOH (250 mM, 20 equiv., 2 mL) were added and the reactor was left for 1 h under stirring by nitrogen bubbling. The reactor was drained off and the resin was washed with 1:1 DMF/MeOH (3 mL, 30 s,  $\times$  4), NMP (3 mL, 30 s,  $\times$  3), 20% piperidine in NMP (3 mL, 30 s,  $\times$  3), NMP (3 mL, 30 s,  $\times$  3), dichloromethane (5 mL, 30 s,  $\times$  3) and NMP (3 mL, 30 s,  $\times$  2).



#### 4c- Study of the intramolecular $N \rightarrow O$ shift kinetics

In order to study the mechanism of the  $N$ -acylation of  $N$ -Hnb-Cys device, Fmoc-Pro-OH was coupled on peptide-resin **S5** using standard automated coupling on the Prelude synthesizer (protocol p S3), for either 30 min, 3 × 30 min or 90 min. Without delay, peptidyl resins were treated with piperidine, in order to remove the Fmoc group while also cleaving any eventual  $O$ -acylated intermediary. Alternatively, before the piperidine treatment, peptide-resin was incubated with 20 equiv.  $iPr_2NEt$  in NMP for 24 h.

This was followed by a 2 h TFA treatment (TFA/H<sub>2</sub>O/ $iPr_3SiH$  93:5:2) to cleave the Rink's amide linker, leading to a mixture of  $N$ -acylated H-Pro-(Hnb)Cys(*StBu*)-Gly-NH<sub>2</sub> (**S6**) and non-acylated H-(Hnb)Cys(*StBu*)-Gly-NH<sub>2</sub> (**S7**) (scheme S5). After evaporation of TFA, the samples were analyzed by HPLC; the  $N$ -acylation yields were quantified by integration at 315 nm, not taking into account eventual differences of molar extinction coefficients between **S6** and **S7** (table S2).



#### Supplementary scheme S5: Studies of the coupling of Fmoc-Pro-OH on **S5**.

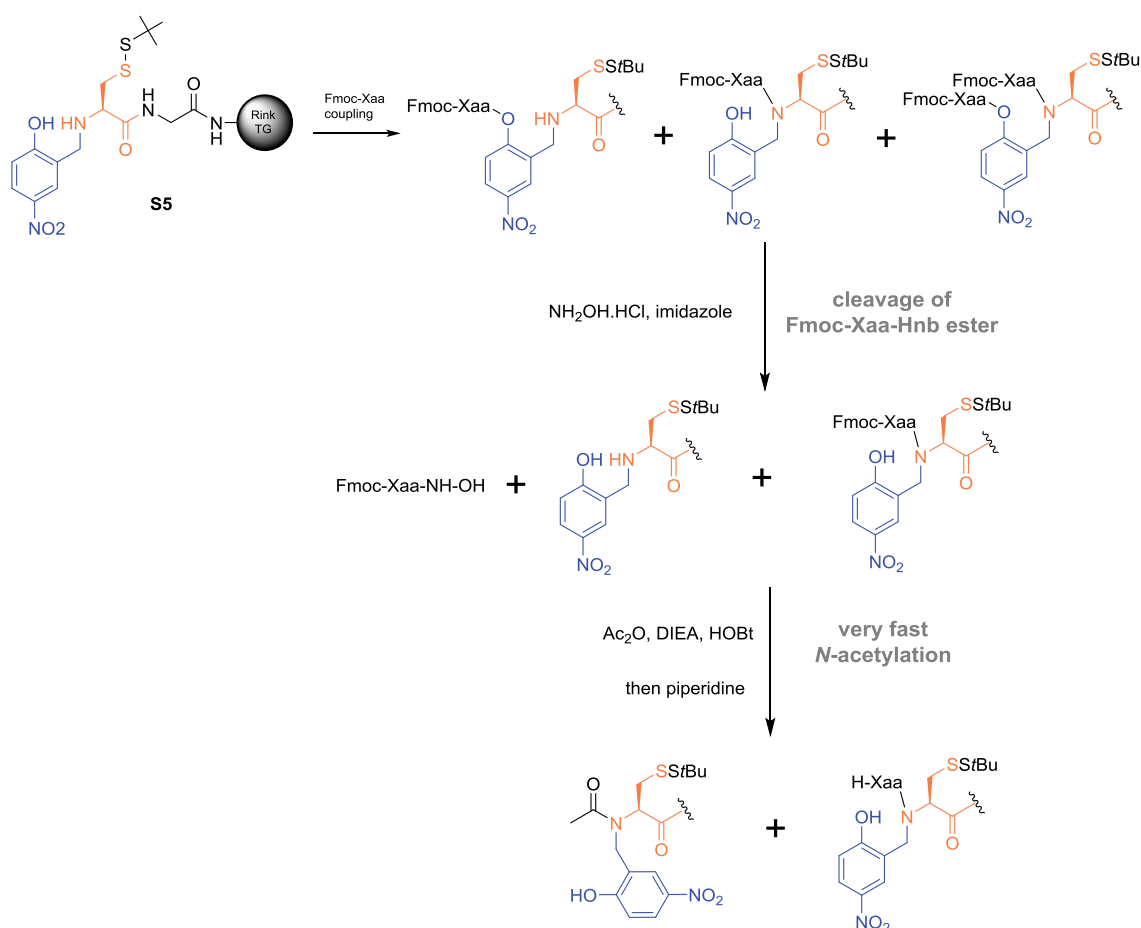
Coupling time (h)	0.5	0.5 × 3	1.5	0.5
$iPr_2NEt$ treatment time (h)	0	0	0	24
$N$ -acylation (%)	16	32	33	61

Supplementary table S2:  $N$ -acylation yield of peptide-resin **S5** with coupling of Fmoc-Pro-OH, eventually followed by a 24 h incubation with  $iPr_2NEt$  in NMP.

These results are in perfect accordance with a slow  $O \rightarrow N$  shift hypothesis: indeed, a 24 h hour incubation after thorough wash of the resin after a 30 min coupling led to a good 61% yield, to be compared with a 16% yield with no extra treatment with  $iPr_2NEt$  in NMP. Note that the  $O$ -acylation is probably relatively slow, and not quantitative after 30 min: threefold coupling for 30 min or 90 min coupling doubled the  $N$ -acylation yield.

#### 4d- Selective cleavage of the Hnb esters

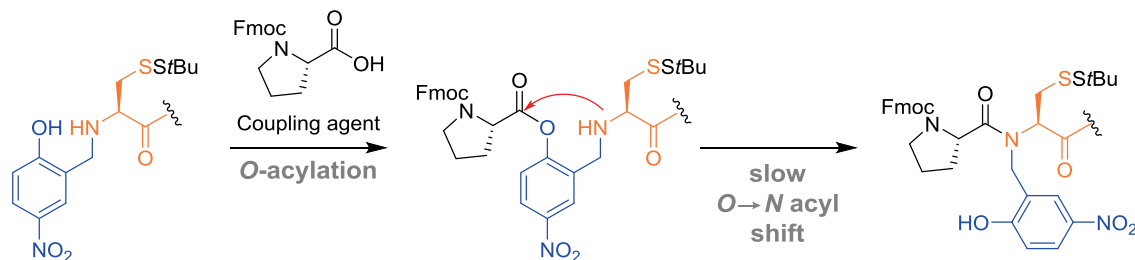
We thought that a simple way to selectively cleave the Hnb esters without affecting the integrity of the Fmoc protective group could rely on a treatment with a nucleophile under moderately basic conditions. Inspired by conditions developed by Bradley *et al.* for the selective nucleophilic cleavage of a Dde group versus a *N*-Fmoc group,<sup>5</sup> we found that a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in a NMP/CH<sub>2</sub>Cl<sub>2</sub> mixture (5:1) cleanly cleaved any Hnb ester present (supplementary scheme S6). Conveniently, this process could be monitored by UV spectrometry of the Fmoc-Pro-NHOH product at 301 nm. We found that two 20 min treatments resulted in complete cleavage of Hnb esters. Importantly, we found that performing the capping step after this treatment resulted in very fast *N*-acetylation of the non-acylated secondary amines, supporting our initial hypothesis of slow capping caused by the steric hindrance of a non-rearranged Hnb ester.



**Supplementary scheme S6:** Nucleophile-mediated cleavage of a putative O-Hnb ester prior to the *N*-acetylation of the Hnb-Cys device after coupling of a Fmoc-Xaa-OH.

<sup>5</sup> Diaz-Mochon, J. J.; Bialy, L.; Bradley, M. *Org. Lett.*, **2004**, *6*, 1127-1129.

All together, these results strongly supports that, at least in the case of Pro, the *N*-acylation process is limited by a slow *O*→*N* acyl shift (see supplementary scheme S7).



### S6

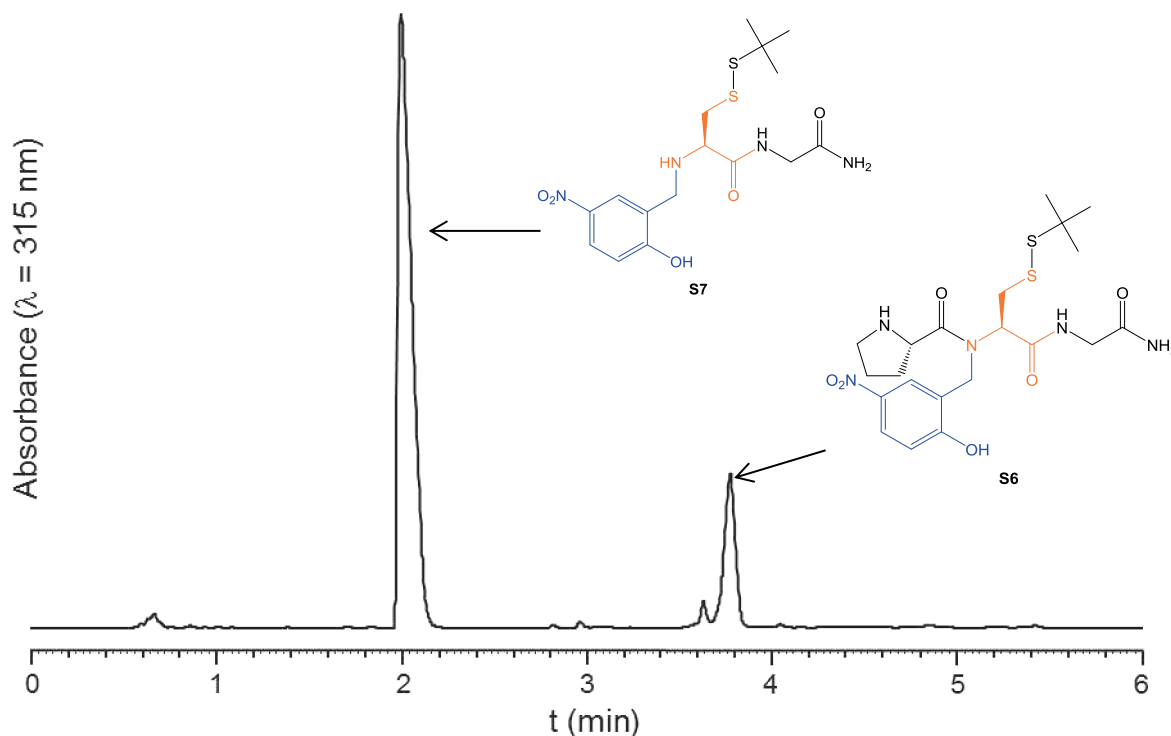
**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{21}H_{32}N_5O_6S_2$ : 514.1794, found: 514.1788.

**HPLC analysis**:  $t_R = 3.86$  min (Chromolith, gradient: 20-50% B/A over 5 min).

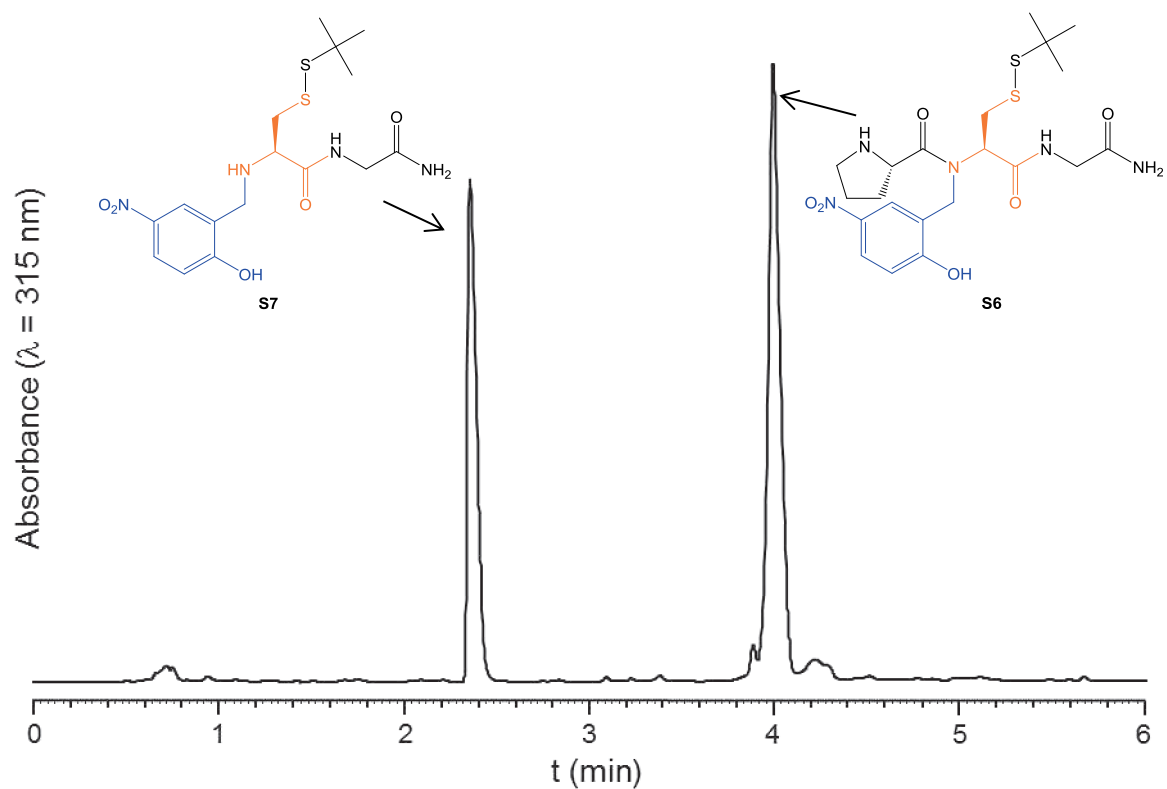
### Non acylated product - S7

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{16}H_{25}N_4O_5S_2$ : 417.1266, found: 417.1261.

**HPLC analysis**:  $t_R = 2.12$  min (Chromolith, gradient: 20-50% B/A over 5 min).



Supplementary figure S6: HPLC trace of crude **S6** after a single standard coupling of Fmoc-Pro-OH (30 min) and without *i*Pr<sub>2</sub>NEt treatment.



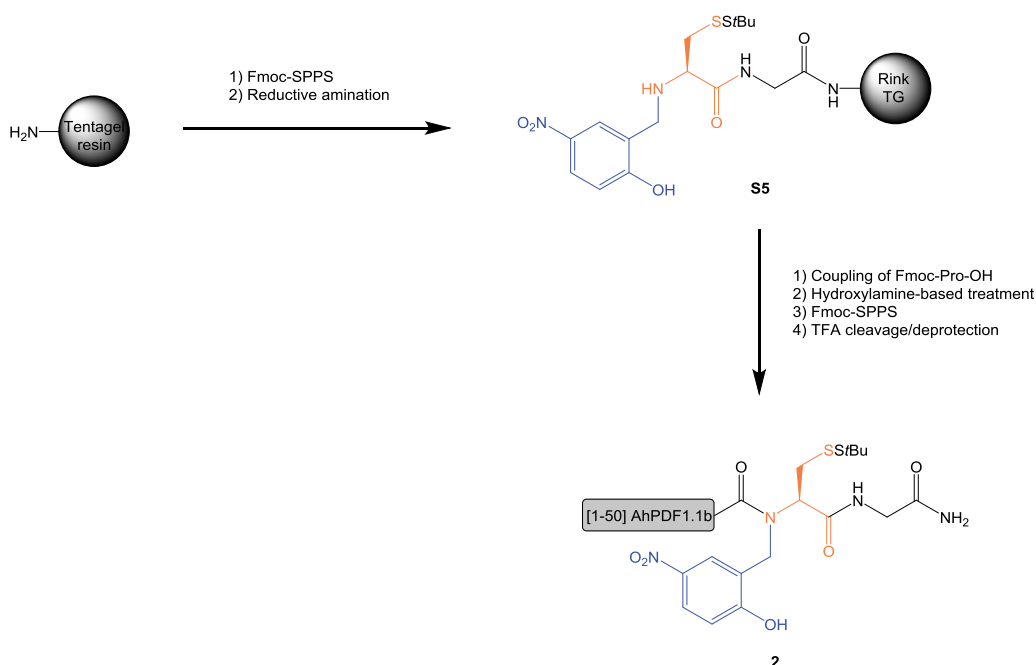
**Supplementary figure S7:** HPLC trace of crude **S6** after a single standard coupling of Fmoc-Pro-OH (30 min) followed by *i*Pr<sub>2</sub>NEt treatment (20 equiv., 24 h).

## 5. Synthesis of [1-50] AhPDF1.1b-(Hnb)C(SfBu)G-NH<sub>2</sub> (**2**)

Amino acid sequence of **2**:

ZRLCEKPSGTWSGVCGNNGACRNQCIRLEKARHGSCNYVFPAHKCICYFP(Hnb)C(SfBu)G-NH<sub>2</sub>

Z = pyroglutamic acid



### Supplementary scheme S8: Synthesis of **2**.

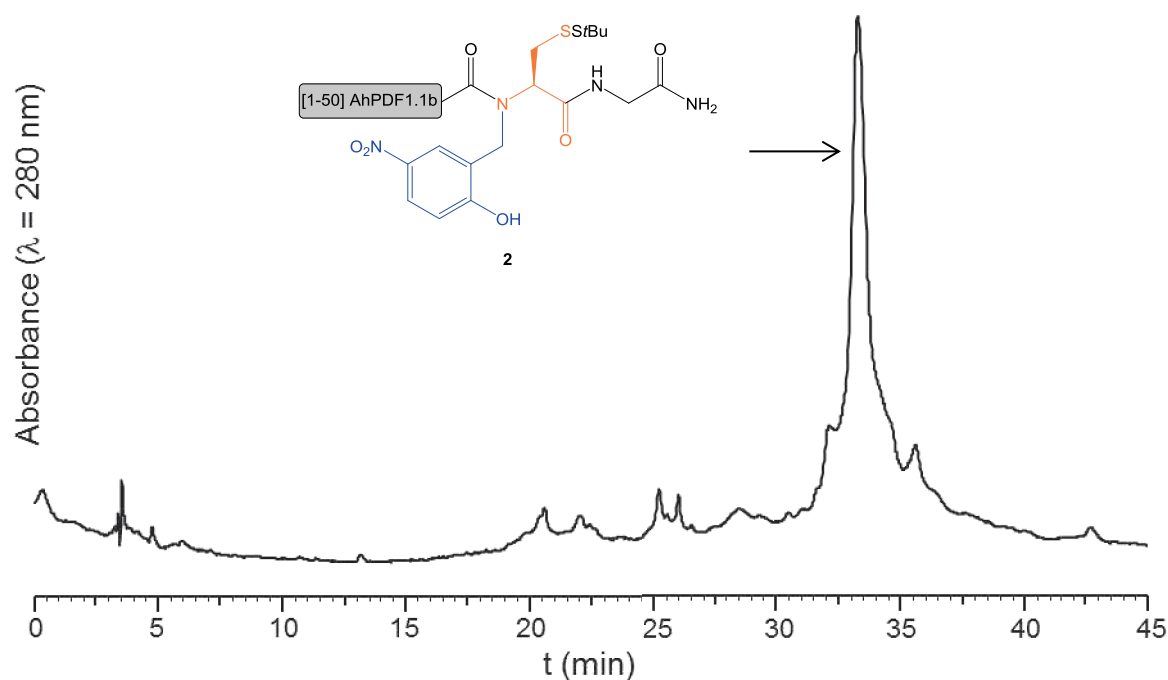
Peptide **2** was obtained through automated SPPS using a Prelude synthesizer (protocol p S3) and using the automated reductive amination (protocol p S11) starting from aminomethyl Tentagel resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol). Fmoc-Pro-OH was coupled five times (see general procedure pS3) on the *N*-Hnb-Cys device and a treatment of the peptide-resin using a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in a NMP / CH<sub>2</sub>Cl<sub>2</sub> mixture (5:1, 20 min, 3 mL,  $\times$  3) was performed. A capping step was performed (protocol p S3) then the *N*-acylation yield (50%) was quantified by UV titration of the fluorenylmethylpiperidine byproduct from the Fmoc removal at  $\lambda = 301$  nm ( $\epsilon^{301} = 7800$  L $\cdot$ mol<sup>-1</sup> $\cdot$ cm<sup>-1</sup>). The peptide sequence was elongated using standard SPPS (protocol p S3), a double coupling being performed for the last residue (pyroglutamic acid), before a final piperidine treatment. Elongation yield (55%) was determined by UV titration of the fluorenylmethylpiperidine byproduct from the last Fmoc removal (Arg<sup>2</sup>), as compared to the first one (Pro<sup>50</sup>). For this purpose, after the coupling of Arg<sup>2</sup> and prior to the piperidine treatment, peptide-resin was treated with a solution of hydroxylamine hydrochloride and imidazole as above described, in order to cleave any Fmoc-Arg(Pbf) ester on the Hnb moiety. After cleavage and deprotection (protocol p S3), a part of crude **2** was purified by semi-preparative RP-HPLC at 70 °C (Nucleosil, gradient: 38-50% B/A over 20 min). Pure peptide **2** was obtained in an overall yield of 15% (quantity of pure **2** determined by UV titration at 280 nm, yield based on loading

of Fmoc-P-(Hnb)C(S*t*Bu)-G-Rink-Tentagel resin and thus corresponding Fmoc SPSS elongation, TFA-mediated cleavage and deprotection, and purification).

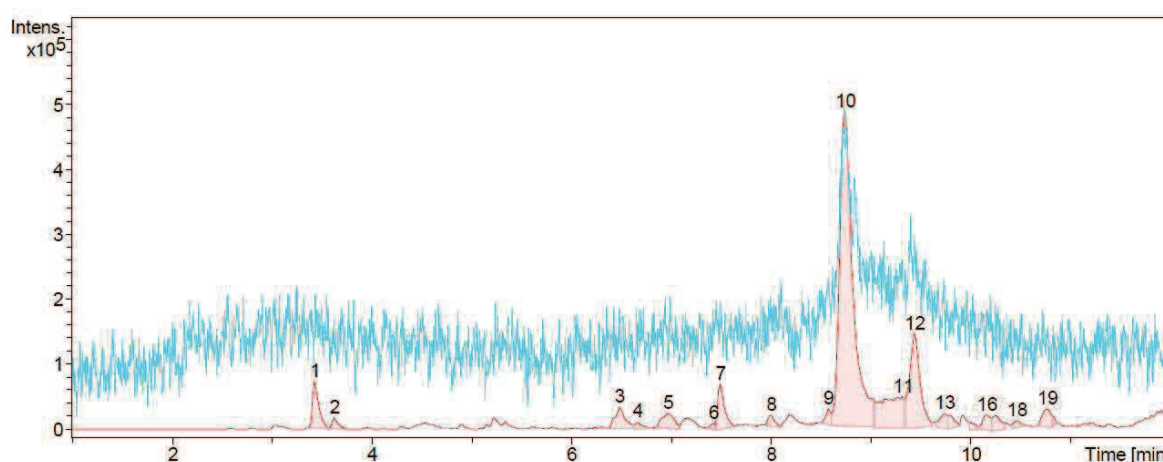
**2**

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{255}H_{390}N_{79}O_{71}S_9$ : 5982.6822, found: 5982.6806.

**HPLC analysis**:  $t_R$  = 33.7 min (Nucleosil, gradient: 20-54% B/A over 45 min, 70 °C).



Supplementary figure S8: HPLC trace (Nucleosil, 70 °C) of crude **2**.



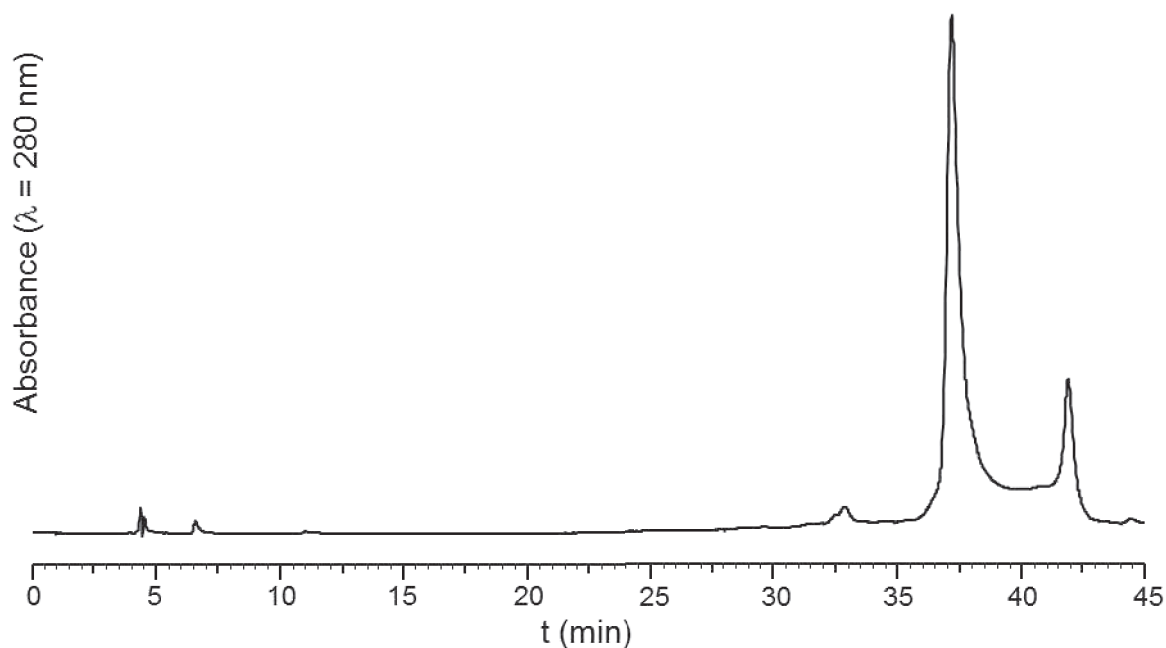
Supplementary figure S9: LC/MS analysis of crude **2**. Blue trace: UV ( $\lambda$  = 214 nm); red trace: base peak ion chromatogram.

Note that the LC/MS analysis of crude **2** showed an unusual profile for the peak corresponding to compound **2**: a first large peak (#10) followed by a smaller one

(#12) separated by a plateau (#11). Such a chromatographic profile is consistent with a conformational equilibrium,<sup>6</sup> that we attributed to a slow *trans/cis* interconversion of the Pro-(Hnb)Cys *N,N*-disubstituted amide bond within the timeframe of the analysis. This hypothesis is relevant with the same peak-plateau-peak profile found in HPLC analysis at room temperature of purified **2** (figure S10), that clearly show a coalescence between the two peaks at higher temperature (figure S11).

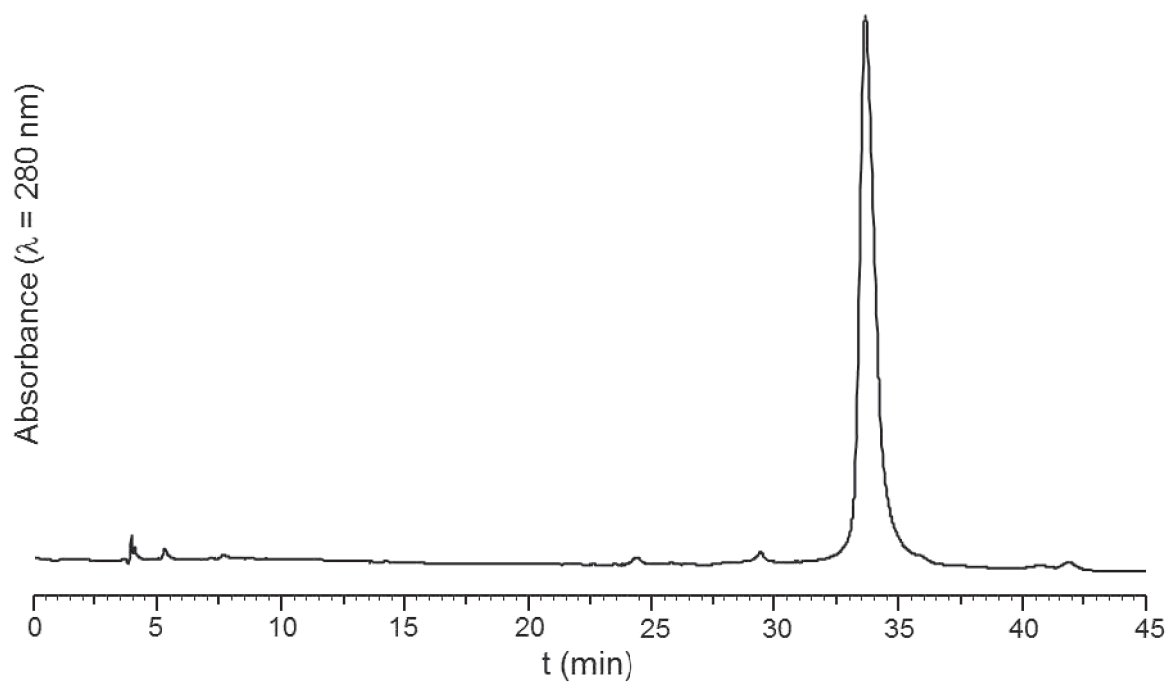
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
7 (7.49)	-	5831.6519	Not attributed
10 (8.74)	5982.6822	5982.6806	<b>2</b>
12 (9.43)	5982.6822	5982.6809	Conformer of <b>2</b>

Supplementary table S3: Attribution of the main peaks observed during LC/MS analysis of crude **2**.



Supplementary figure S10: HPLC trace of purified **2**, 20 °C.

<sup>6</sup> See for example: a) Oxelbark, J.; Allenmark, S. *J. Org. Chem.* **1999**, *64*, 1483–1486. b) Shibukawa, Masami; Miyake, A.; Eda, S.; Saito, S. *Anal. Chem.* **2015**, *87*, 9280–9287.

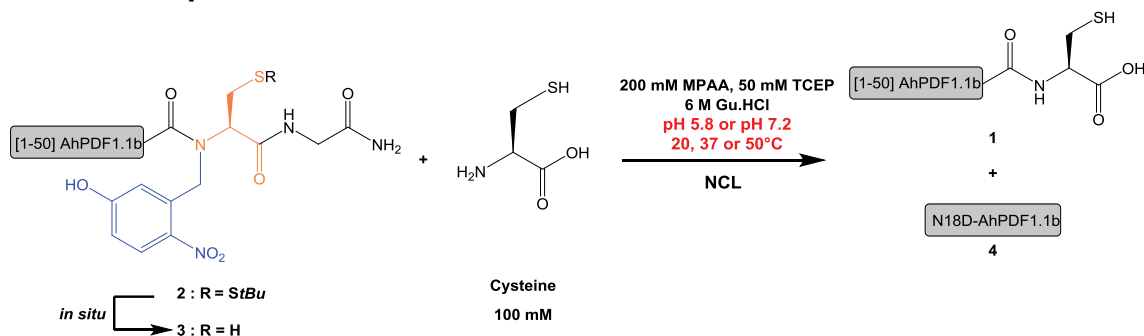


Supplementary figure S11: HPLC trace of purified **2**, 70 °C.



## 6. Synthesis of [1-51] AhPDF1.1b (1) via NCL

### 6a- Optimization of the NCL conditions



Supplementary scheme S9: NCL using HPLC-purified 2 to obtain 1.

#### NCL conditions optimization (see table 1 of the article body):

80  $\mu\text{L}$  of a deoxygenated<sup>a</sup> 0.2 M (pH = either 5.8 or 7.2, see table 1) sodium phosphate buffer containing 200 mM 4-mercaptophenylacetic acid (MPAA), 50 mM *tris*-carboxyethylphosphine (TCEP), 6 M guanidine hydrochloride (Gu.HCl) and 100 mM cysteine was added to HPLC-purified peptide **2** (0.08  $\mu\text{mol}^{\text{b}}$ ) (final concentration 1 mM), under argon. Ligations were carried out at either 20, 37 or 50 °C and monitored by RP-HPLC. For this, 5  $\mu\text{L}$  aliquots of the reaction mixture were diluted with 100  $\mu\text{L}$  of 2% TFA in water and injected in HPLC. Yields were quantified by integrations of the HPLC peaks ( $\lambda = 280 \text{ nm}$ ) corresponding to compounds **3** and **1/4**, taking into account the absorbance of the N-Hnb-Cys device ( $\epsilon^{280} \text{ Hnb} = 2360 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ,  $\epsilon^{280} \text{ 3} = 10440 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ,  $\epsilon^{280} \text{ 1 and 4} = 8080 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$  in 8:2:0.01  $\text{H}_2\text{O}/\text{MeCN}/\text{TFA}$ ). The relative amount of **1** and **4** –not well separated under the analytical HPLC conditions used for monitoring NCL– was determined by integration of the extracted ion chromatogram from the LC/MS analyses corresponding to both compounds, neglecting the eventual differences in terms of ionization of the two compounds.

Reaction at pH = 5.8 and 50 °C for 48 h were found to be optimal conditions for this reaction (see Table 1 of the article).

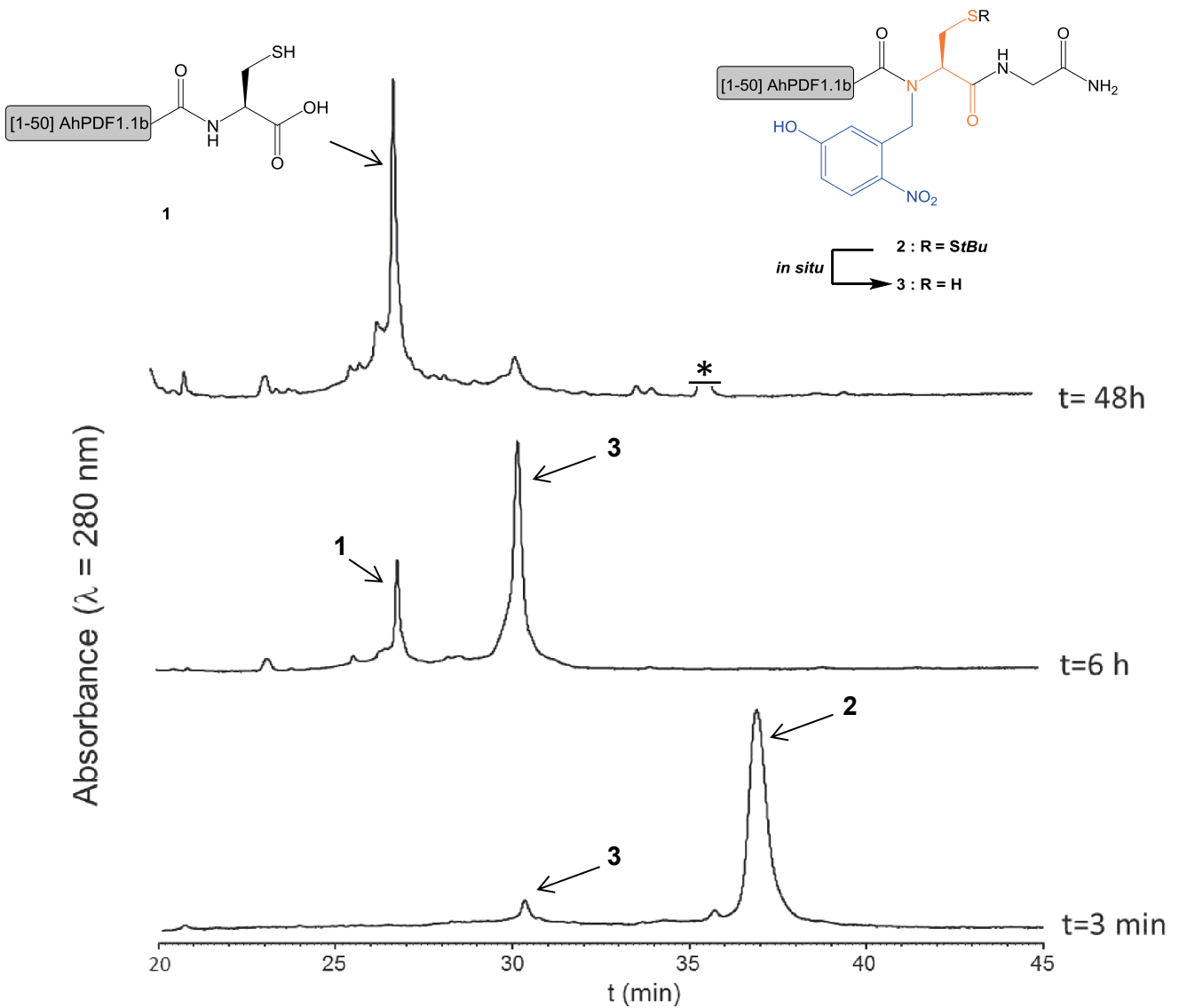
<sup>a</sup>: Deoxygenation was performed through four consecutive vacuum/argon cycles.

<sup>b</sup>: Quantified by UV titration at 280 nm of pure **2**.

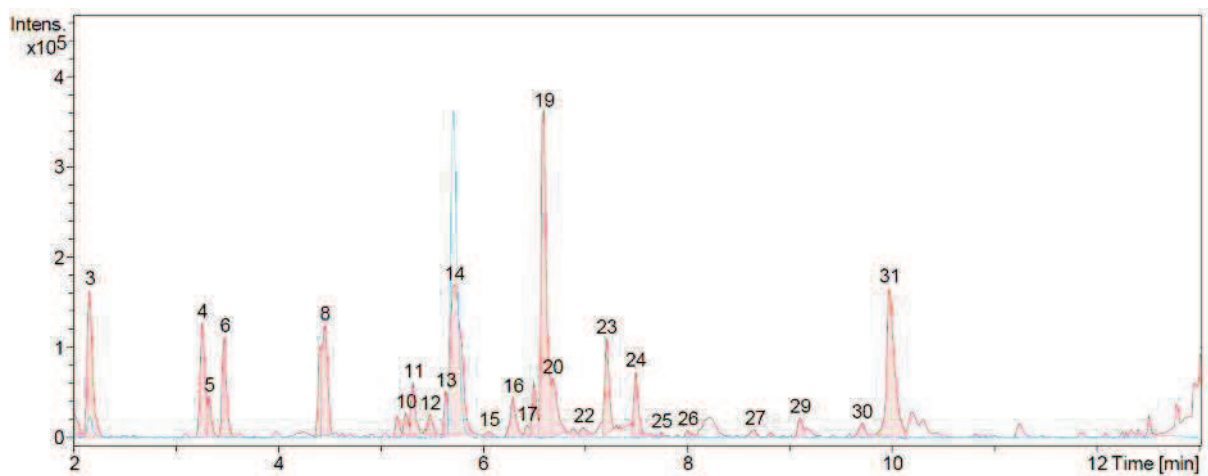
1

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd. for  $\text{C}_{242}\text{H}_{373}\text{N}_{76}\text{O}_{68}\text{S}_8$ : 5687.5831, found: 5687.5864.

**HPLC analysis**:  $t_{\text{R}} = 26.9 \text{ min}$  (Nucleosil, gradient: 10-50% B/A over 40 min, 70 °C).



Supplementary figure S12: Analytical HPLC monitoring at 70 °C of the NCL reaction at pH 5.8 at 50 °C with purified **2** to obtain peptide **1**. \*: MPAA disulfide



Supplementary figure S13: LC/MS analysis of NCL to obtain **1** after 48h at pH 5.8 at 50 °C. Blue trace: UV (λ = 214 nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
3 (2.16)	-	409.1870	Not attributed
4 (3.26)	-	373.0867	Not attributed
6 (3.48)	-	502.2439	Not attributed
8 (4.46)	-	418.9847	Not attributed
14 (5.72)	289.0357 <sup>a</sup>	289.0344 <sup>a</sup>	MPAA <sup>a</sup>
16 (6.29)	3659.6889	3659.6897	H-[21-51] AhPDF1.1b <sup>b</sup>
18 (6.50)	3158.4771	3158.4790	H-[25-51] AhPDF1.1b <sup>b</sup>
18 (6.50)	5584.5740	5584.5595	[1-50] AhPDF1.1b <sup>c</sup>
19 (6.60)	5687.5831	5687.5864	<b>1</b>
20 (6.68)	5688.5909	5688.5874	<b>4</b> <sup>d</sup>
20 (6.68)	5670.5561	5670.5652	<b>1</b> -17 Da <sup>e</sup>
23 (7.20)	848.3145	848.3128	H-[45-51] AhPDF1.1b <sup>b</sup>
24 (7.49)	1907.8105	1907.8092	H-[35-51] AhPDF1.1b <sup>b</sup>
31 (9.97)	352.0677 <sup>f</sup>	352.0661 <sup>f</sup>	MPAA disulfide

**Supplementary table S4:** Attribution of the main peaks observed during LC/MS analysis of NCL to obtain **1** after 48h at pH 5.8 at 50°C.

<sup>a</sup>: Attribution based on the characteristic UV spectrum of this compound, and on the corresponding HPLC chromatogram showing a peak at the expected retention time for MPAA. Note that under our MS conditions, no peaks corresponding to the expected  $m/z$  for protonated MPAA neither cation adducts were observed. A possible explanation for the observed  $m/z$  could be oxidation into disulfide then decarboxylation to yield a benzyl cation ( $C_{15}H_{13}O_2S_2$ , calcd.  $m/z$  289.0357). Note that minor peaks corresponding to  $M+Na^+$  and  $M+NH_4^+$  relative to the disulfide were also observed.

<sup>b</sup>: Traces of non-acetylated truncated peptides containing a N-terminal Cys residue were observed in the LC/MS analysis of the mixture. These co-products were attributed to *in situ* N→S shift at Xaa-Cys sites followed by trans-thioesterification, as reported by MacMillan.<sup>7</sup> Note that the ion chromatogram signal is not representative of the actual proportion of these peptides, as compared with the HPLC-UV analysis. Smaller peptides are expected to lead to a higher BPC signal due to greater ionization.

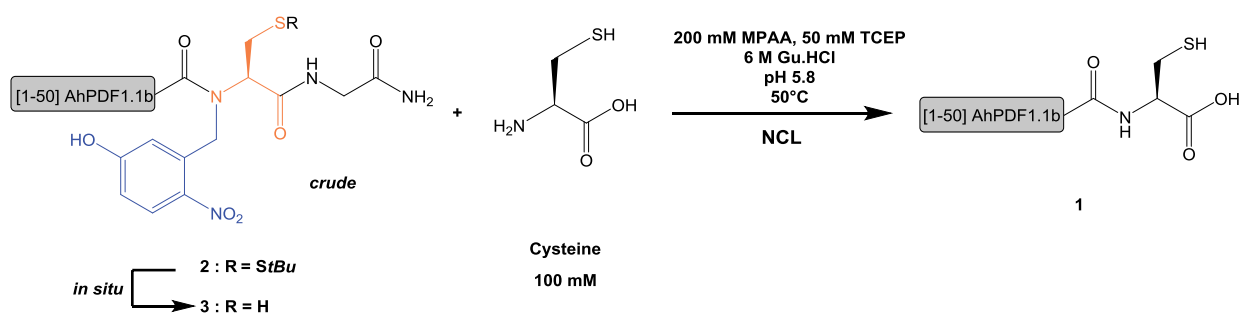
<sup>c</sup>: [1-50] AhPDF1.1b corresponds to a co-product arising from hydrolysis of peptide crypto-thioester **3**. [1-50] AhPDF1.1b proportion was evaluated to 6% relative to **1** by integration of the corresponding extracted ion chromatograms, neglecting the eventual differences in terms of ionization of the two compounds.

<sup>d</sup>: Compound **4** proportion was evaluated to 13% relative to **1** by integration of the extracted ion chromatogram corresponding to both **1** and **4**, neglecting the eventual differences in terms of ionization of the two compounds.

<sup>e</sup>: The proportion of this compound, attributed to aspartimide formation at the N<sup>18</sup>G site, was evaluated to 4.6% relative to **1** by integration of the corresponding extracted ion chromatograms, neglecting the eventual differences in terms of ionization.

<sup>f</sup>: The calculated  $m/z$  value corresponds to a  $[M+NH_4]^+$  form ( $C_{16}H_{18}NO_4S_2$ ). A  $[M+Na]^+$  species is also observed. A minor peak consistent with decarboxylation to yield a benzyl cation (see note a) is also observed.

<sup>7</sup> Cowper, B.; Shariff, L.; Chen, W.; Gibson, S. M.; Diab, W.-L.; Macmillan, D. *Org. Biomol. Chem.*, **2015**, *13*, 7469–7476.

6b- Application to crude crypto-thioester **2**Supplementary scheme S10: NCL using crude **2** to obtain **1**.

## NCL under optimized conditions:

800  $\mu\text{L}$  of a deoxygenated<sup>a</sup> 0.2 M pH 5.8 sodium phosphate buffer containing 200 mM MPAA, 50 mM TCEP, 6 M guanidine hydrochloride and 100 mM cysteine was added to crude dry peptide **2** (8 mg, corresponding to 1.3  $\mu\text{mol}$  of the crude crypto-thioester mixtures,<sup>b</sup> and 0.72  $\mu\text{mol}$  of **2**<sup>c</sup>) under argon. Ligation was carried out at 50 °C and monitored by RP-HPLC. After 48h, TCEP (100 mM final concentration) was added to the reaction mixture and pH was adjusted to 6.0 using a 10 M NaOH solution to reduce disulfides possibly formed due to oxygen contamination. After 1 h, pH was adjusted to 1 with 1.5% TFA in water and ligation product **1** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 30-36% B/A over 12 min). **1** was isolated with an overall yield of 13.4% (based on the loading of Fmoc-Pro-(Hnb)C(StBu)-G-Rink-Tentagel resin, and thus corresponding Fmoc SPPS elongation, TFA-mediated cleavage and deprotection, NCL and purification).

<sup>a</sup>: deoxygenation was performed through four consecutive vacuum/argon cycles.

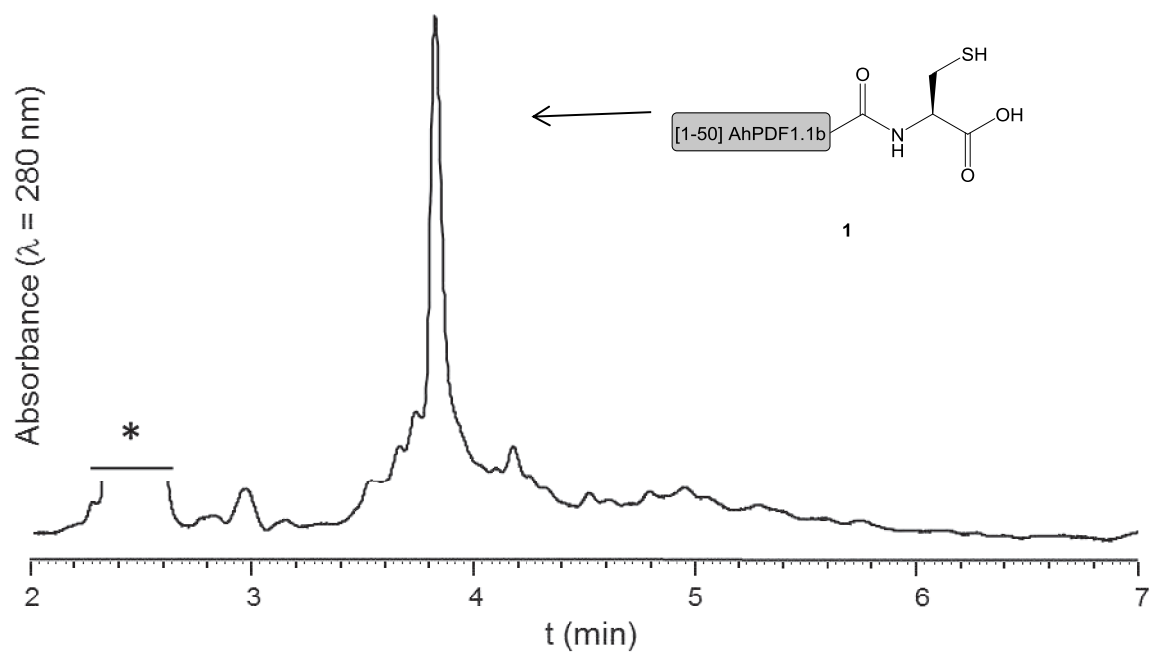
<sup>b</sup>: Quantity of crude peptide crypto-thioester mixture evaluated from the Fmoc-Pro-(Hnb)C(StBu)-G-Rink-Tentagel resin loading, as determined by UV titration of the fluorenylmethylpiperidine byproduct from the Fmoc removal ( $\epsilon^{301} = 7800 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ).

<sup>c</sup>: As evaluated from the SPPS elongation yield (55%).

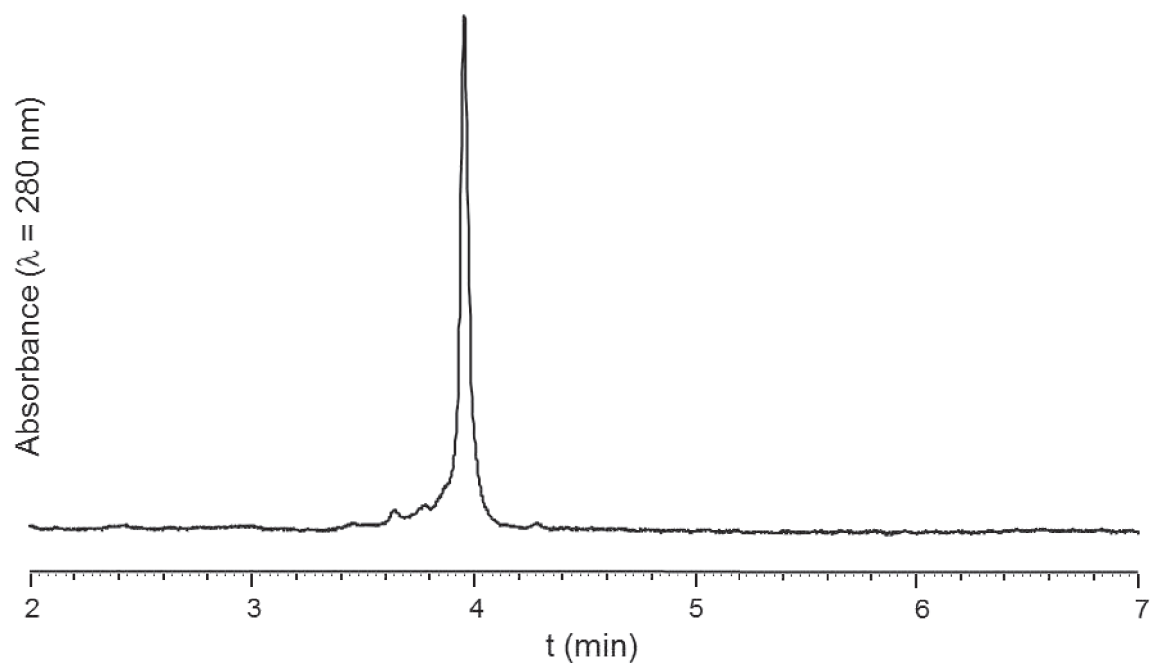
**1**

**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd. for  $\text{C}_{242}\text{H}_{373}\text{N}_7\text{O}_{68}\text{S}_8$ : 5687.5831 found: 5687.5751.

**HPLC analysis**:  $t_R = 3.9$  min (Chromolith, gradient: 20-50% B/A over 9 min).



Supplementary figure S14: HPLC trace of crude reduced **1** after 42 h of NCL. (\*: MPAA).



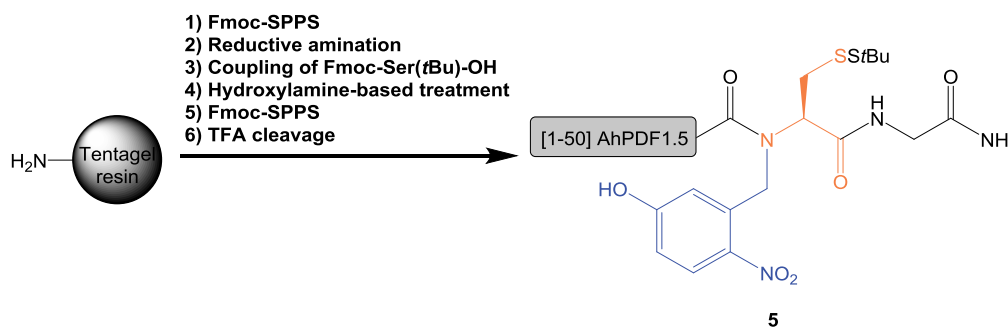
Supplementary figure S15: HPLC trace of purified **1**.

## 7. Application to the synthesis of [1-51] AhPDF1.5 (6)

### 7a- Synthesis of [1-50] AhPDF1.5-(Hnb)C(S*t*Bu)G-NH<sub>2</sub> (5)

Amino acid sequence of **5**:

ZLCKRESETWSGRCVNDYQCRDHCINNDRGNDGYCAGGYPWYRGCFCCFFS(Hnb)C(S*t*Bu)G-NH<sub>2</sub>  
 Z = pyroglutamic acid



Supplementary scheme S11: Synthesis of **5**.

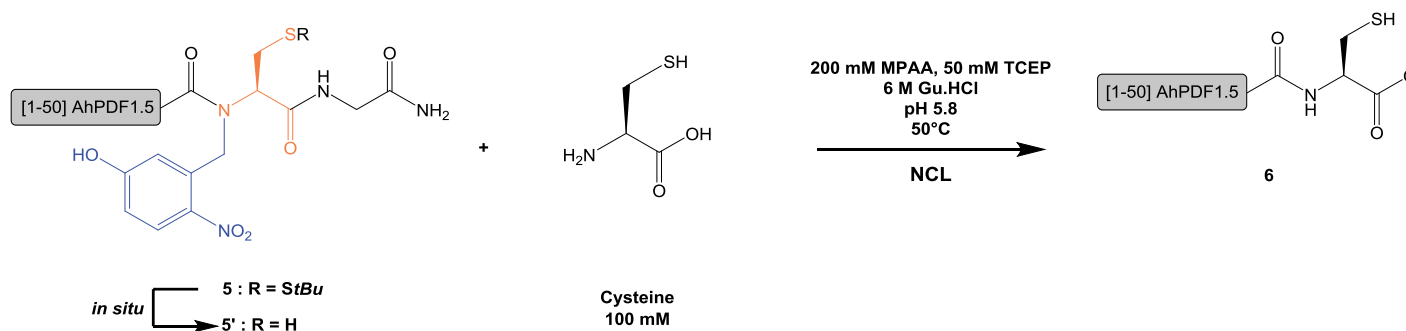
Peptide **6** was obtained through automated SPPS using a Prelude synthesizer (protocol p S3) and using the automated reductive amination (protocol p S11) starting from aminomethyl Tentagel resin (120 mg, 0.21 mmol/g, 25  $\mu$ mol). Fmoc-Ser(*t*Bu)-OH was coupled five times (see general procedure pS3) on the *N*-Hnb-Cys device and a treatment of the peptide-resin using a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in a NMP / CH<sub>2</sub>Cl<sub>2</sub> mixture (5:1, 20 min, 3 mL,  $\times$  3) was performed. A capping step was performed (protocol p S3) then the *N*-acylation yield (85%) was quantified by UV titration of the fluorenylmethylpiperidine byproduct from the Fmoc removal ( $\epsilon^{301} = 7800 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ). The peptide sequence was elongated using standard SPPS (protocol p S3), a double coupling being performed for the last residue (pyroglutamic acid), before a final piperidine treatment. Elongation yield (56%) was determined by UV titration of the fluorenylmethylpiperidine byproduct from the last Fmoc removal (Leu<sup>2</sup>), as compared to the first one (Ser<sup>50</sup>). For this purpose, after the coupling of Leu<sup>2</sup> and prior to the piperidine treatment, peptide-resin was treated with a solution of hydroxylamine hydrochloride and imidazole as above described, in order to cleave any Fmoc-Leu ester on the Hnb moiety. After TFA-mediated cleavage of the peptide from the resin and deprotection (protocol p S3), **5** was directly used in a NCL reaction without further purification.

### 5

**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> calcd. for C<sub>266</sub>H<sub>376</sub>N<sub>79</sub>O<sub>81</sub>S<sub>9</sub>: 6260.5218, found: 6260.5363.

**HPLC analysis**:  $t_R = 7.03$  min (Chromolith, gradient: 20-50% B/A over 9 min).



7b- Synthesis of **6** via NCLSupplementary scheme S12: NCL using crude **5** to obtain [1-51] AhPDF1.5 **6**.

800  $\mu\text{L}$  of a deoxygenated<sup>a</sup> 0.2 M pH 5.8 sodium phosphate buffer containing 200 mM MPAA, 50 mM TCEP, 6 M guanidine hydrochloride and 100 mM cysteine was added to crude dry peptide **5** (8 mg, corresponding to 1.0  $\mu\text{mol}$  of the crude cryptothioester mixtures,<sup>b</sup> and 0.56  $\mu\text{mol}$  of **5'**) under argon. Ligation was carried out at 50  $^\circ\text{C}$  and monitored by RP-HPLC. For this, 5  $\mu\text{L}$  aliquots of the reaction mixture were diluted with 100  $\mu\text{L}$  of 2% TFA in water and injected in HPLC. After 5 h, pH was adjusted to 1 with 1.5% TFA in water and ligation product **6** was purified by semi-preparative RP-HPLC (Nucleosil, gradient: 35-45% B/A over 20 min). **6** was isolated with an overall yield of 15.8% (based on the loading of Fmoc-Ser(*t*Bu)-(Hnb)C(*t*Bu)-G-Rink-Tentagel resin, and thus corresponding Fmoc SPPS elongation, TFA-mediated cleavage and deprotection, NCL and purification).

<sup>a</sup>: Deoxygenation was performed through four consecutive vacuum/argon cycles.

<sup>b</sup>: Quantity of crude peptide cryptothioester mixture evaluated from the Fmoc-Ser-(Hnb)C(*t*Bu)-G-Rink-Tentagel resin loading, as determined by UV titration of the fluorenylmethylpiperidine byproduct from the Fmoc removal ( $\epsilon^{301} = 7800 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ).

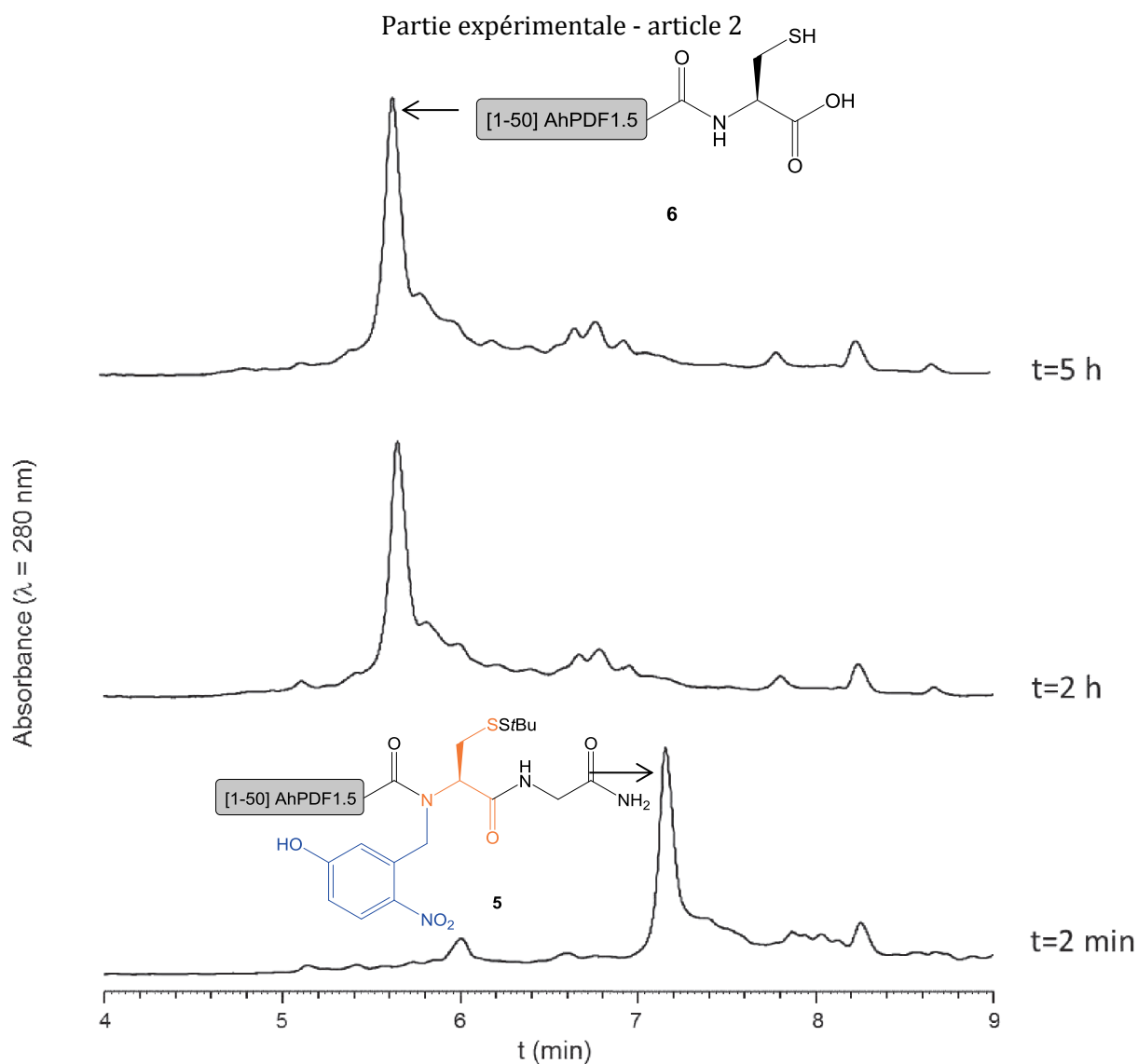
<sup>c</sup>: As evaluated from the SPPS elongation yield (56%).

**6**

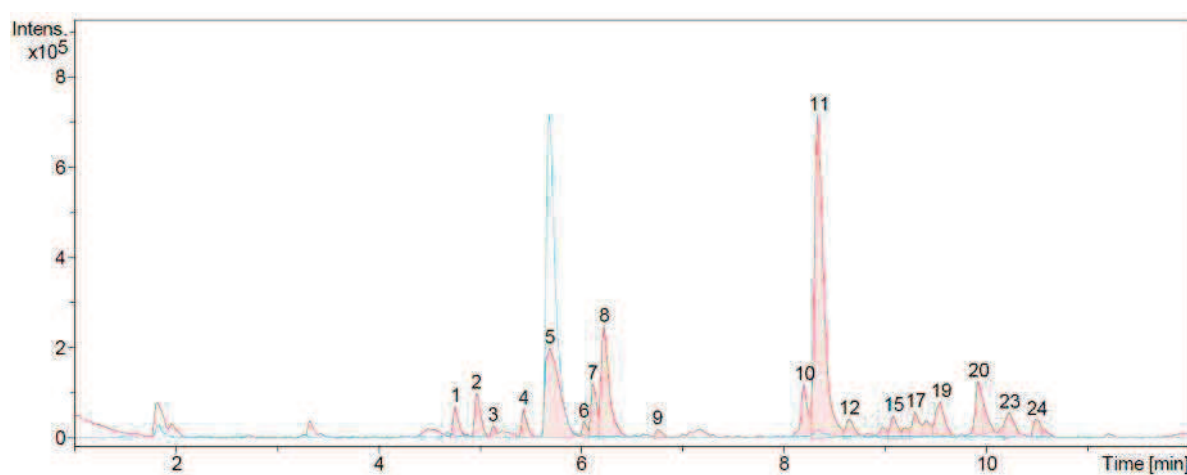
**ESI-HRMS** ( $m/z$ ):  $[\text{MH}]^+$  calcd. for  $\text{C}_{253}\text{H}_{359}\text{N}_{76}\text{O}_{78}\text{S}_8$ : 5965.4227, found: 5965.4221.

**HPLC analysis**:  $t_R = 5.65 \text{ min}$  (Chromolith, gradient: 20-50% B/A over 9 min).





**Supplementary figure S18:** Analytical HPLC monitoring of the NCL reaction with crude **5** to obtain **6** showing completion of the reaction in < 2 h.



**Supplementary figure S19:** LC/MS analysis of NCL to obtain **6** after 6 h. Blue trace: UV ( $\lambda = 214 \text{ nm}$ ); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
5 (5.69)	289.0357 <sup>a</sup>	289.0442 <sup>a</sup>	MPAA <sup>a</sup>
8 (6.23)	-	469.1979	Not attributed
10 (8.20)	856.2832	856.2811	H-[45-51] AhPDF1.5 <sup>b</sup>
10 (8.20)	5862.4130	5862.4092	H-[1-50] AhPDF1.5 <sup>c</sup>
11 (8.33)	5965.4227	5965.4221	<b>6</b>
20 (9.92)	352.0677 <sup>d</sup>	352.0656 <sup>d</sup>	MPAA disulfide

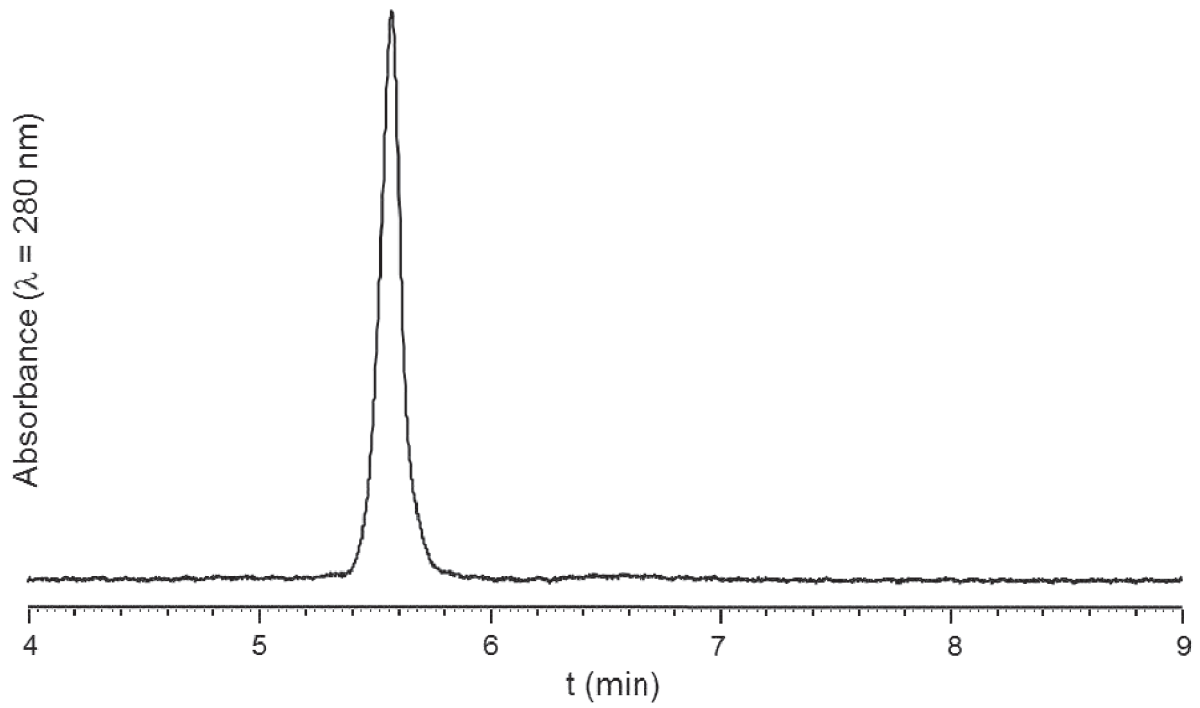
Supplementary table S6: Attribution of the main peaks observed during LC/MS analysis of NCL to obtain **6** after 6h.

<sup>a</sup>: Attribution based on the characteristic UV spectrum of this compound, and on the corresponding HPLC chromatogram showing a peak at the expected retention time for MPAA. Note that under our MS conditions, no peaks corresponding to the expected  $m/z$  for protonated MPAA neither cation adducts were observed. A possible explanation for the observed  $m/z$  could be oxidation into disulfide then decarboxylation to yield a benzyl cation ( $C_{15}H_{13}O_2S_2$ , calcd.  $m/z$  289.0357). Note that minor peaks corresponding to  $M+Na^+$  and  $M+NH_4^+$  relative to the disulfide were also observed.

<sup>b</sup>: Traces of a non-acetylated truncated peptide containing a N-terminal Cys residue were observed in the LC-MS analysis of the mixture. This co-product was attributed to *in situ* N→S shift at Xaa-Cys sites followed by trans-thioesterification, as reported by MacMillan.<sup>8</sup> Note that the ion chromatogram signal is not representative of the actual proportion of this peptide, as compared with the HPCL-UV analysis. Smaller peptides very probably lead to a higher BPC signal due to greater ionization.

<sup>c</sup>: [1-50] AhPDF1.5 corresponds to a co-product arising from hydrolysis of peptide crypto-thioester **6'**. [1-50] AhPDF1.5 proportion was evaluated to 2% relative to **7** by integration of the corresponding extracted ion chromatograms, neglecting the eventual differences in terms of ionization of the two compounds.

<sup>d</sup>: The calculated  $m/z$  value corresponds to a  $[M+NH_4]^+$  form ( $C_{16}H_{18}NO_4S_2$ ), a  $[M+Na]^+$  species being also observed but not  $[M+H]^+$ . A minor peak consistent with decarboxylation to yield a benzyl cation (see note a) is also observed.



Supplementary figure S20: HPLC trace of purified **6**.

## Partie expérimentale article 3 :

Highly efficient synthesis of cysteine-rich cyclic peptides through intramolecular native chemical ligation of *N*-Hnb-Cys peptide crypto-thioesters

Terrier V. P.; Delmas A. F.; Aucagne V., *Chem. Commun.*, **2016**, article soumis.

### Supporting Information

#### **Table of contents:**

1) General information	293
2) General procedures for solid phase peptide synthesis	293
3) Syntheses of peptide linear precursors <b>1a-5a</b>	294
4) Peptides cyclization <i>via</i> intramolecular NCL	307

## 1- General information

All reagents and solvents were used without further purification. Protected amino acids, Rink's linker and HCTU were purchased from Merck Biosciences (Nottingham, UK). Aminomethyl TentaGel R resin was purchased from Rapp polymers (Tuebingen, Germany). Peptide synthesis grade DMF and HATU were purchased 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).

High resolution ESI-MS analyses were performed on a maXis<sup>TM</sup> ultra-high-resolution Q-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany), using the positive mode.

HPLC analyses and 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 autosampler. Nucleosil C18 (300 Å, 5 µm, 250 × 4.6 mm, 1 mL/min flow rate) or Chromolith HighResolution RP-18e (150 Å, 10 × 4.6 mm, 3 mL/min flow rate) columns were used for analysis and Nucleosil C18 (300 Å, 5 µm, 250 × 10 mm, 3 mL/min flow rate) for purification. Chromatography was conducted at room temperature unless otherwise mentioned. Solvents A and B are 0.1% TFA in H<sub>2</sub>O and 0.1% TFA in MeCN, respectively. Each gradient was followed by a washing step (up to 95% B/A over 0.5 min for the HR Chromolith; over 1 min for the Nucleosil C18 column) to identify eventual co-products not eluted during the gradient. LC/HRMS analyses were carried out on an Ultimate<sup>®</sup> 3000 RSLC HPLC system (Dionex, Germering, Germany), coupled with the maXis<sup>TM</sup> mass spectrometer and fitted with a Zorbax 300 SB-C18 RRHD (300 Å, 1.8 µm, 100 × 2.1 mm, 0.5 mL/min flow rate, 40°C) column. Solvents A and B were 0.1% formic acid in H<sub>2</sub>O 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.

For yield calculations purposes, the quantities of purified peptides were determined by weight, taking into account a molecular mass including trifluoroacetate counterions (one per Arg, His, Lys and N-terminal amine of the peptide sequence) but not eventual hydration.

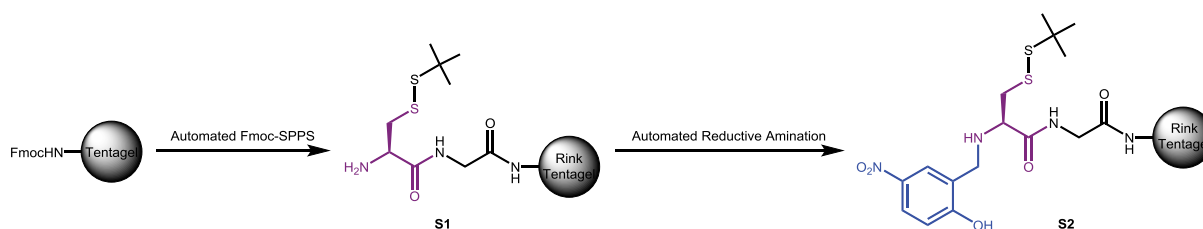
## 2- General procedures for solid phase peptide synthesis

Fmoc-based solid phase peptide syntheses (SPPS) were carried out on a Prelude synthesizer from Protein Technologies. Standard side-chain protecting groups were used: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Cys(StBu), Glu(OtBu), Gln(Trt), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc) and Tyr(tBu).

Syntheses were performed at a 25  $\mu\text{mol}$  scale. Protected amino acids (0.25 mmol, 10 equiv.) were coupled using HCTU (98 mg, 0.238 mmol, 9.5 equiv.) and *i*Pr<sub>2</sub>NEt (87  $\mu\text{L}$ , 0.5 mmol, 20 equiv.) in NMP (3 mL) for 30 min. Capping of eventual unreacted amine groups was achieved by treatment with acetic anhydride (143  $\mu\text{L}$ , 1.51 mmol, 60 equiv.), *i*Pr<sub>2</sub>NEt (68  $\mu\text{L}$ , 0.39 mmol, 15.5 equiv.) and HOBt (6 mg, 0.044 mmol, 1.8 equiv.) in NMP (3 mL) for 7 min. Fmoc group was removed by three successive treatments with 20% piperidine in NMP (3 mL) for 3 min.

The crude peptides were deprotected and cleaved from the resin through a treatment with TFA/H<sub>2</sub>O/*i*Pr<sub>3</sub>SiH/phenol, 88/5/2/5 for 2 h, then precipitated by dilution into an ice-cold 1:1 diethyl ether/petroleum ether mixture, recovered by centrifugation and further washed three times with diethyl ether.

### 3- Syntheses of peptide linear precursors 1a-5a



Supplementary scheme S1: Synthesis of peptide-resin **S2**.

The Rink's linker, Gly and Cys(*t*Bu) were successively coupled by automated SPPS on a Tentagel R resin (120 mg, 0.21 mmol/g, 25  $\mu\text{mol}$ ) in order to obtain peptide-resin **S1**.

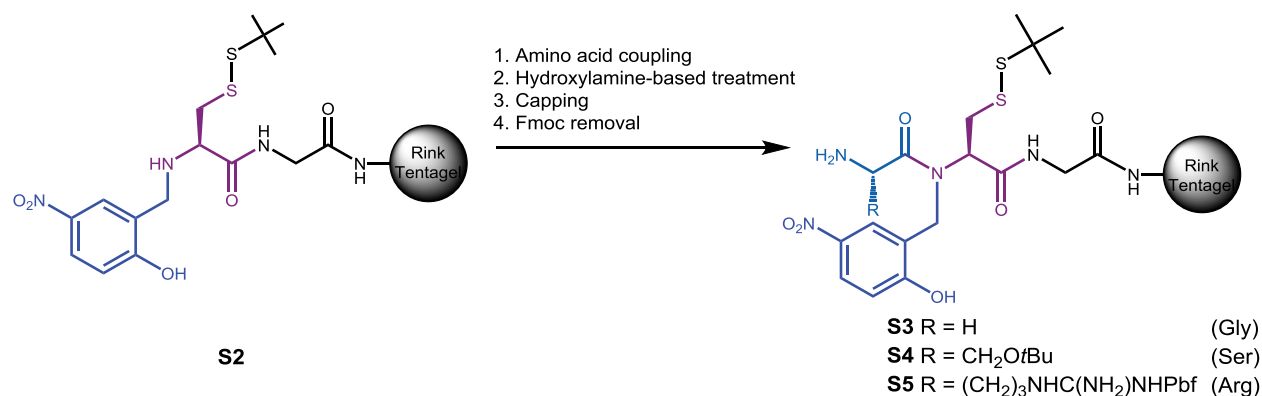
#### Reductive amination:

Peptide-resin **S1** (25  $\mu\text{mol}$ ) was washed two times with 3 mL of a 1:1 DMF/MeOH mixture for 30 s, then swollen in 3 mL of a 9:9:2 DMF/MeOH/AcOH mixture for 5 min. The reactor was drained off and the resin was washed four times with 3 mL of a 1:1 DMF/MeOH mixture for 30 s. This process forms the acetic acid salt of the amine group of cysteine.

2-Hydroxy-5-nitrobenzaldehyde (HNBA) in 1:1 DMF/MeOH (125 mM, 10 equiv., 2 mL) was then added then the reactor was left for 1 h under stirring by nitrogen bubbling. The reactor was drained off and the resin was washed four times with 3 mL of 1:1 DMF/MeOH for 15 s.

Without delay, a fresh solution of sodium cyanoborohydride in 9:9:2 DMF/MeOH/AcOH (250 mM, 20 equiv. 2 mL) were added and the reactor was left for 1 h under stirring by nitrogen bubbling. The reactor was drained off and the resin was washed with 1:1 DMF/MeOH (3 mL, 30 s,  $\times$  4), NMP (3 mL, 30 s,  $\times$  3), 20%

piperidine in NMP (3 mL, 30 s, × 3), NMP (3 mL, 30 s, × 3), dichloromethane (5 mL, 30 s, × 3) and NMP (3 mL, 30 s, × 2).

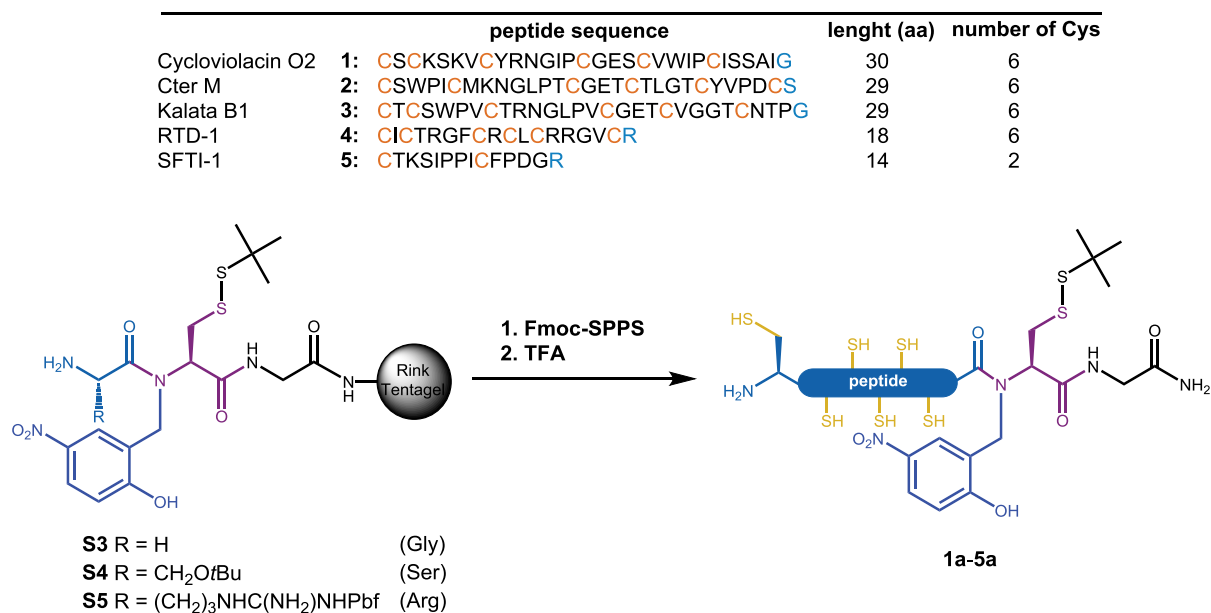


**Supplementary scheme S2:** Installation of the C-terminal amino acid of the sequence and *N*-acylation quantification.

The C-terminal Fmoc-protected amino acid of the target sequence was coupled twice (except for Fmoc-Ser(*t*Bu)-OH: three times) on peptide-resin **S2** (general procedure p S3). Peptide-resin was treated with a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in 5:1 NMP/CH<sub>2</sub>Cl<sub>2</sub> (3 mL, 20 min, × 3). After a capping step (general procedure p S3), the Fmoc group was removed through a standard piperidine treatment (p S3) and the *N*-acylation yield was quantified using UV titration of the fluorenylmethylpiperidine byproduct at 301 nm ( $\epsilon = 7800 \text{ L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ ) (see table S1).

Peptide sequence	Cycloviolacin O2	Cter M	Kalata B1	RTD-1	SFTI-1
Peptide-resin	<b>S3</b>	<b>S4</b>	<b>S3</b>	<b>S5</b>	<b>S5</b>
Amino acid	Gly	Ser( <i>t</i> Bu)	Gly	Arg(Pbf)	Arg(Pbf)
<i>N</i> -acylation yield (%)	89	85	89	77	77

**Supplementary table S1:** *N*-acylation yield for the introduction of the C-terminal amino acid of the sequence.

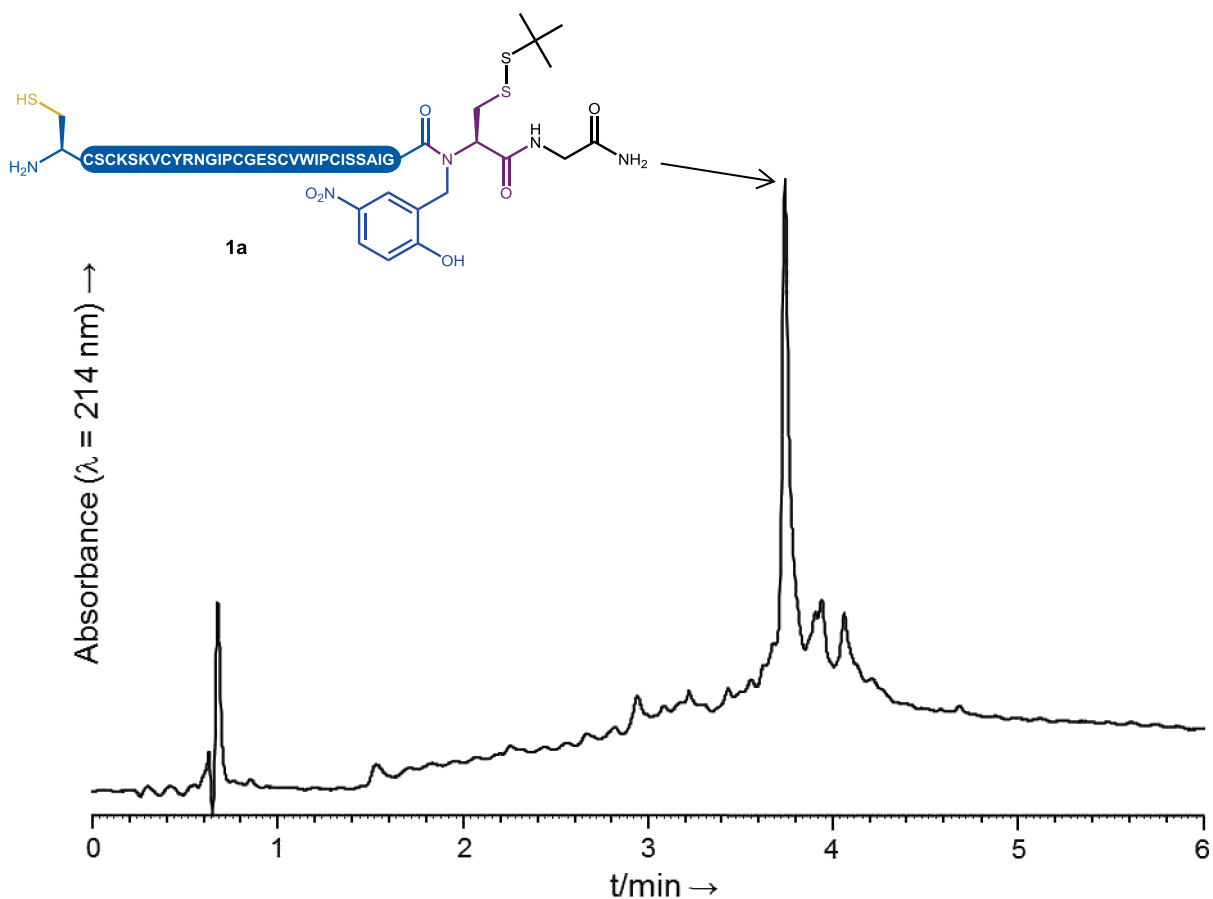
**Peptide elongation:****Supplementary scheme S3: Fmoc-based SPPS elongation of crypto-thioester peptides 1a-5a.**

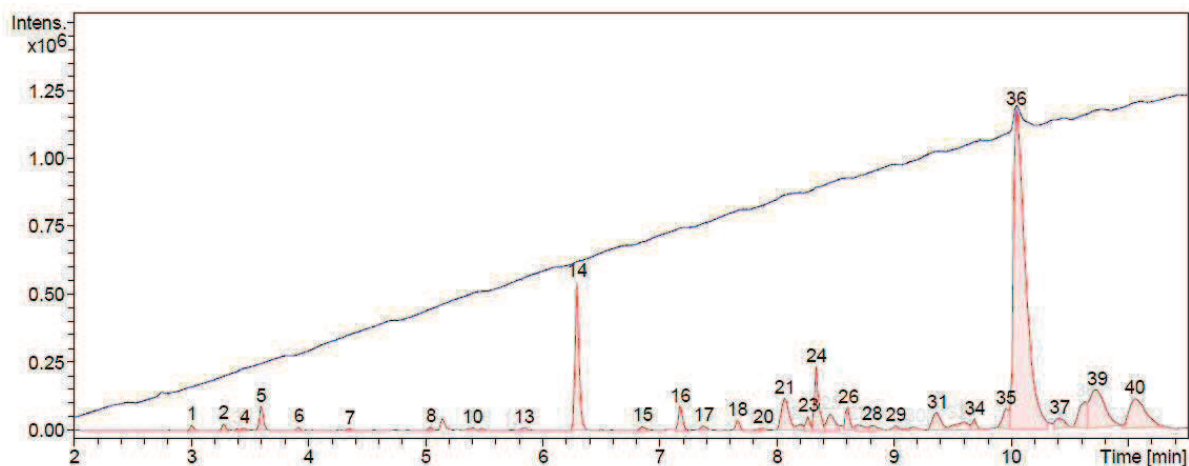
Then, the five different peptide sequences were elongated through standard Fmoc-based SPPS (general procedure p S3). Elongation yields were determined by UV titration of the deprotection of the first and last amino acid residues of the sequence. For this purpose, after the coupling of the N-terminal cysteine, peptide-resins were treated with a solution of hydroxylamine hydrochloride (0.3 M) and imidazole (0.225 M) in 5:1 NMP/CH<sub>2</sub>Cl<sub>2</sub> (3 mL, 20 min, × 3) prior to the piperidine treatment, in order to cleave any Fmoc-Cys(Trt) ester on the Hnb moiety. A final TFA treatment (general procedure p S3) afforded peptides **1a-5a** that were purified by RP-HPLC.

Peptide	1a	2a	3a	4a	5a
Elongation yield (%)	56	65	56	74	85
Yield / cycle (%)	98.1	98.5	98.0	98.4	98.8
Isolated yield (%)	18	10	16	11	21

**Supplementary table S2: Yields for peptides 1a-5a.**



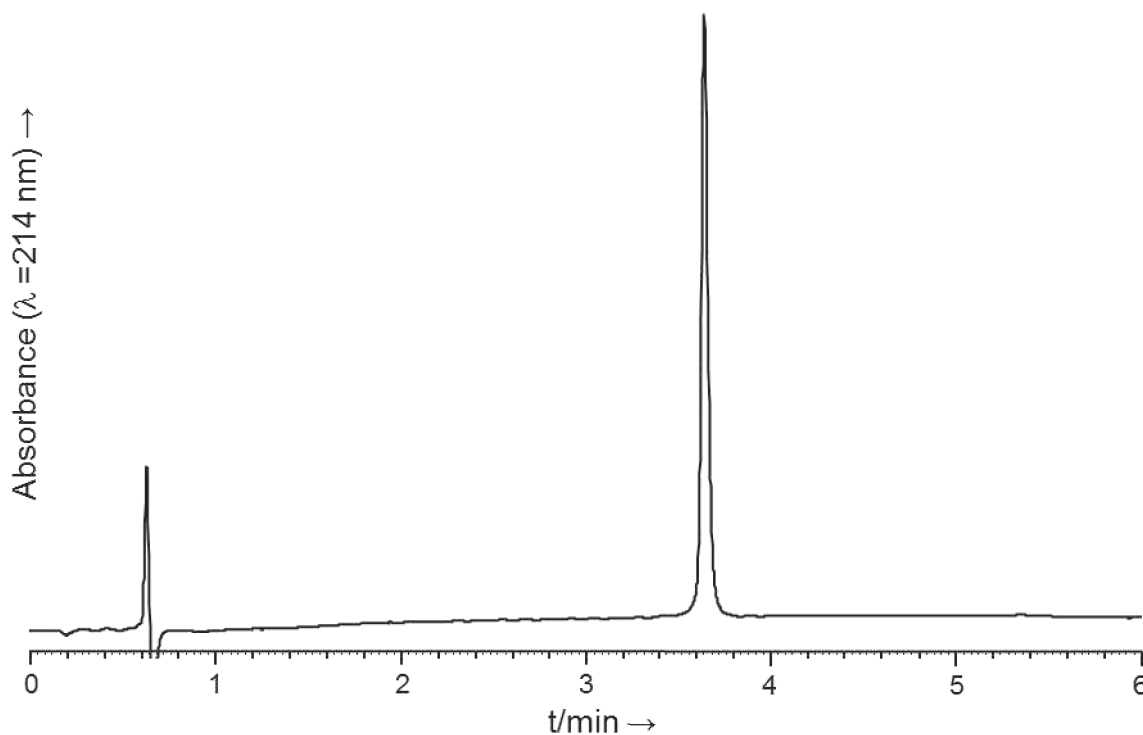
Linear crypto-thioester precursor of Cycloviolacin O2 (1a):Sequence: H-CSCKSKVCYRNGIPCGESCWIPCISSAIG-(Hnb)C(S<sup>t</sup>Bu)-G-NH<sub>2</sub>**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd. for C<sub>149</sub>H<sub>238</sub>N<sub>41</sub>O<sub>44</sub>S<sub>8</sub>: 3561.5412, found: 3561.5415.**HPLC analysis:** *t<sub>R</sub>* = 3.64 min (Chromolith, gradient: 20-70% B/A over 5 min).**HPLC purification:** Nucleosil, gradient: 40-50% B/A over 10 min, 70 °C.**Yield:** 18% (overall yield, based on the initial loading of the resin).Supplementary figure S1: HPLC trace of crude 1a.



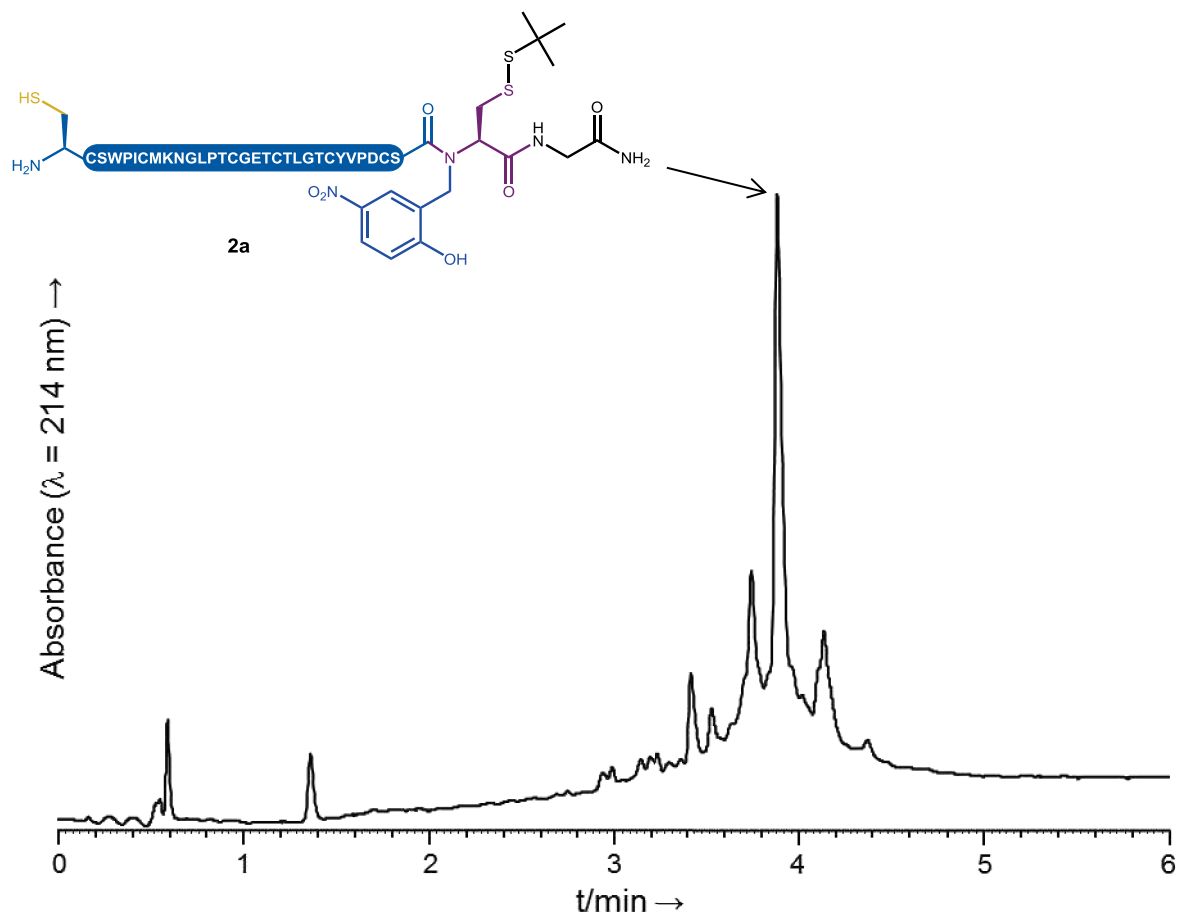
Supplementary figure S2: LC/MS analysis of crude **1a**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

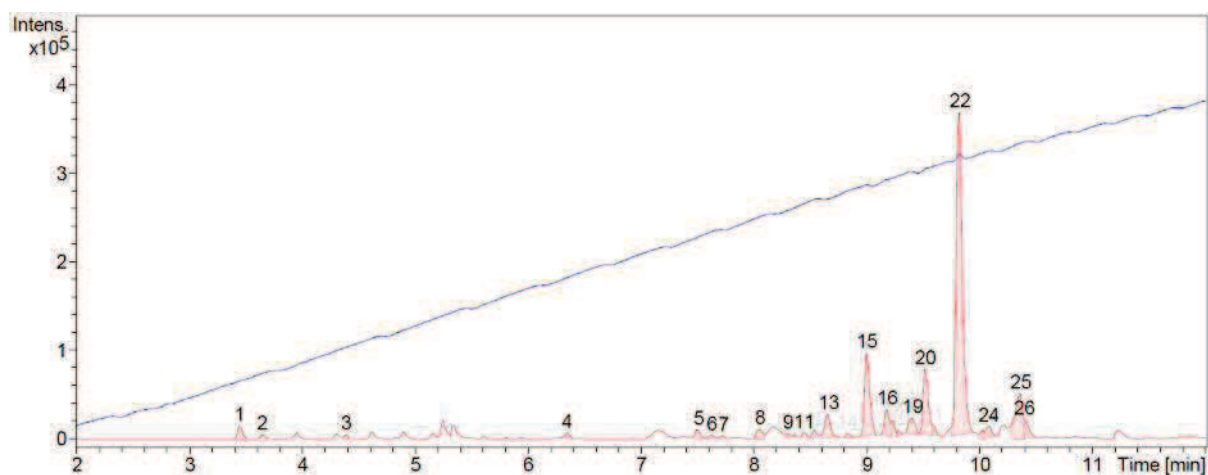
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
14 <sup>a</sup> (6.29)	417.1266	417.1251	(Hnb)C( <i>St</i> Bu)-G-NH <sub>2</sub>
36 (10.04)	3561.5412	3561.5415	<b>1a</b>
38 (10.63)	3617.6038	3617.6042	<b>1a</b> + <i>t</i> Bu
39 (10.73)	3617.6038	3617.6042	<b>1a</b> + <i>t</i> Bu
40 (11.05)	3617.6038	3617.6040	<b>1a</b> + <i>t</i> Bu

Supplementary table S3: Attribution of the main peaks observed during LC/MS analysis of crude **1a**. <sup>a</sup>: Note that the large LC/MS peak is not representative of the actual quantity.



Supplementary figure S3: HPLC trace of purified **1a**.

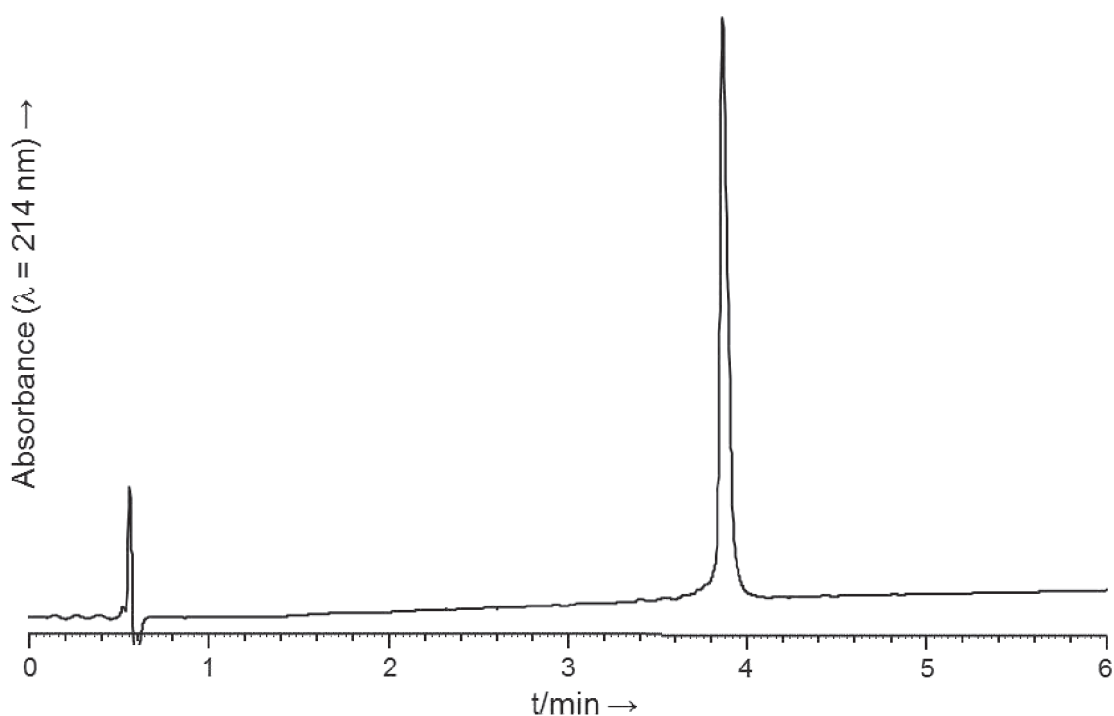
Linear crypto-thioester precursor of Cter M (2a):Sequence: H-CSWPICMKNGLPCTGGETCTLGTCYVPDCS-(Hnb)C(S*t*Bu)-G-NH<sub>2</sub>**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd. for C<sub>144</sub>H<sub>223</sub>N<sub>36</sub>O<sub>46</sub>S<sub>9</sub>: 3480.3704, found: 3480.3686.**HPLC analysis:** *t*<sub>R</sub> = 3.88 min (Chromolith, gradient: 20-60% B/A over 5 min).**HPLC purification:** Nucleosil, gradient: 40-50% B/A over 10 min.**Yield:** 10% (overall yield, based on the initial loading of the resin).Supplementary figure S4: HPLC trace of crude 2a.



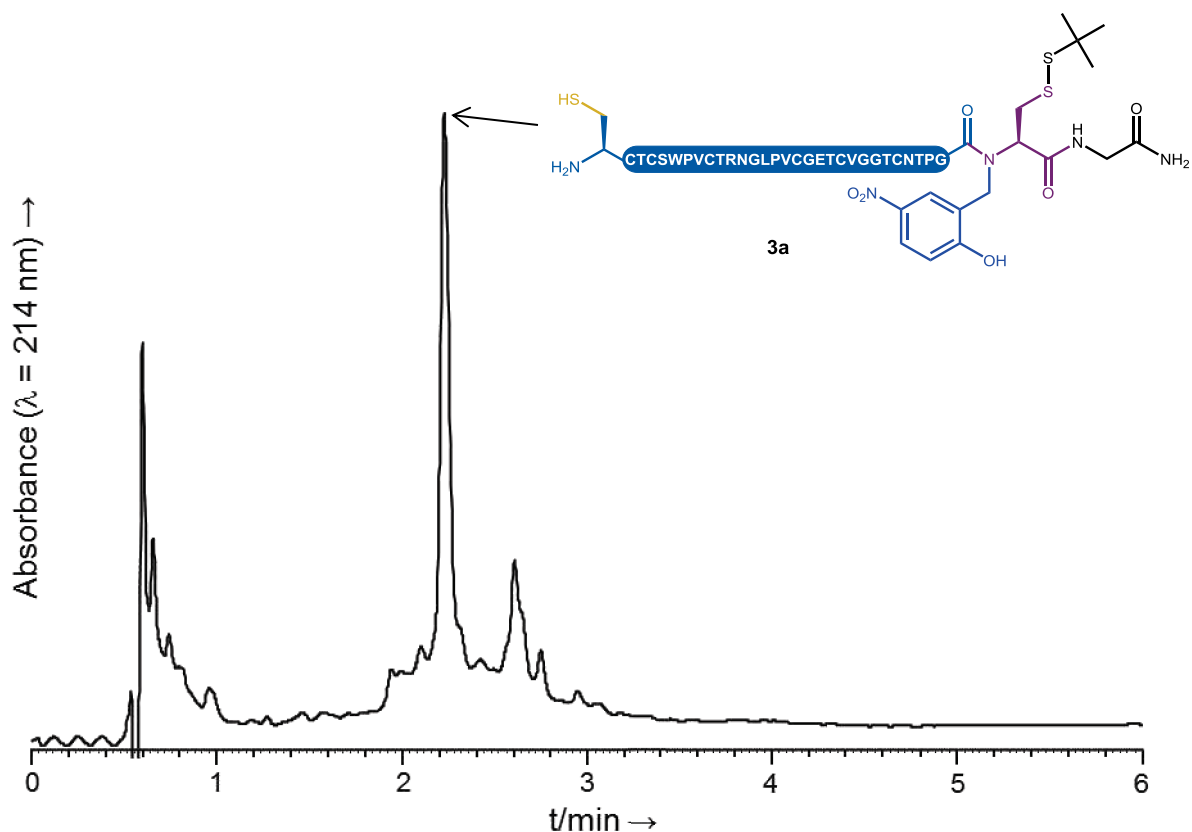
Supplementary figure S5: LC/MS analysis of crude **2a**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

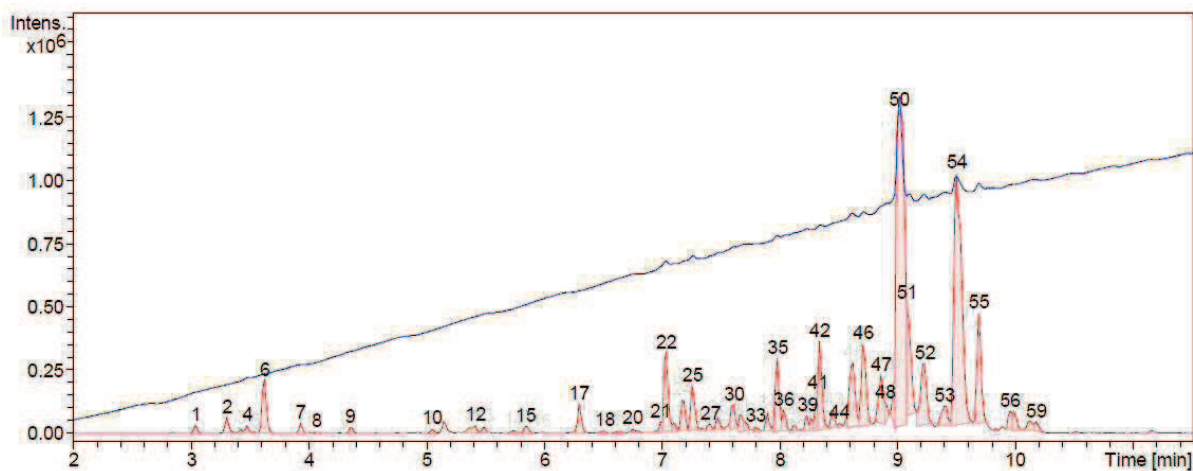
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
4 (6.35)	417.1266	417.1253	(Hnb)C( <i>St</i> Bu)-G-NH <sub>2</sub>
15 <sup>a</sup> (8.99)	3329.3434	3329.3418	<b>2a</b> - Hnb
20(9.52)	3496.3653	3496.3643	<b>2a</b> with oxidized methionine
22 (9.82)	3480.3704	3480.3686	<b>2a</b>
25 (10.36)	3536.4330	3536.4325	<b>2a</b> + <i>t</i> Bu
26 (10.39)	3536.4330	3536.4325	<b>2a</b> + <i>t</i> Bu

Supplementary table S4: Attribution of the main peaks observed during LC/MS analysis of crude **2a**. <sup>a</sup>: Note that the large LC/MS peak is not representative of the actual quantity.



Supplementary figure S6: HPLC trace of purified **2a**.

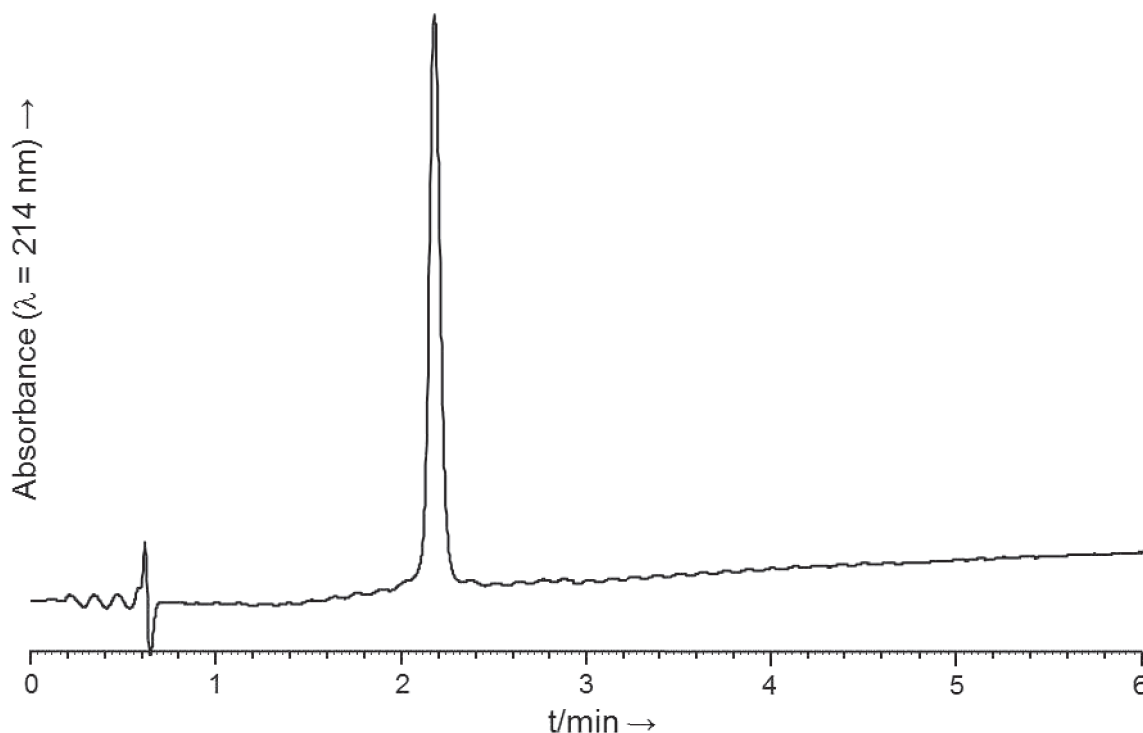
Linear crypto-thioester precursor of Kalata B1 (3a):Sequence: H-CTCSWPVCTRNGLPVCGETCVGGTCNTPG-(Hnb)C(S*t*Bu)-G-NH<sub>2</sub>**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd. for C<sub>133</sub>H<sub>210</sub>N<sub>39</sub>O<sub>44</sub>S<sub>8</sub>: 3313.3160, found: 3313.3153.**HPLC analysis:** *t*<sub>R</sub> = 2.10 min (Chromolith, gradient: 30-60% B/A over 5 min).**HPLC purification:** Nucleosil, gradient: 35-45% B/A over 10 min.**Yield:** 16% (overall yield, based on the initial loading of the resin).Supplementary figure S7: HPLC trace of crude **3a**.



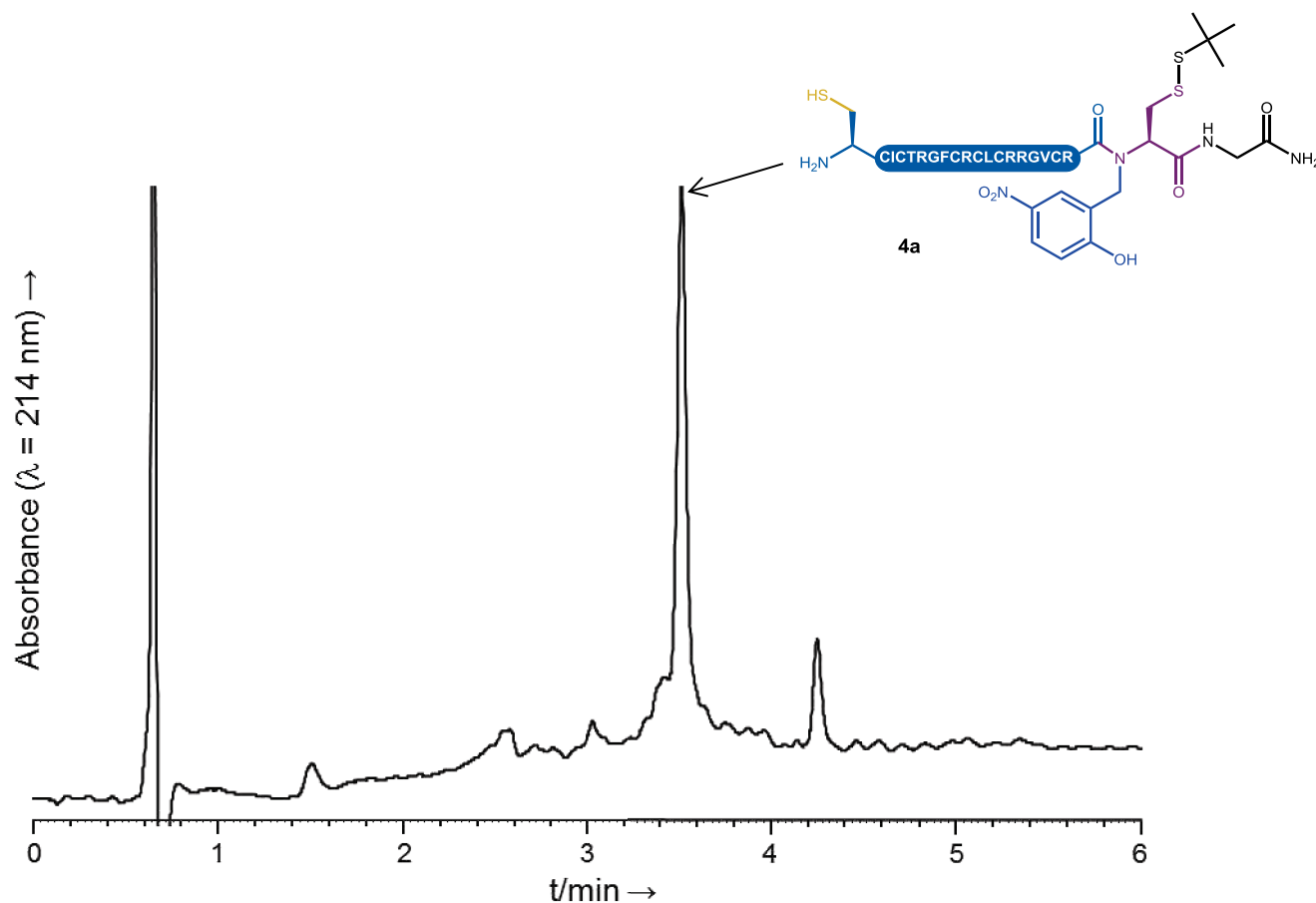
Supplementary figure S8: LC/MS analysis of crude **3a**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

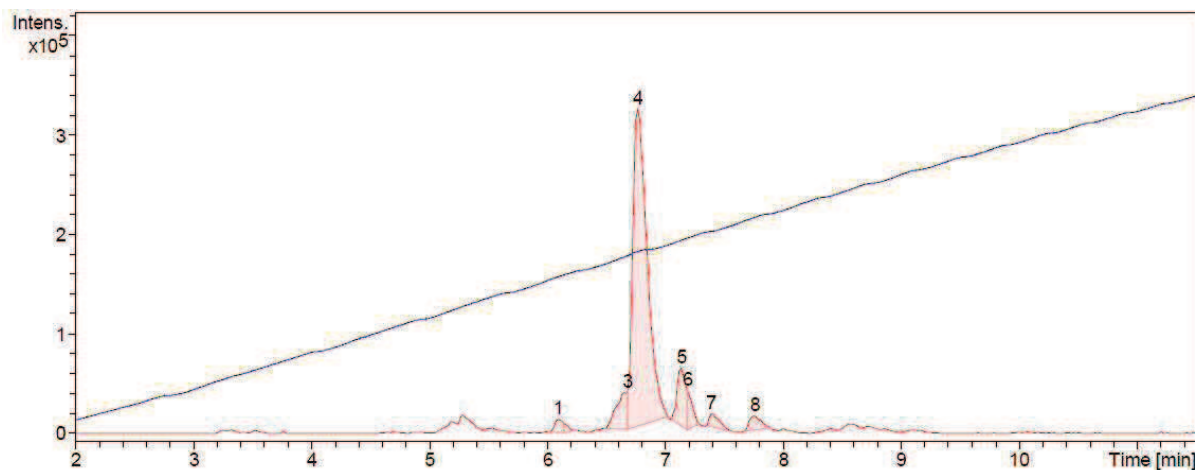
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
50 (9.01)	3313.3160	3313.3153	<b>3a</b>
54 <sup>a</sup> (9.50)	3369.3786	3369.3777	<b>3a + tBu</b>
55 (9.68)	3369.3786	3369.3764	<b>3a + tBu</b>

Supplementary table S5: Attribution of the main peaks observed during LC/MS analysis of crude **3a**. <sup>a</sup>: Note that the large LC/MS peak is not representative of the actual proportion of **3a** / **3a + tBu**, see HPLC trace on figure S7 for quantification.



Supplementary figure S9: HPLC trace of purified **3a**.

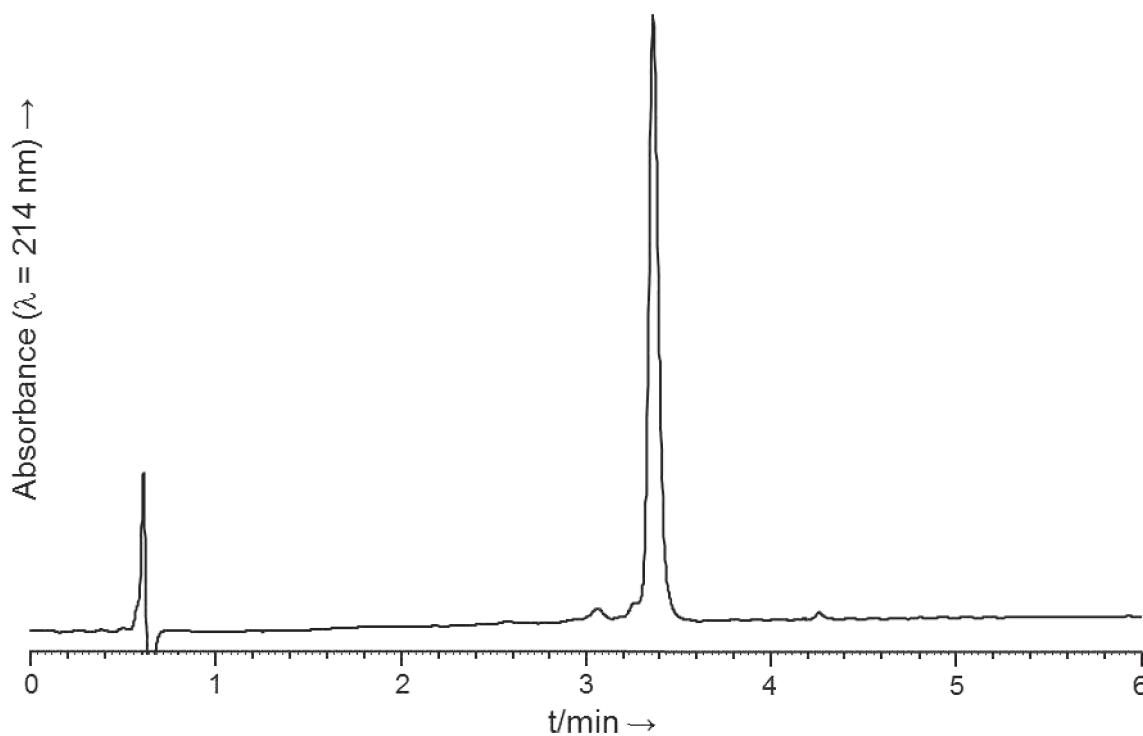
Linear crypto-thioester precursor of RTD-1 (**4a**):Sequence: H-CICTRGFCRCLCRRGVCR-(Hnb)C(S*t*Bu)-G-NH<sub>2</sub>**ESI-HRMS** (*m/z*): [MH]<sup>+</sup> calcd. for C<sub>98</sub>H<sub>168</sub>N<sub>37</sub>O<sub>24</sub>S<sub>8</sub>: 2503.0829, found: 2503.0873.**HPLC analysis**: *t*<sub>R</sub> = 3.52 min (Chromolith, gradient: 20-50% B/A over 5 min).**HPLC purification**: Nucleosil, gradient: 30-35% B/A over 5 min.**Yield**: 11% (overall yield, based on the initial loading of the resin).



**Supplementary figure S11:** LC/MS analysis of crude **4a**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

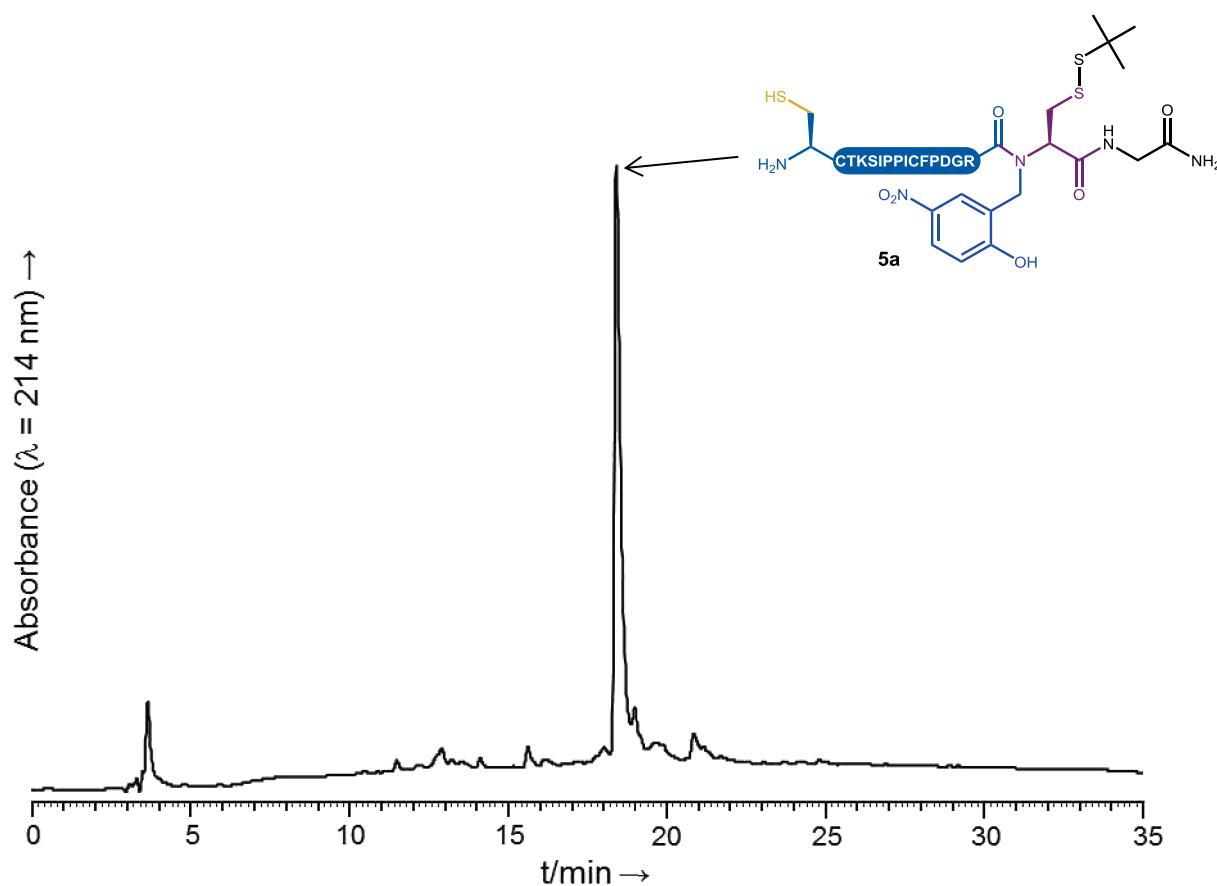
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
4 (6.76)	2503.0829	2503.0873	<b>4a</b>
5 (7.14)	615.2383	615.2369	Ac-R-(Hnb)C(StBu)-G-NH <sub>2</sub>
7 (7.39)	1505.6421	1505.6435	Ac-[10-18]RTD-1-(Hnb)C(StBu)-G-NH <sub>2</sub>
8 (7.76)	1968.8422	1968.8445	Ac-[6-18]RTD-1-(Hnb)C(StBu)-G-NH <sub>2</sub>

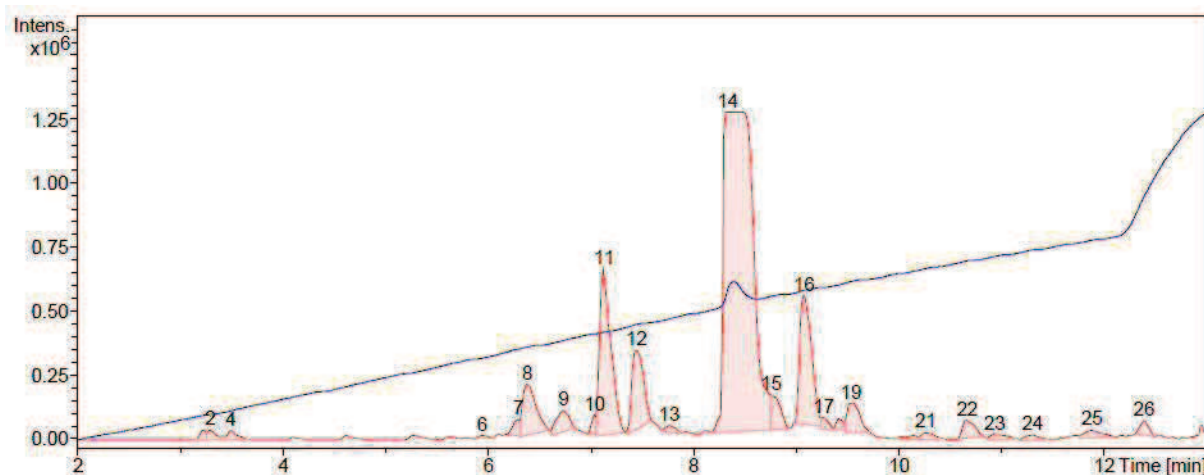
**Supplementary table S6:** Attribution of the main peaks observed during LC/MS analysis of crude **4a**.



**Supplementary figure S12:** HPLC trace of purified **4a**.



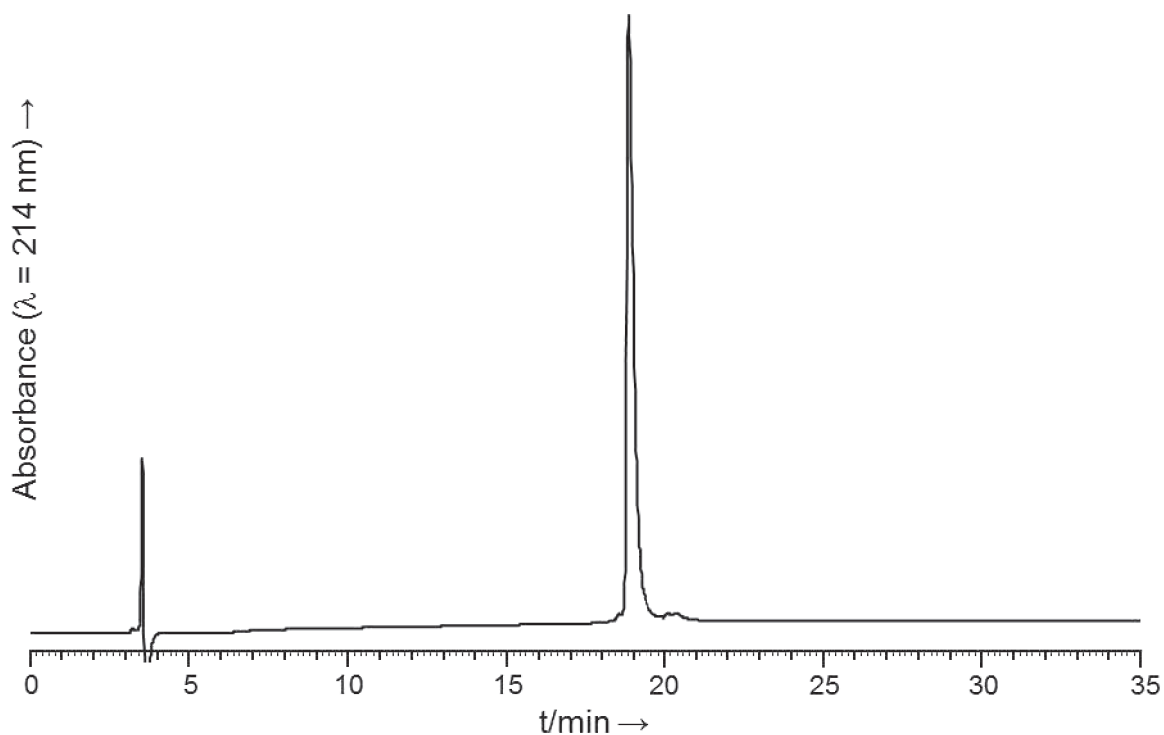
Linear crypto-thioester precursor of SFTI-1 (**5a**):Sequence: H-CTKSIPPICFPDGR-(Hnb)C(*St*Bu)-G-NH<sub>2</sub>**ESI-HRMS** ( $m/z$ ): [MH]<sup>+</sup> calcd. for C<sub>83</sub>H<sub>131</sub>N<sub>22</sub>O<sub>23</sub>S<sub>4</sub>: 1931.8640, found: 1931.8651.**HPLC analysis:**  $t_R$  = 18.45 min (Nucleosil, gradient: 20-60% B/A over 30 min, 70 °C). HPLC analysis was performed at high temperature due to the presence of large peaks when analysing **5a** on a chromolith column at room temperature, probably due to an equilibrium between two conformers of **5a**.**HPLC purification:** Nucleosil, gradient: 35-40% B/A over 5 min.**Yield:** 21% (overall yield, based on the initial loading of the resin).Supplementary figure S13: HPLC trace of crude **5a**.



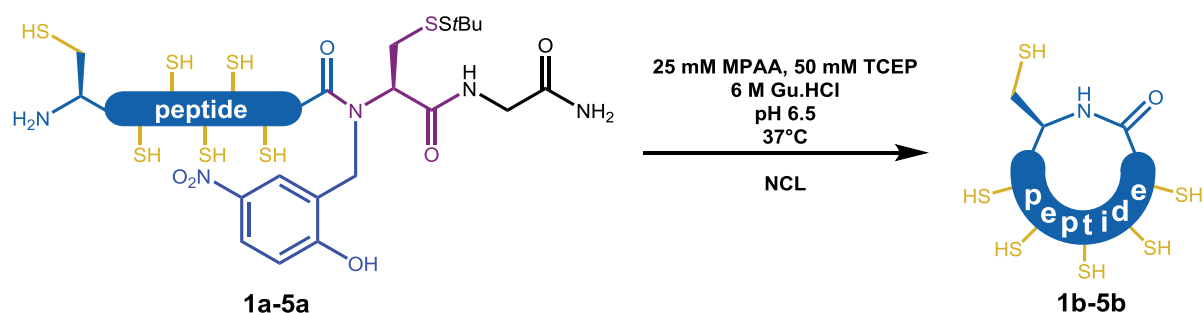
Supplementary figure S14: LC/MS analysis of crude **5a**. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
11 (7.12)	615.2383	615.2368	Ac-R-(Hnb)C( <i>St</i> Bu)-G-NH <sub>2</sub>
12 <sup>a</sup> (7.44)	1780.8370	1780.8394	<b>5a</b> - Hnb
14 (8.33)	1931.8640	1931.8651	<b>5a</b>
16 (9.07)	1987.9266	1987.9285	<b>5a</b> + <i>t</i> Bu

Supplementary table S7: Attribution of the main peaks observed during LC/MS analysis of crude **5a**. <sup>a</sup>: Note that the large LC/MS peak is not representative of the actual quantity.



Supplementary figure S15: HPLC trace of purified **5a**.

4- Peptide cyclization *via* intramolecular NCL

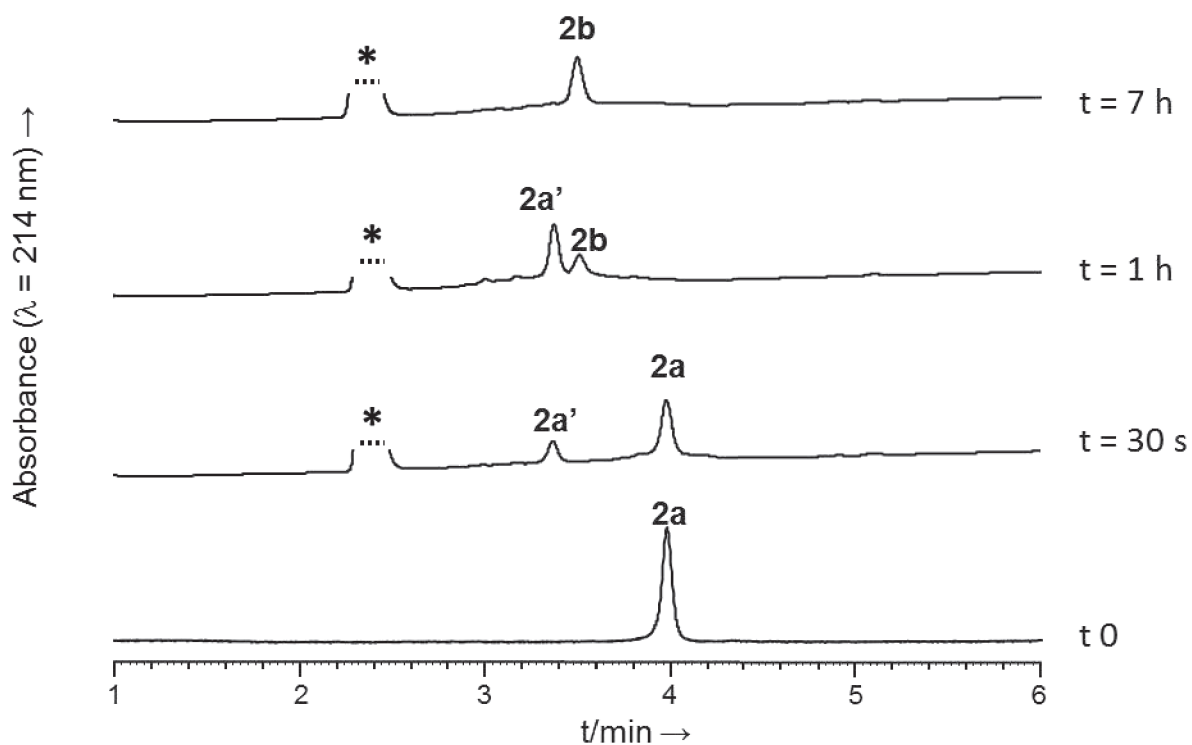
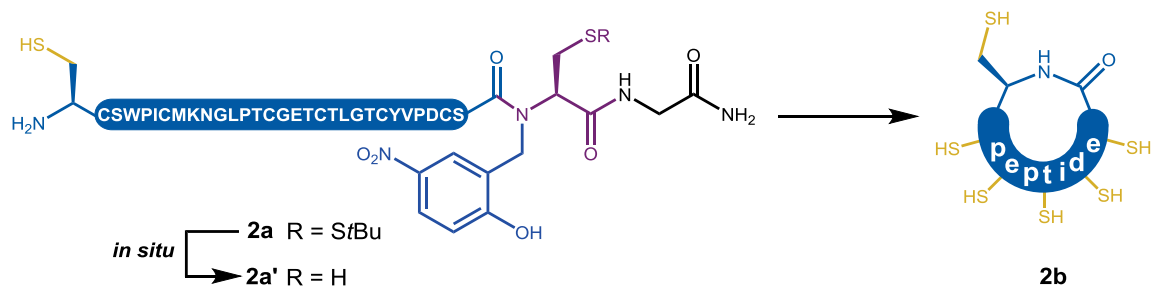
Supplementary scheme S4: Syntheses of peptides **1b-5b** *via* intramolecular NCL.

**General procedure:**

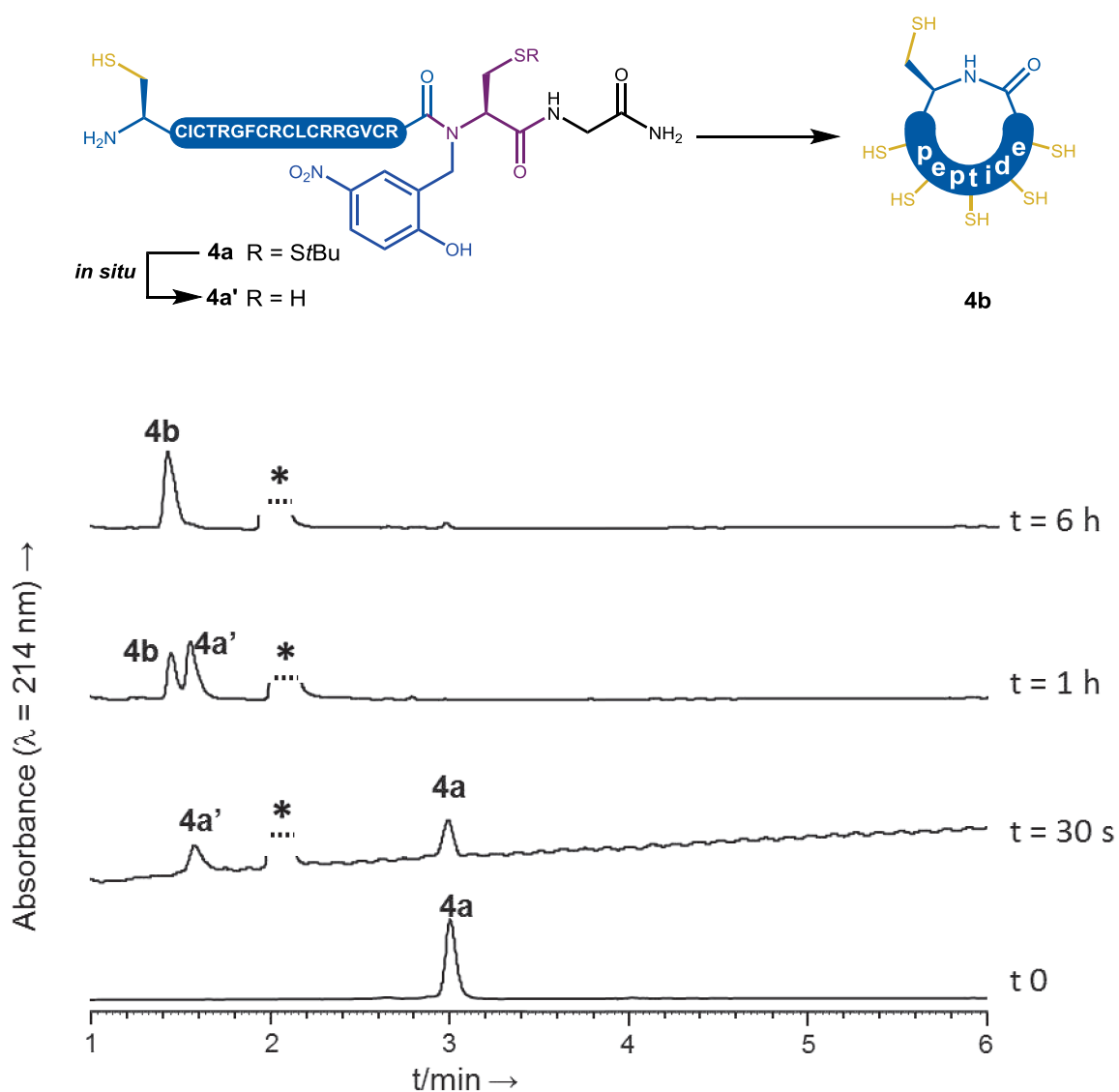
A deoxygenated<sup>a</sup> 0.2 M pH 6.5 sodium phosphate buffer containing 25 mM MPAA, 50 mM TCEP and 6 M guanidine hydrochloride was added to HPLC-purified dry peptides **1a-5a** (final concentration 1 mM) under argon. The ligations were carried out at 37°C and monitored by RP-HPLC. For this, 2  $\mu\text{L}$  aliquots of the reaction mixture were diluted with 108  $\mu\text{L}$  of 1.5% TFA in water and 100  $\mu\text{L}$  were injected in HPLC. After completion, the reaction mixtures were acidified to pH 1 using 1.5% TFA in water in order to obtain a final volume of 1.5 mL. These solutions were then extracted with diethyl ether (4  $\times$  10 mL) to remove MPAA. In case of disulfide formation due to oxygen contamination, TCEP was added (final concentration 100 mM) and pH was adjusted to 5.0 using a 10 M NaOH solution; after 20 min, pH was adjusted to 1 with 1.5% TFA in water. The ligation products were purified by semi-preparative RP-HPLC.

For **3a** and **5a**, the intramolecular cyclization was also conducted at 0.1 mM peptide concentration, in order to minimize the formation of cyclic dimers **3c** and **5c**.

<sup>a</sup>: deoxygenation was performed through four consecutive vacuum/argon cycles.

Examples of reaction monitoring for the cyclization of **2b** and **4b**:Reduced form of Cter M (**2b**):

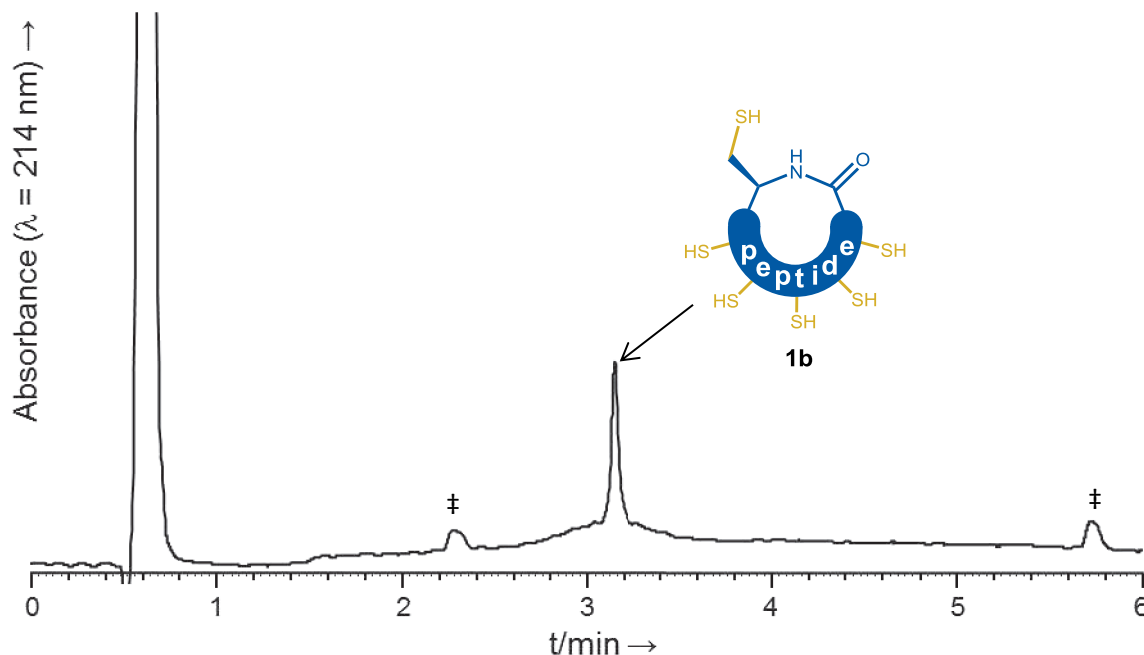
Supplementary figure S16: Analytical HPLC monitoring of the NCL-based cyclization of **2a**. Chromolith column, gradient: 20-60% B/A over 5 min. (\* = MPA).

**Reduced form of RTD-1 (4b):**

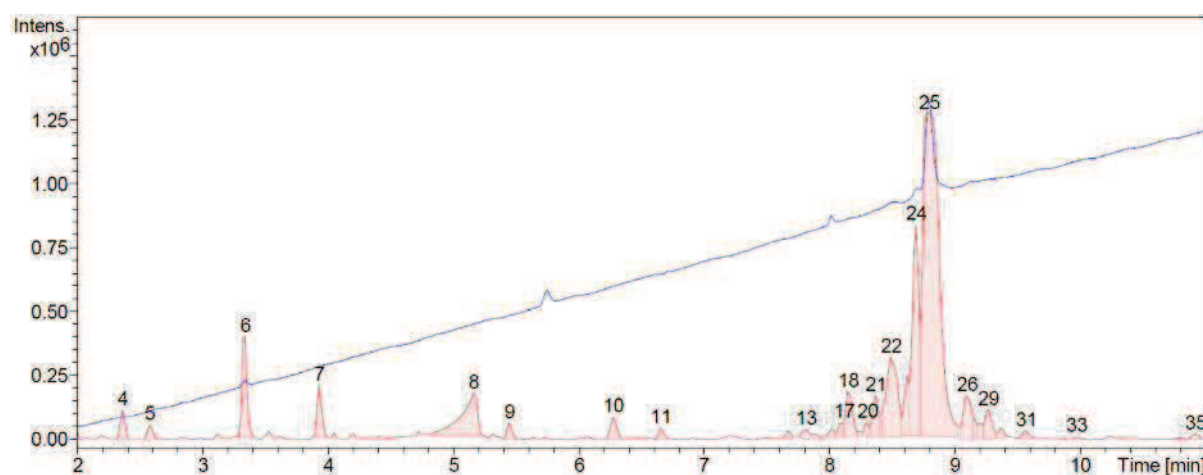
**Supplementary figure S17:** Analytical HPLC monitoring of the NCL-based cyclization of **4a**. Chromolith column, gradient: 25-45% B/A over 5 min. (\* = MPAA).

Reduced form of Cycloviolacin O2 (1b):

Sequence: cyclo(CSCKSKVCYRNGIPCGESCWIPCISSAIG)

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{133}H_{214}N_{37}O_{39}S_6$ : 3145.4224, found: 3145.4175.**HPLC analysis:**  $t_R$  = 3.15 min (Chromolith, gradient: 20-70% B/A over 5 min).**HPLC purification:** Nucleosil, gradient: 38-45% B/A over 7 min.**Yield:** 97%.

Supplementary figure S18: HPLC trace of **1b** after work-up (‡: non-peptidic contaminants).

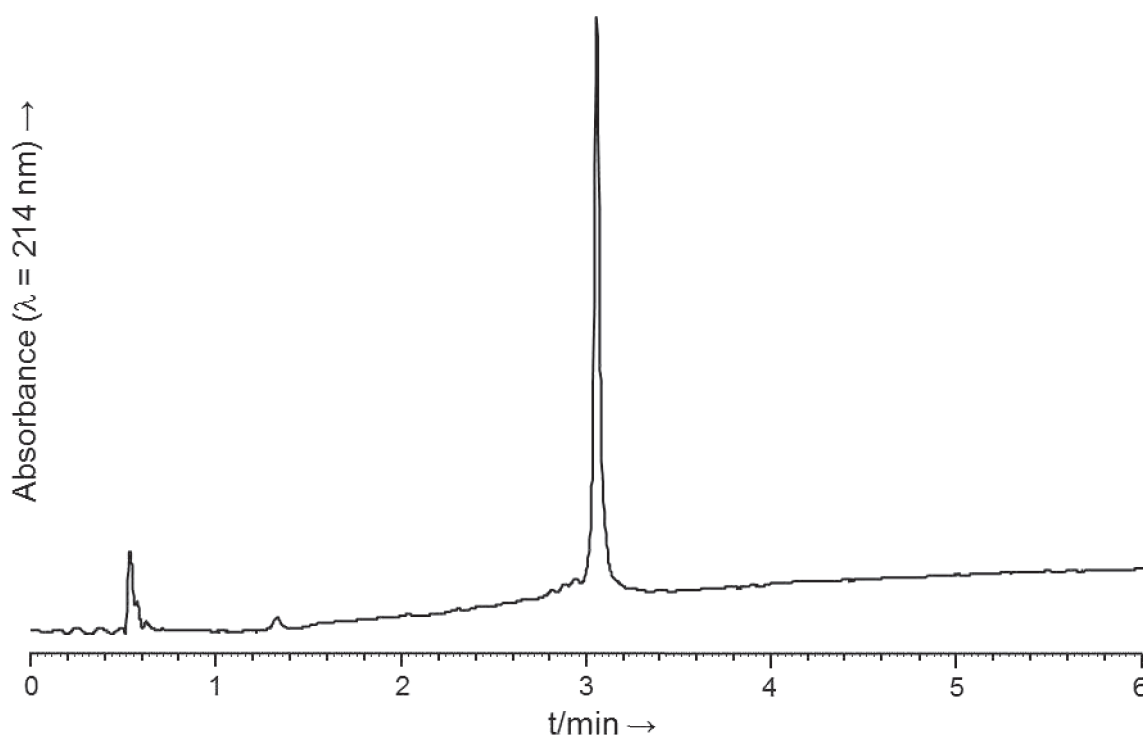


Supplementary figure S19: LC/MS analysis of crude **1b** after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
24 <sup>a</sup> (8.69)	-	3149.4145	Not attributed
25 (8.79)	3145.4224	3145.4175	<b>1b</b>
27 (9.14)	6289.8370	6289.8281	Cyclic dimer <b>S6</b> <sup>b</sup>

Supplementary table S8: Attribution of the main peaks observed during LC/MS analysis of crude **1b** after work-up. <sup>a</sup>: Note that the large LC/MS peak is not representative of the actual quantity.

<sup>b</sup>: **S6** = cyclo(CSCKSKVCYRNGIPCGESCVWIPCISSAIGCSCKSKVCYRNGIPCGESCVWIPCISSAIG).



Supplementary figure S20: HPLC trace of pure **1b**.

Reduced form of Cter M (2b):

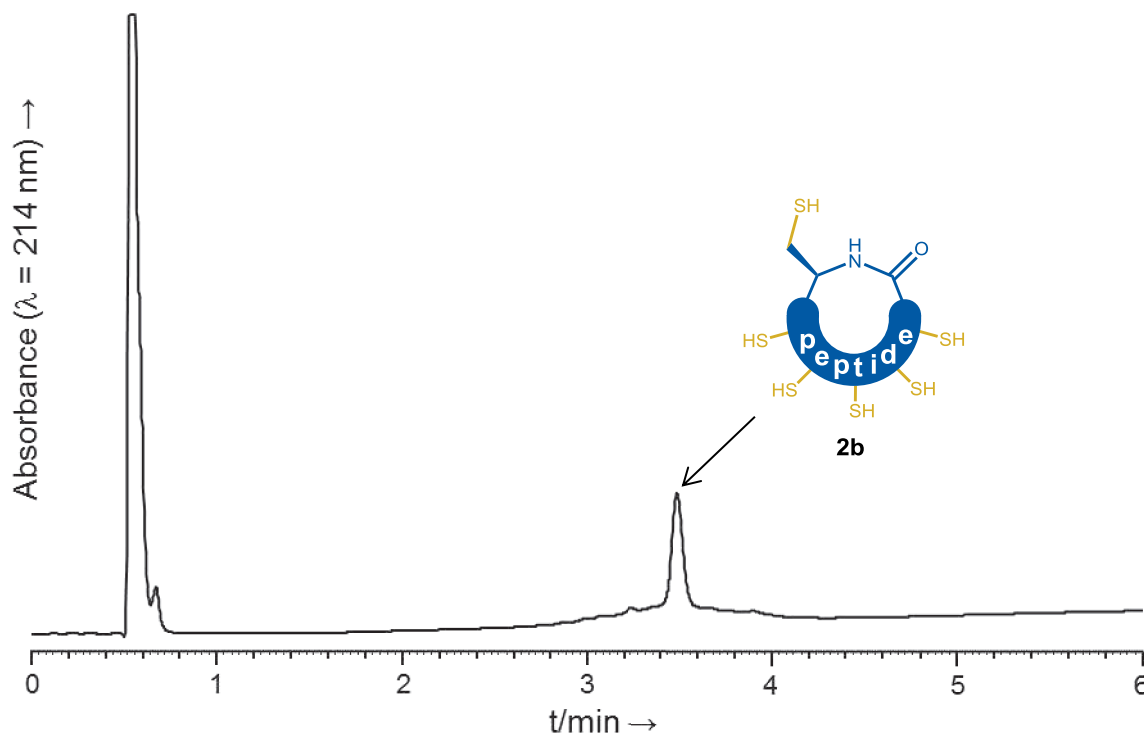
Sequence: cyclo(CSWPICMKNGLPCTCGETCTLGTCYVPDCS).

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{128}H_{199}N_{32}O_{41}S_7$ : 3064.2516, found: 3064.2467.

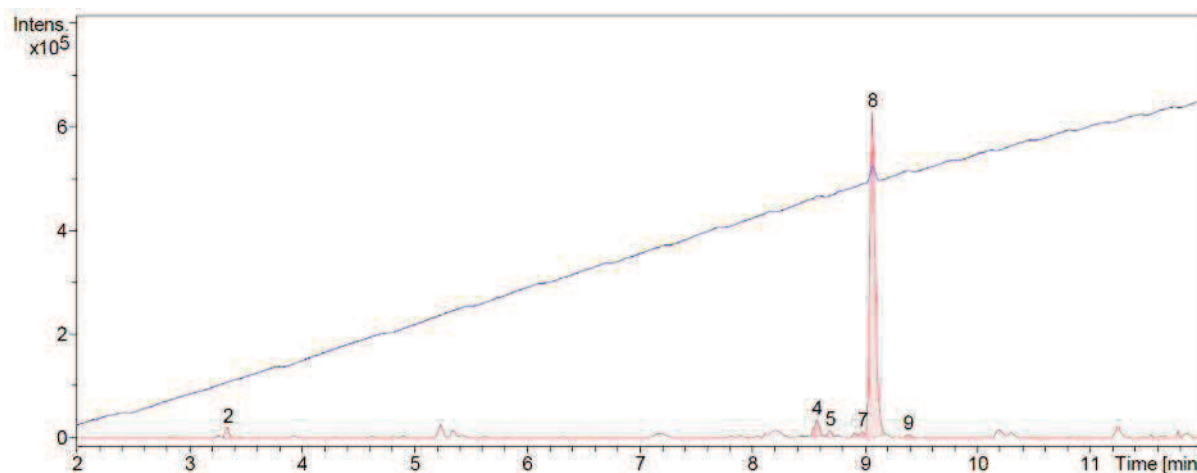
**HPLC analysis**:  $t_R = 3.49$  min (Chromolith, gradient: 20-60% B/A over 5 min).

**HPLC purification**: Nucleosil, gradient: 35-43% B/A over 8 min.

**Yield**: 94%.



Supplementary figure S21: HPLC trace of crude **2b** after work-up.

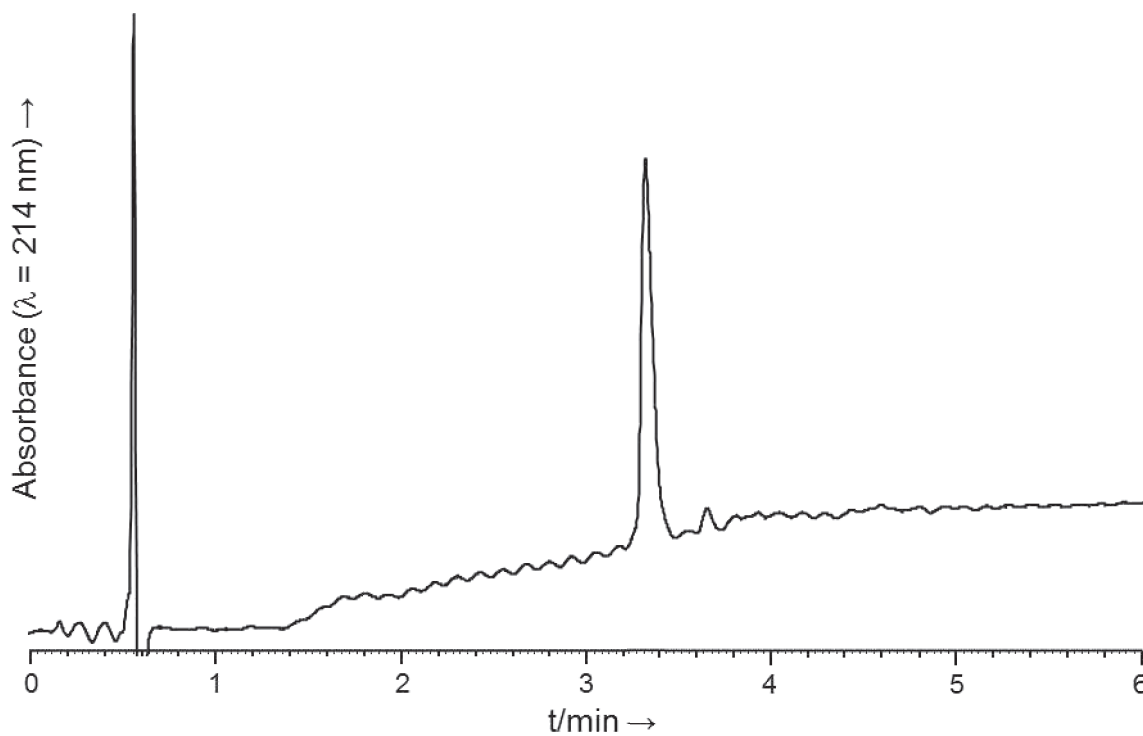


Supplementary figure S22: LC/MS analysis of crude **2b** after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.



Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
4 (8.57)	3080.2465	3080.2385	<b>2b</b> with oxidized methionine
8 (9.06)	3064.2516	3064.2467	<b>2b</b>

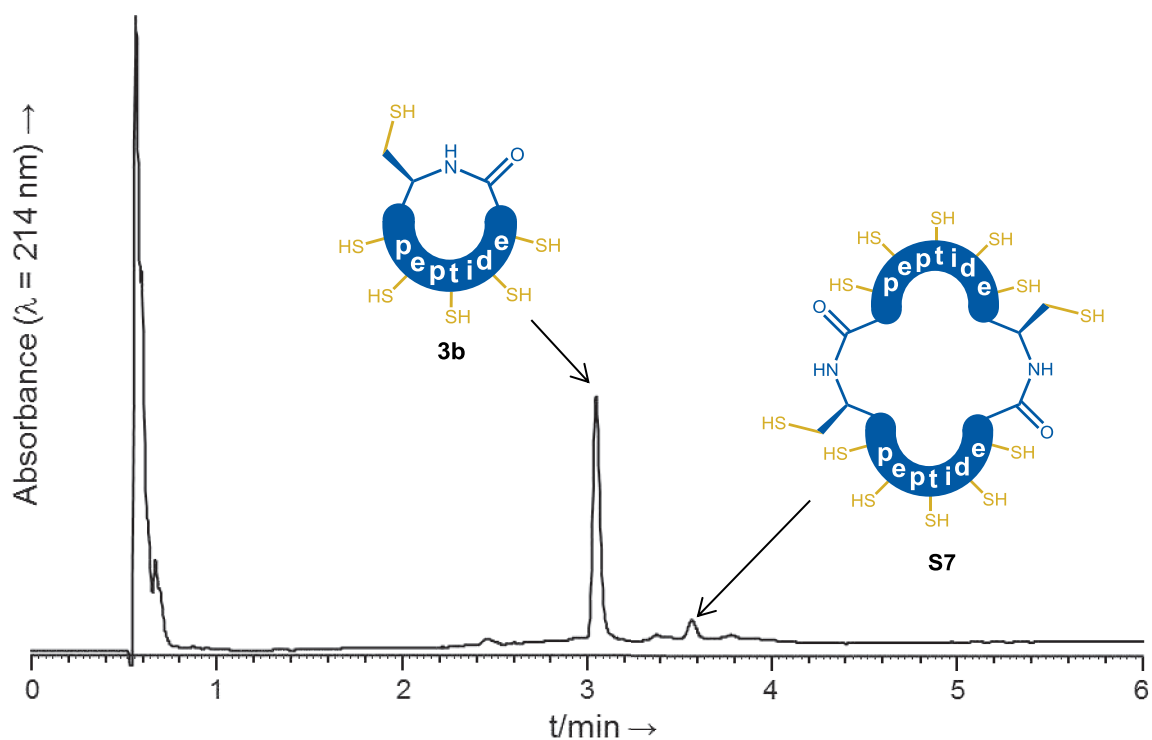
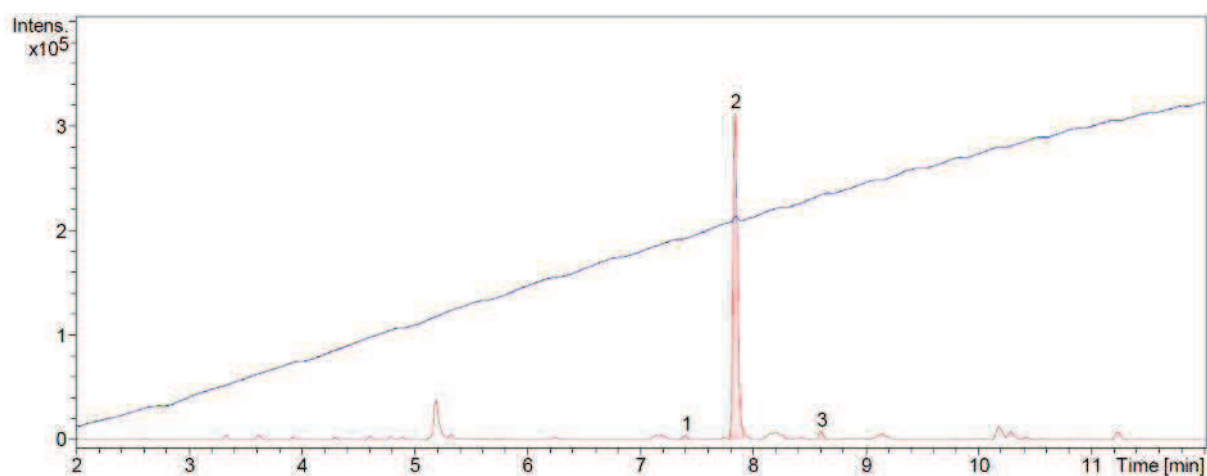
Supplementary table S9: Attribution of the main peaks observed during LC/MS analysis of crude **2b** after work-up.



Supplementary figure S23: HPLC trace of purified **2b**.

Reduced form of Kalata B1 (**3b**):

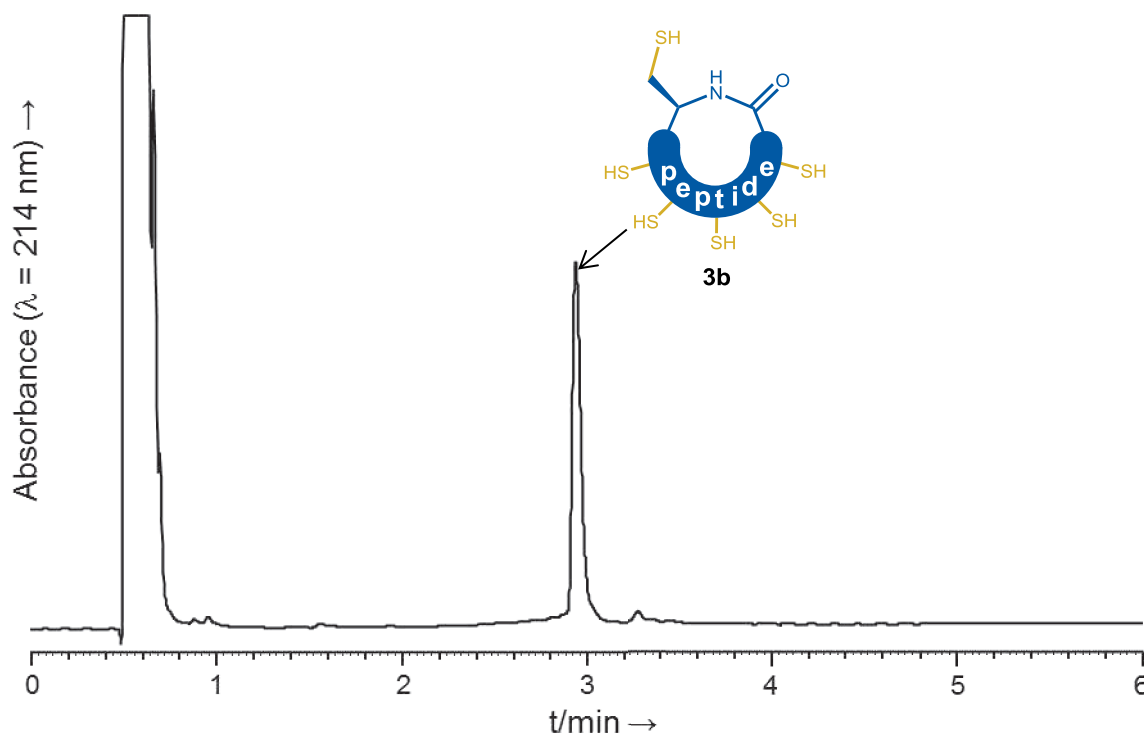
Sequence: cyclo(CTCSWPVCTRNGLPVCGETCVGGTCNTPG).

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{117}H_{186}N_{35}O_{39}S_6$ : 2897.1976, found: 2897.1973.**HPLC analysis**:  $t_R$  = 3.04 min (Chromolith, gradient: 20-50% B/A over 5 min).**HPLC purification**: Nucleosil, gradient: 28-38% B/A over 10 min, 70°C.**Yield**: 85%.Supplementary figure S24: HPLC trace of crude **3b** (NCL at 1 mM) after work-up.Supplementary figure S25: LC/MS analysis of crude **3b** (NCL at 1 mM) after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

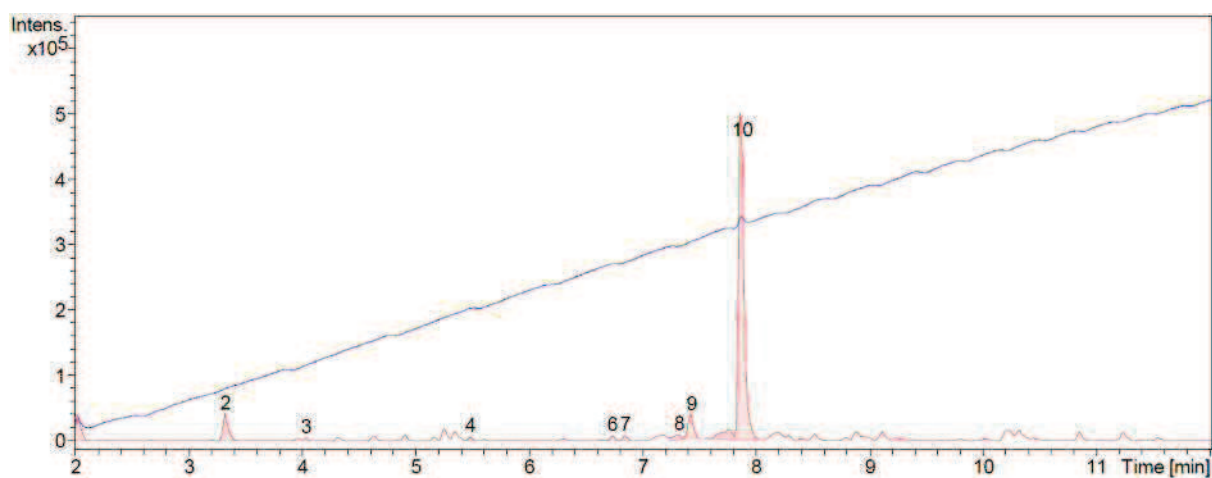
Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
2 (7.85)	2897.1976	2897.1973	<b>3b</b>
3 (8.60)	5793.3874	5793.3902	Cyclic dimer <b>S7<sup>a</sup></b>

Supplementary table S10: Attribution of the main peaks observed during LC/MS analysis of crude **3b** (NCL at 1 mM) after work-up.

<sup>a</sup>: **S7** = cyclo(CTCSWPVCTRNGLPVCGETCVGGTCNTPGCTCSWPVCTRNGLPVCGETCVGGTCNTPG).



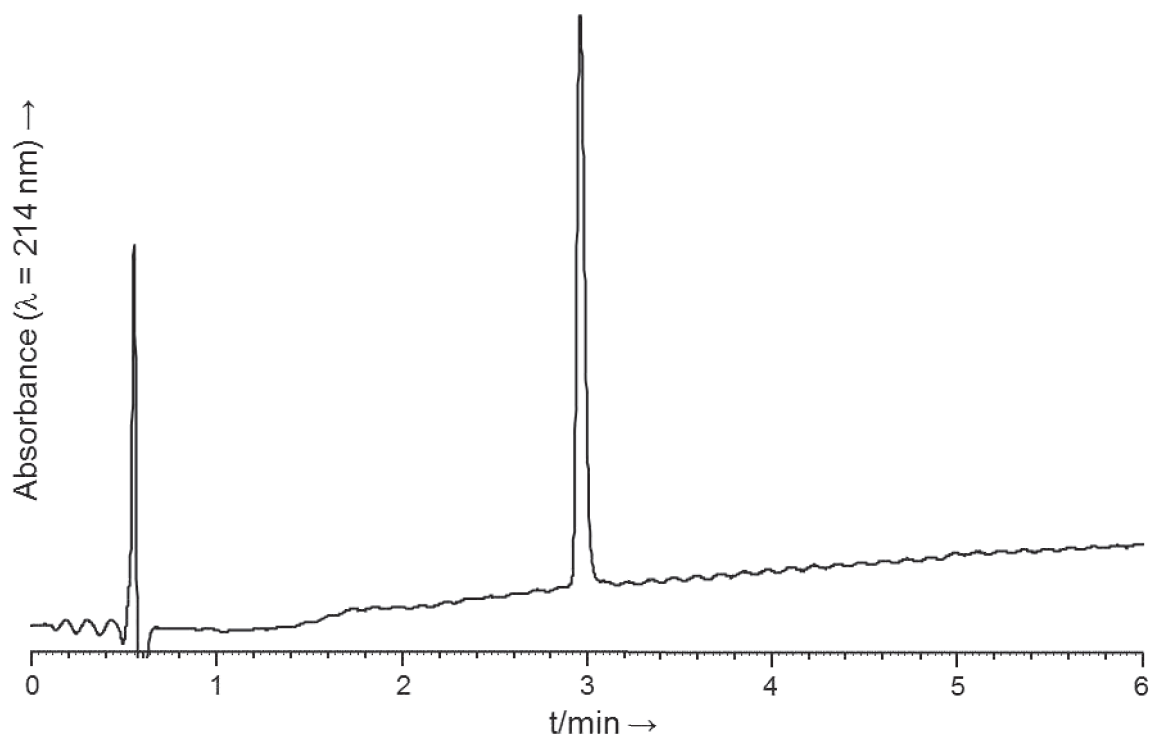
Supplementary figure S26: HPLC trace of crude **3b** (NCL at 0.1 mM) after work-up.



Supplementary figure S27: LC/MS analysis of crude **3b** (NCL at 0.1 mM) after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
10 (7.85)	2897.1976	2897.1952	<b>3b</b>

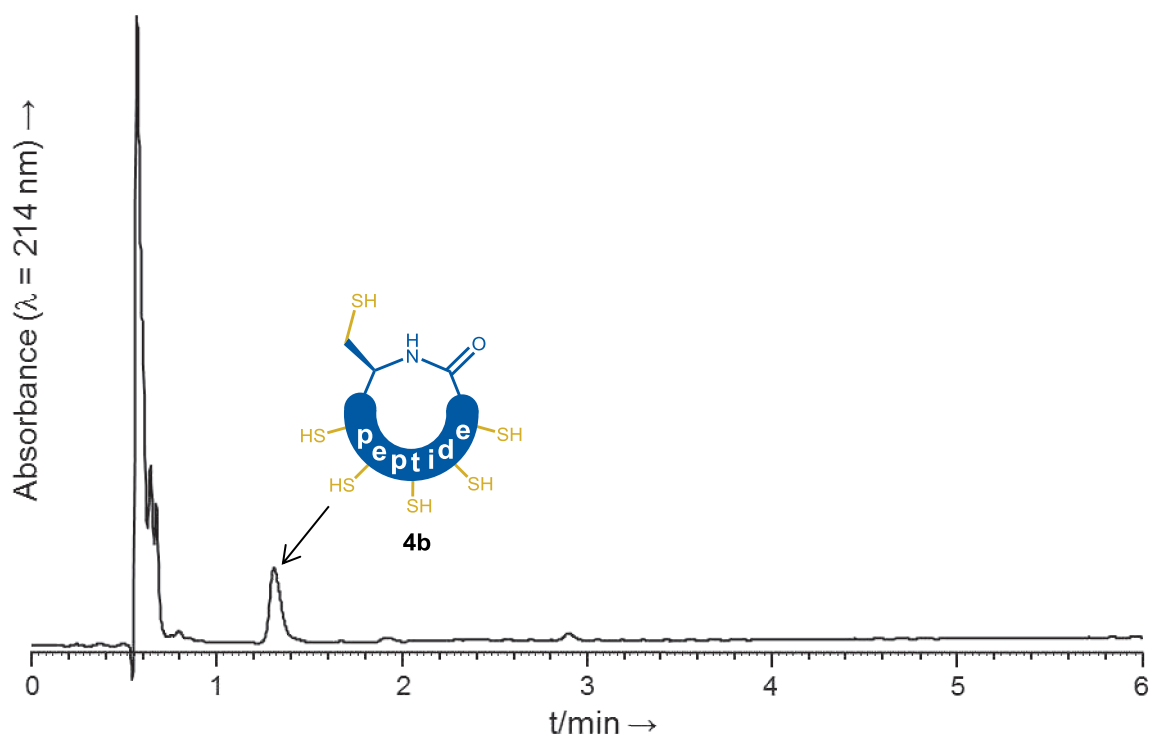
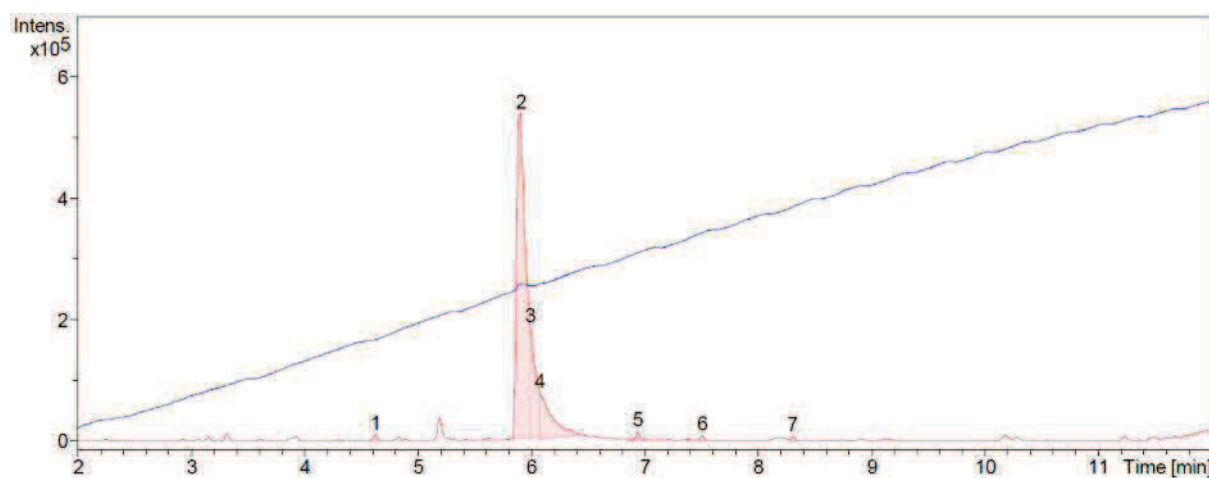
Supplementary table S11: Attribution of the main peak observed in LC/MS analysis of crude **3b** (NCL at 0.1 mM) after work-up.



Supplementary figure S28: HPLC trace of purified **3b**.

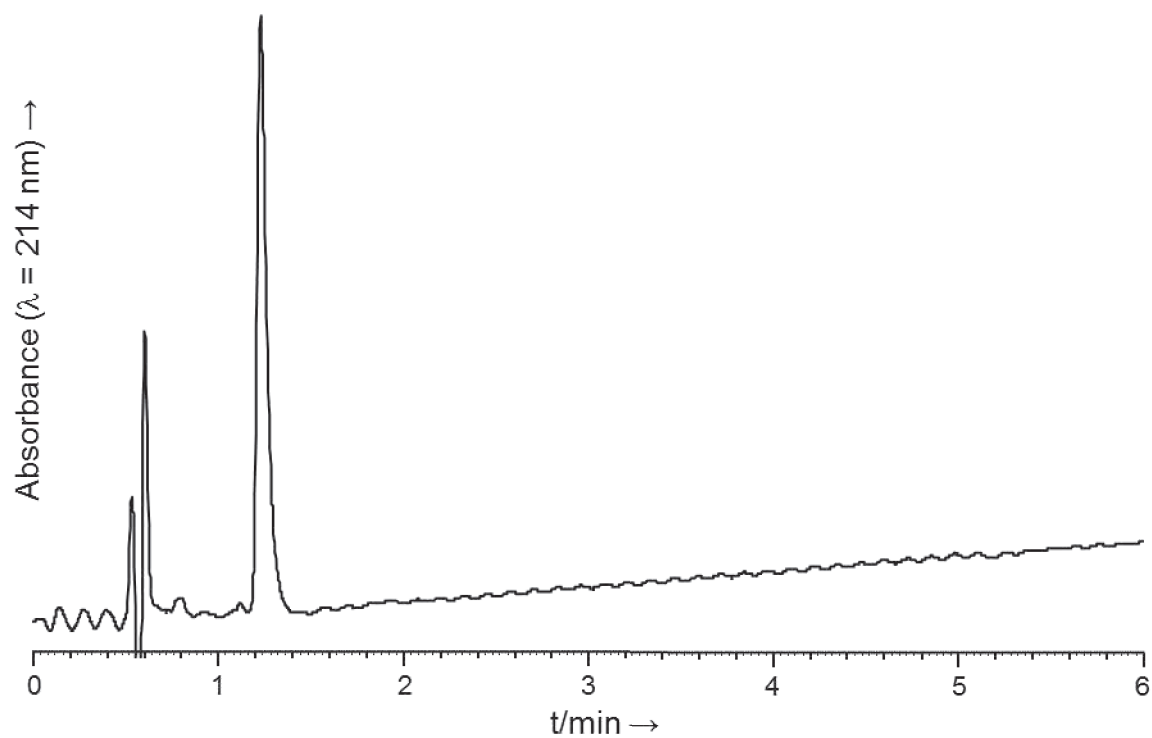
Reduced form of RTD-1 (4b):

Sequence: cyclo(CICTRGFCRCLCRRGVCR).

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{82}H_{144}N_{33}O_{19}S_6$ : 2086.9641 found: 2086.9637.**HPLC analysis:**  $t_R = 1.31$  min (Chromolith, gradient: 25-45% B/A over 5 min).**HPLC purification:** Nucleosil, gradient: 25-35% B/A over 10 min.**Yield:** 90%.Supplementary figure S29: HPLC trace of crude **4b** after work-up.Supplementary figure S30: LC/MS analysis of crude **4b** after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
2 / 3 / 4 (5.91-6.07)	2086.9641	2086.9637	<b>4b</b>

Supplementary table S12: Attribution of the main peaks observed during LC/MS analysis of crude **4b** after work-up.



Supplementary figure S31: HPLC trace of purified **4b**.

Reduced form of SFTI-1 (5b):

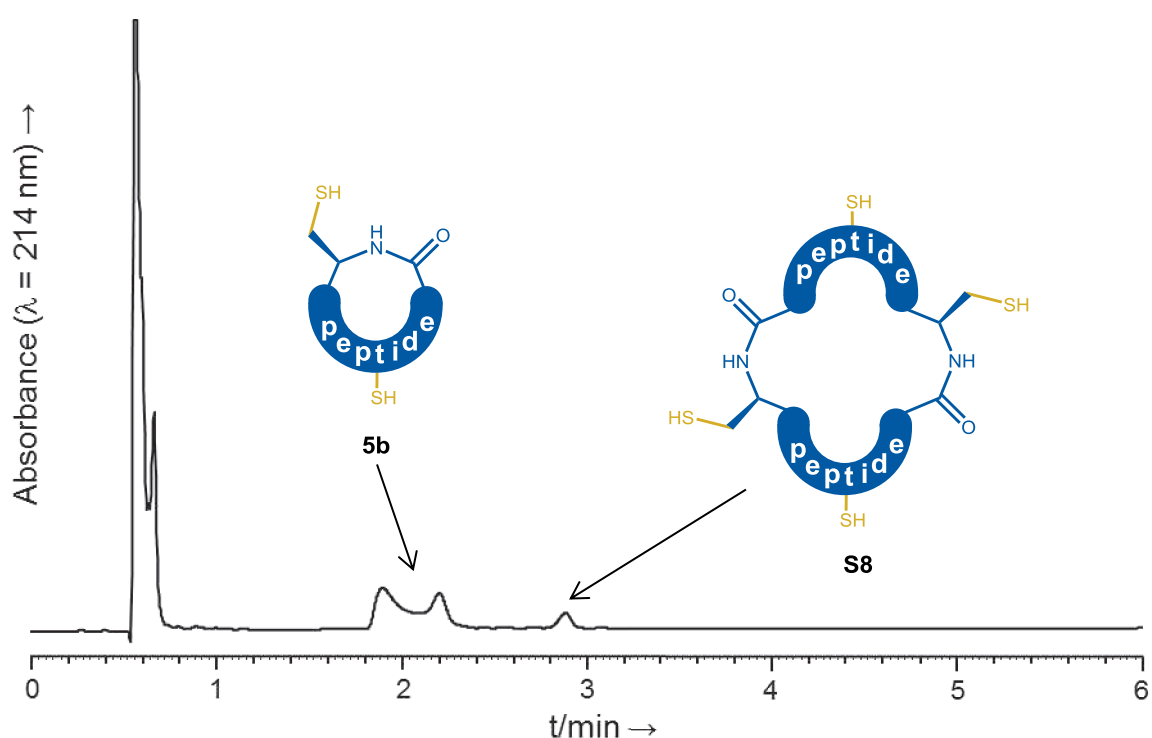
Sequence: cyclo(CTKSIPPICFPDGR).

**ESI-HRMS** ( $m/z$ ):  $[MH]^+$  calcd. for  $C_{67}H_{107}N_{18}O_{18}S_2$ : 1515.7452 found: 1515.7450.

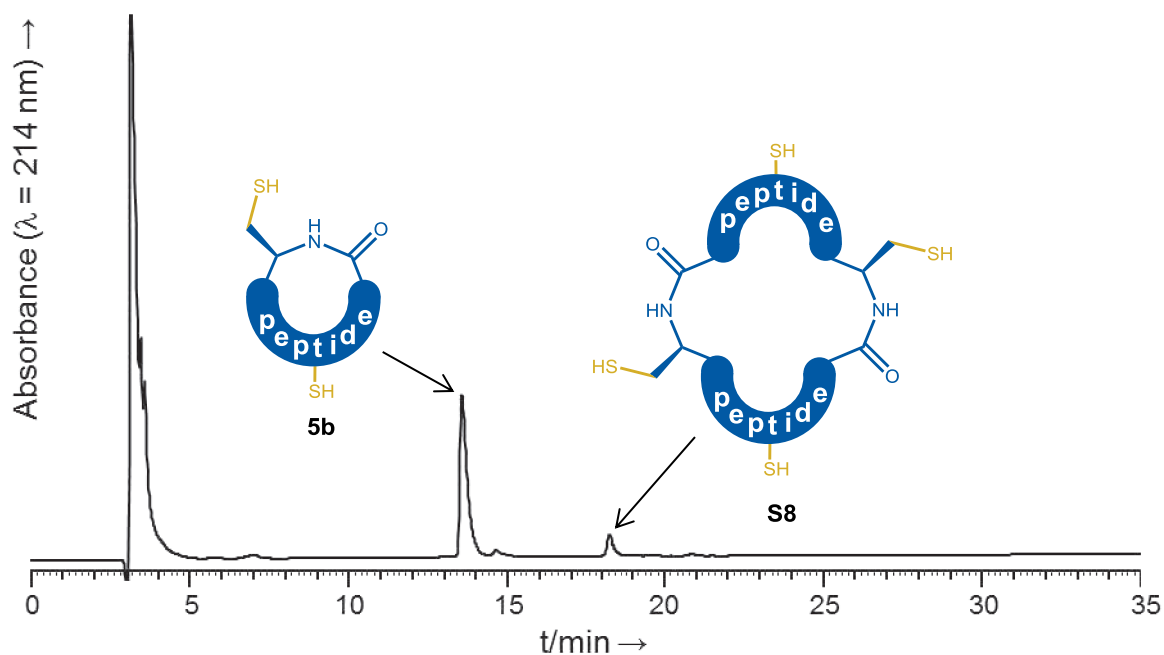
**HPLC analysis:**  $t_R$  = 1.81-2.31 min (Chromolith, gradient: 25-45% B/A over 5 min).  $t_R$  = 13.72 min (Nucleosil, gradient: 25-45% B/A over 30 min, 70 °C). HPLC analysis was performed at high temperature due to the presence of large peaks when analysing **5b** on a chromolith column at room temperature, probably due to an equilibrium between two conformers of **5b**.

**HPLC purification:** Nucleosil, gradient: 28-38% B/A over 10 min, 70°C.

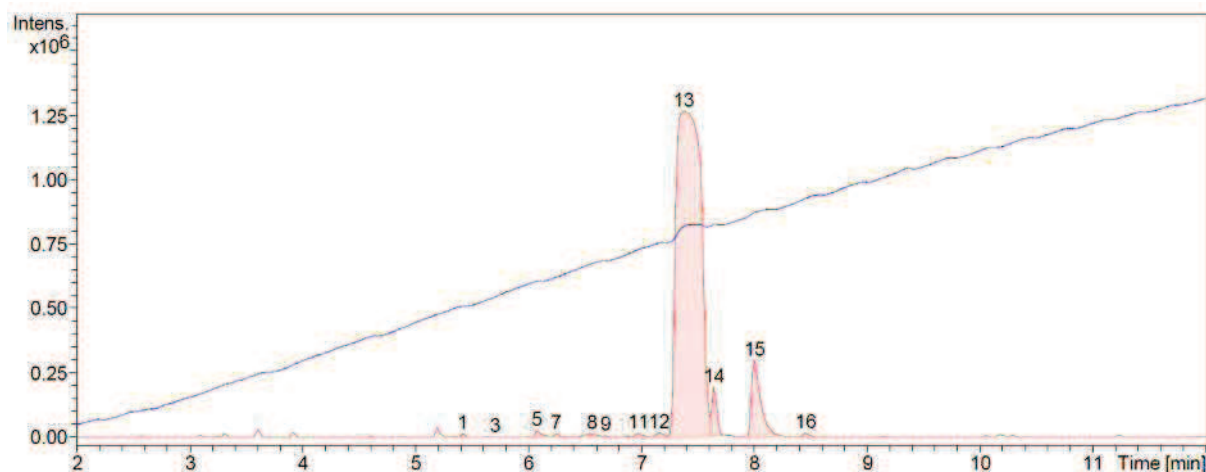
**Yield:** 91%.



Supplementary figure S32: HPLC trace of crude **5b** (NCL at 1 mM) after work-up (Chromolith, gradient: 25-45% B/A over 5 min).



Supplementary figure S33: HPLC trace of crude **5b** (NCL at 1 mM) after work-up (Nucleosil, gradient: 25-45% B/A over 30 min, 70 °C).



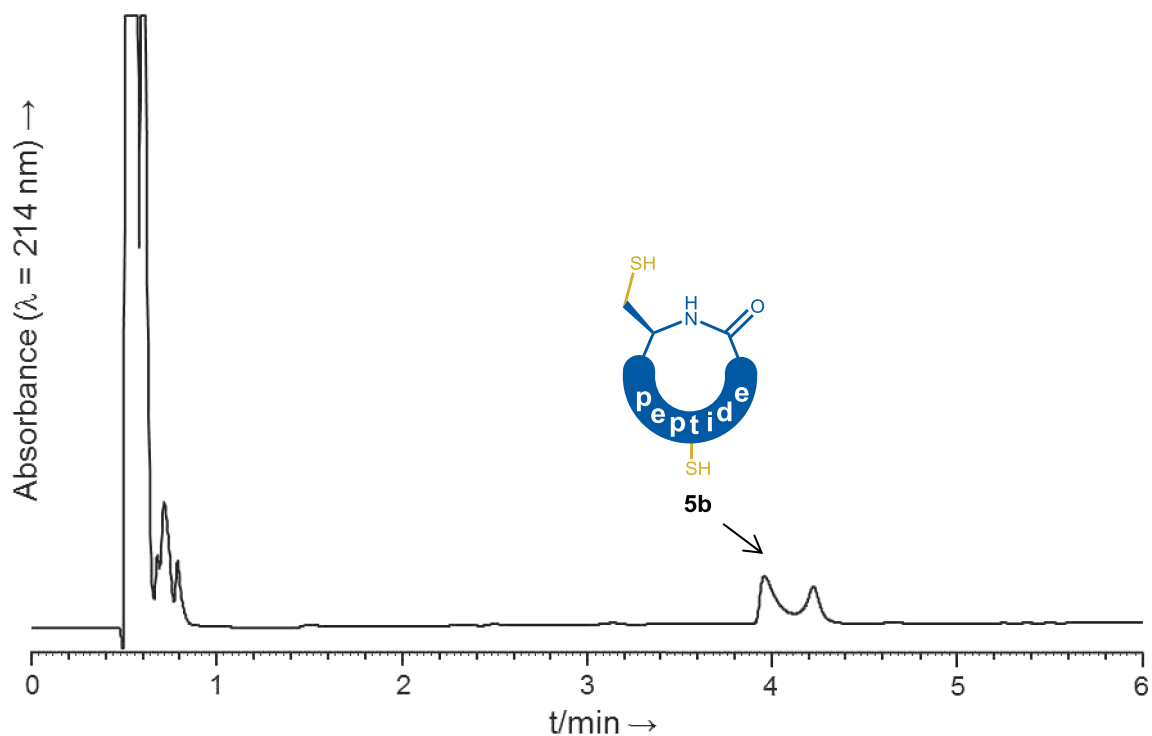
Supplementary figure S34: LC/MS analysis of crude **5b** (NCL at 1 mM) after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak ion chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
13 (7.37)	1515.7452	1515.7450	<b>5b</b>
14 (7.64)	1497.7436	1497.7348	Aspartimide derivative of <b>5b</b>
15 (8.00)	3030.4825	3030.4858	Cyclic dimer <b>S8</b> <sup>a</sup>

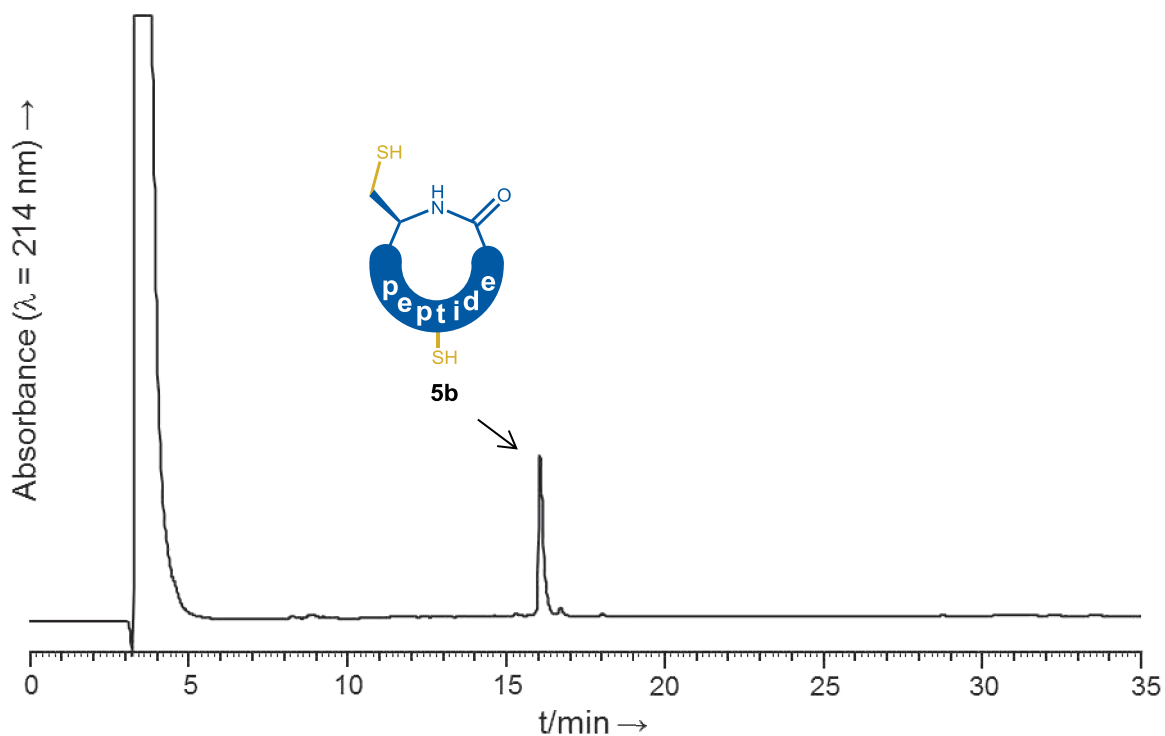
Supplementary table S13: Attribution of the main peaks observed in LC/MS analysis of crude **5b** (NCL at 1 mM) after work-up.

<sup>a</sup>: **S8** = cyclo(CTKSIPPICFPDGRCTKSIPPICFPDGR)

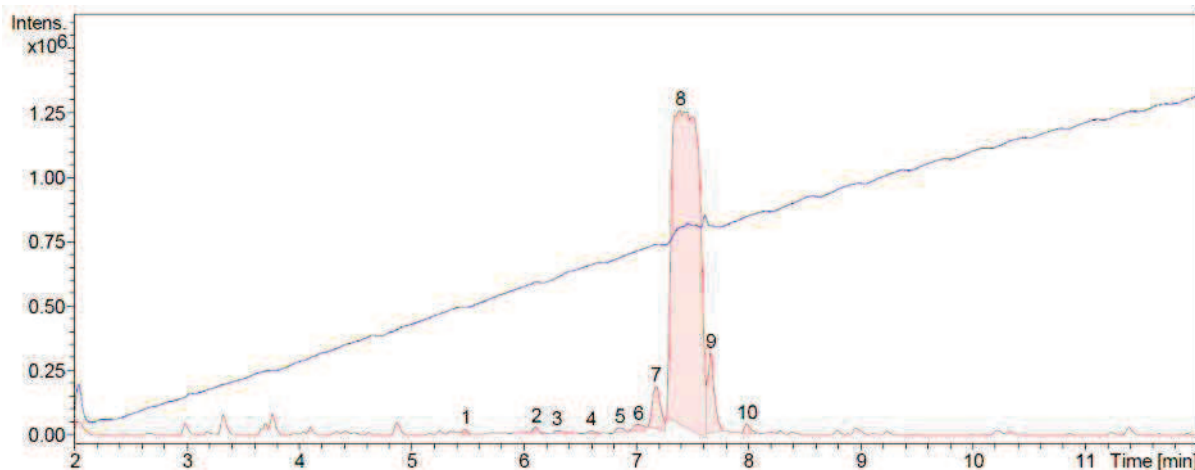




Supplementary figure S35: HPLC trace of crude **5b** (NCL at 0.1 mM) after work-up (Chromolith, gradient: 15-40% B/A over 5 min).



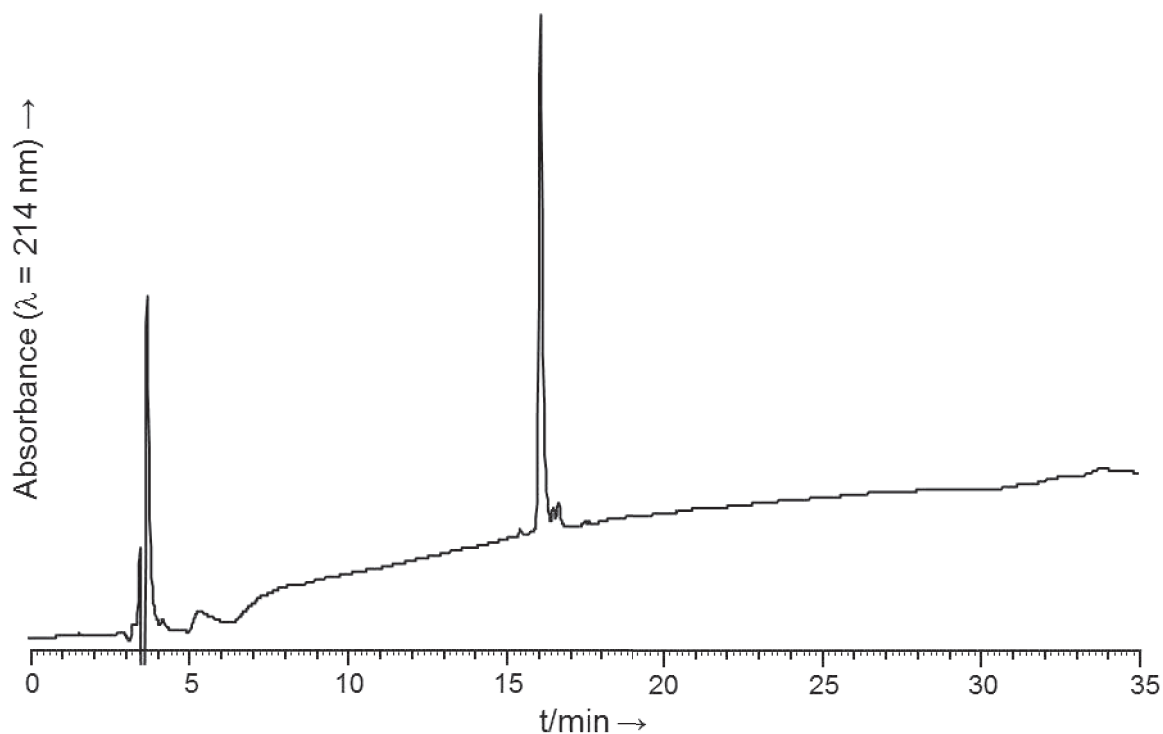
Supplementary figure S36: HPLC trace of crude **5b** (NCL at 0.1 mM) after work-up (Nucleosil, gradient: 20-60% B/A over 30 min, 70 °C).



**Supplementary figure S37:** LC/MS analysis of crude **5b** (NCL at 0.1 mM) after work-up. Blue trace: UV ( $\lambda = 214$  nm); red trace: base peak chromatogram.

Peak number ( $t_R$ (min))	$[MH]^+$ ( $m/z$ ) calcd.	$[MH]^+$ ( $m/z$ ) found	Attributed to
8 (7.39)	1515.7452	1515.7440	<b>5b</b>
9 (7.66)	1497.7436	1497.7350	Aspartimide derivative of <b>5b</b>
10 (7.98)	3030.4825	3030.4827	Cyclic dimer <b>S8</b>

**Supplementary table S14:** Attribution of the main peaks observed in LC/MS analysis of crude **5b** (NCL at 0.1 mM) after work-up.



**Supplementary figure S38:** HPLC trace of purified **5b**.

# ANNEXE



## Annexe : Note technique

### Fmoc-based synthesis of *N*-(2-hydroxy-5-nitrobenzyl)cysteine peptides as efficient thioester surrogates for native chemical ligation.

Victor P. Terrier, Agnès F. Delmas, Vincent Aucagne

**Running Head:** *N*-Hnb-Cys peptides: synthesis and use as crypto-thioesters for NCL

#### Summary

Peptide thioesters play a prominent role in the chemical synthesis of proteins, using strategies based on the convergent coupling of unprotected peptide segments. The native chemical ligation reaction (NCL) has revolutionized the field, since it provides access to native proteins by coupling peptide C $\alpha$ -thioesters with N-terminal cysteine-containing segments. However, the synthesis of peptide thioesters by Fmoc-based SPPS remains a real bottleneck to expand the limits of chemical protein synthesis. In this chapter, we describe the fully automatable Fmoc-based synthesis of *N*-(2-hydroxy-5-nitrobenzyl)cysteine (*N*-Hnb-Cys) peptide thioester precursors, which can be directly used in NCL reactions as “crypto-thioesters”. This strategy does not require any post-synthetic chemical step reaction and is extremely simple to implement. The herein described synthetic protocols enable non-experts to perform the entire process for the chemical synthesis of small proteins *via* the NCL of two segments.

**Key words:** Fmoc SPPS; thioester; crypto-thioester; NCL; ligation; chemical protein synthesis.

#### 1. Introduction

The discovery of the native chemical ligation (NCL) reaction has constituted a major advance in the field of chemical protein synthesis (1); this reaction allows obtaining very long peptide sequences by chemoselectively coupling peptide C $\alpha$ -thioester with N-terminal-cysteine containing peptides. While the synthesis of cysteinyl peptides is straightforward, the synthesis of peptide thioesters is complicated when using the widely used Fmoc-based solid phase peptide synthesis (SPPS) strategy. These difficulties arise from the reactivity of the thioester function towards piperidine treatments repeatedly used to remove the Fmoc group during peptide chain

elongation. Most Fmoc-based methodologies developed so far for peptide thioesters synthesis rely in a post-SPPS conversion of a piperidine-stable precursor into a thioester, but none of these methods is universal. Conception of widely-applicable and simple alternatives is still a major challenge.

Numerous strategies based on  $\beta$ - or  $\gamma$ -mercapto amide thioesterification devices have recently emerged (2-13); these methods exploit a post-SPPS amide-to-thioester rearrangement of a piperidine-stable precursor through an intramolecular  $N \rightarrow S$  acyl shift (Fig. 1). The mechanism of this rearrangement is reminiscent to the first step of *in vivo* protein splicing promoted by inteins. In most cases, the  $N \rightarrow S$  shift requires acidic conditions to proceed, and thus the peptides have to be first converted into thioesters then purified before being engaged into an NCL reaction. Only a few systems can operate at neutral pH (NCL conditions) thus enabling their direct use in NCL: these peptide thioester precursors are called crypto-thioesters (5-6, 8-9, 11-13).

In this chapter, we detail the synthesis and the use of an  $N$ -(2-hydroxy-5-nitrobenzyl)cysteine ( $N$ -Hnb-Cys) intramolecular thioesterification device for the synthesis of peptide crypto-thioesters (Fig. 2), a simple and general methodology that we recently introduced (14).

From a technical point of view, the device can be readily synthesized on solid support, including a key solid-supported reductive amination step to introduce the  $N$ -Hnb group onto a Cys(*StBu*)-Gly dipeptide (Fig. 3). Importantly, the phenol group of the device provides assistance for its  $N$ -acylation when coupling the C-terminal amino acid of the target sequence, and thus greatly facilitating this step. This allows for full automation of the process: device synthesis, introduction of the C-terminal amino acid and further Fmoc-SPPS elongation.

During each coupling, the Hnb hydroxyl group can react with activated Fmoc-amino acid to form an ester that is cleaved during the next piperidine treatment. This is similar to the synthesis of peptides containing an  $N$ -Hmb (2-hydroxy-4-methoxybenzyl) or related groups used to prevent aggregation during SPPS (15, 16). In our hands, this competing  $O$ -acylation did not raise any drop in yield during peptide elongation under the SPPS conditions we currently use, as compared to peptide sequences not incorporating an  $N$ -Hnb-Cys device. However, it interferes with the monitoring of peptide elongation through UV titration of the fluorenylmethylpiperidine byproduct of the Fmoc deprotection. We also found that it can inhibit the acetylation-mediated capping step when the coupling of the C-terminal amino acid of the sequence is not quantitative, as observed with sterically-hindered side chains (*see* section 3.2.1, part 5). Utilizing conditions developed by Bradley *et al.* (17) for the selective cleavage of a Dde protective group in the presence of a Fmoc carbamate, we bypassed these problems, and set up an automated hydroxylamine-based treatment that allows selective cleavage of Hnb esters without affecting

the Fmoc group. This simple treatment ensures complete acetylation of any unreacted secondary amine of the device, and allows performing accurate UV quantification of the elongation yield based on the deprotection of the final N-terminal residue.

Our process yields peptide crypto-thioesters that can be conveniently handled and purified using standard procedures, due to the use of a *StBu* protecting group for the cysteine residue of the device. Such a disulfide is stable to peptide cleavage/deprotection, analysis and purification conditions and thus prevents any unwanted premature *N*→*S* shift. The *StBu* protective group is readily removed within minutes under NCL conditions, allowing the *in situ* conversion into an active crypto-thioester.

We herein describe simple, optimized protocols for the Fmoc-based synthesis of *N*-Hnb-Cys peptide and their use as crypto-thioesters under standard NCL conditions. This should allow non-specialists to perform the entire process for the synthesis of cysteine-containing protein sequences through the ligation of two peptide segments.

## 2. Materials

All commercially available reagents and solvents are used without further purification.

Prepare all aqueous solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of 18 M  $\Omega$  cm at 25°C) and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing waste materials.

### 2.1. Solid phase peptide synthesis

1. Resin for peptide synthesis. note that the protocols described here for the synthesis of *N*-Hnb-Cys peptide crypto-thioesters were optimized for an aminomethyl TentaGel R resin (Rapp polymers) but could very probably be used for other commercially-available solid supports.

2. Fmoc-Rink's linker and standard Fmoc-protected L-amino acid derivatives : Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Phe-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Pro-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, Fmoc-Trp(Boc)-OH , Fmoc-Tyr(*t*Bu)-OH, Fmoc-Val-OH , and Fmoc-Cys(*St*Bu)-OH as a special building block.

3. *N*-Methyl-2-pyrrolidone (NMP).

4. Dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>).

5. Amino acid solutions: 0.125 M Fmoc-protected amino acid in NMP. Store at -20 °C and use within a week delay.
6. Piperidine solution: 20 % piperidine in NMP (v/v).
7. Activator solution: 0.475 M HCTU in DMF. In a 100 mL graduated cylinder, weigh 19.65 g of HCTU, complete to 100 mL with DMF (peptide synthesis grade). Add a magnetic stirring bar and vigorously stir until complete dissolution. Store at 4 °C and use within a week delay.
8. Base solution: 1 M DIEA in NMP. Mix 17.4 mL diisopropylethylamine (DIEA) and 82.6 mL NMP.
9. Acetylation solution: 0.5 M Ac<sub>2</sub>O, 0.13 M DIEA and 0.013 M HOBt in NMP. In a 100 mL graduated cylinder, weigh 200 mg of 1-hydroxybenzotriazole hydrate (HOBt), add 30-40 mL NMP and sonicate until complete dissolution. Add 4.75 mL acetic anhydride, 2.25 mL DIEA then NMP up to a 100 mL final volume. Store at 4 °C and use within a two days delay.
10. Peptide synthesizer (the work described here has been performed on a Prelude synthesizer from Protein Technologies but can be implemented either manually or using other manufacturers instruments).

## 2.2. Solid-supported reductive amination solutions

1. Solution for resin washes: DMF/MeOH (1:1 v/v).
2. Resin conditioning solution: DMF/MeOH/AcOH (9:9:2 v/v/v).
3. Aldehyde solution: 0.125 M 2-hydroxy-5-nitrobenzaldehyde in DMF/MeOH (1:1 v/v). Weigh 209 mg of 2-hydroxy-5-nitrobenzaldehyde in a graduated cylinder and complete to 10 mL final volume with DMF/MeOH (1:1 v/v). Sonicate or vortex until complete dissolution.
4. Reducing agent solution: 0.25 M sodium cyanoborohydride (NaBH<sub>3</sub>CN) in DMF/ MeOH/AcOH (9:9:2 v/v/v). Just before use, weigh 157 mg of NaBH<sub>3</sub>CN in a graduated cylinder and complete to 10 mL final volume with DMF/MeOH/AcOH (9:9:2 v/v/v). Sonicate or vortex until complete dissolution. (*see Note 1*).

## 2.3. Hydroxylamine-based reagent

0.3 M Hydroxylamine hydrochloride, 0.225 M imidazole in NMP/CH<sub>2</sub>Cl<sub>2</sub> (5:1 v/v). Weigh 12.5 g of hydroxylamine hydrochloride and 9.18 g of imidazole in a conical flask. Add 50 mL of NMP and stir vigorously with a magnetic stirring bar or sonicate until complete dissolution. The dissolution process is rather slow (typically 1-3 h). This stock solution can be stored at -20°C for several weeks. Just before use, add 1 volume CH<sub>2</sub>Cl<sub>2</sub> for 5 volume stock solution.

## 2.4. TFA-based cleavage cocktail solution and peptide precipitation solvents



1. Trifluoroacetic acid (TFA)-based cleavage cocktail: TFA/PhOH/water/*i*Pr<sub>3</sub>SiH (88:5:5:2 v/w/v/v).
2. Rotary evaporator.
3. Mixture of diethyl ether/petroleum ether (1:1 v/v) for precipitation.
4. Diethyl ether for further peptide wash.
5. Centrifuge.
6. Vacuum pump.

## 2.5. HPLC analysis and semi-preparative purification

1. Solvent A: water / TFA (99.9:0.1 v/v).
2. Solvent B: acetonitrile / TFA (99.9:0.1 v/v).
3. HPLC system equipped with reverse phase C<sub>18</sub> analytical and semi-preparative columns.
4. Rotary evaporator.
5. Freeze-dryer.

## 2.6. NCL

1. Phosphate and guanidine hydrochloride buffer: 200 mM di-sodium hydrogen phosphate (Na<sub>2</sub>HPO<sub>4</sub>), 6 M guanidine hydrochloride (Gu.HCl) in water. In a 10 mL graduated cylinder, weigh 356 mg Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O and 5.7g Gu.HCl, add ultra-pure water up to 10 mL. (*see Note 2*). Transfer the solution in a round bottom flask sealed with a rubber septum. This solution can be stored at -20°C for several months.
2. 4-Mercaptophenylacetic acid (MPAA)
3. *Tris*-(2-carboxyethyl)phosphine (TCEP)
4. NaOH as a fine powder. (*see Note 3*).
5. Argon cylinder.
6. Quenching solution for analyzing aliquots: water / TFA (98.5:1.5 v/v).
7. 5 µl microsyringe
8. Quenching solution: water / acetonitrile / acetic acid (79:20:1).
9. Diethyl ether.
10. System for deoxygenation of the stock solutions and NCL reaction media. We conveniently use in our lab four successive vacuum degassing/argon saturation cycles, based on a simple three-way stopcock-based apparatus connecting the solution to be deoxygenated to an argon-filled balloon and a 1-20 mbar membrane vacuum pump (1-20 mbars). (*see Note 4*).

## 3. Methods

We briefly describe here the SPPS protocols currently in use in our laboratory, but other can be used. For general details about SPPS, refer to *Methods in Molecular Biology* Volume 298 chapter by Subra *et al.* (16).

### 3.1. Cysteinyl peptide synthesis

#### 3.1.1 Fmoc-SPPS (Fig. 4)

1. Place 25  $\mu\text{mol}$  of a resin preloaded with the first amino acid in a SPPS reactor and swell it in NMP.

2. Peptide elongation:

a) Coupling: add 2 mL of the appropriate amino acid solution (sol. **2.1.5**), 500  $\mu\text{L}$  activator solution (sol. **2.1.7**), 500  $\mu\text{L}$  base solution (sol **2.1.8**), mix for 30 min and then remove solution by filtration. Wash thoroughly with NMP.

b) Capping: add 3 mL acetylation solution (sol. **2.1.9**), mix for 7 min and then remove solution by filtration. Wash thoroughly with NMP.

c) Fmoc removal: add 3 mL piperidine solution (sol. **2.1.6**), mix for 3 min and then remove solvent by filtration ( $\times 3$ ). Wash thoroughly with NMP.

d) Repeat this deprotection/amino acid coupling/capping cycle to perform the complete elongation of the target sequence.

3. Peptide cleavage and deprotection:

a) Wash thoroughly the resin with  $\text{CH}_2\text{Cl}_2$  and further let dry the resin to remove the maximum of  $\text{CH}_2\text{Cl}_2$ . Add 5 mL TFA cocktail solution (sol. **2.4.1**) and stir gently the resin by rotation for 3h. (*see Note 5*).

b) Collect the resulting solution in a 50 mL round bottom flask, wash the resin with pure TFA ( $3 \times 5 \text{ mL}$ ) and add the wash solutions to the flask. Concentrate the solution by evaporating the TFA on a rotary evaporator down to a 2-4 mL final volume.

c) Add the solution to 40 mL of an ice-cold precipitation solution (sol. **2.4.3**) into a conical polypropylene tube to precipitate the peptide.

d) Centrifuge the conical tube (3000 G, 3 min) and remove the supernatant.

e) Add 40 mL diethyl ether, sonicate or vortex upon obtaining a homogeneous suspension, then centrifuge the tube and remove the supernatant. ( $\times 3$ )

f) Dry the crude solid product by removing remaining diethyl ether using a vacuum pump.

#### 3.1.2 Analysis and purification

As the HPLC analysis and purification depend on the peptide sequence and on the HPLC system and columns used, there is not any general method, adapt the gradient and column to the peptide sequence.

1. Analysis: dissolve 1 mg crude product in 1 mL solvent A / solvent B (sol. **2.5.1** & **2.5.2**) (8:2 v/v) and analyze by RP-HPLC and mass spectrometry.

2. Purification: purify the product using the HPLC system fitted with a semi-preparative column. Remove acetonitrile and TFA by concentration using a rotary evaporator and then remove water by freeze-drying.

### **3.2. (Hnb)Cys Peptide crypto-thioester synthesis** (Fig. 5)

3.2.1 Synthesis of the *N*-Hnb-Cys device and coupling of the C-terminal amino acid of the peptide sequence.

The following protocols are adapted to a 25  $\mu$ mol scale synthesis using aminomethyl TentaGel R resin, but they could be further adapted to other sources of aminomethyl resins.

1. Couple successively Rink's linker, Fmoc-Gly-OH and Fmoc-Cys(*St*Bu)-OH on the resin (*see* Subheading **3.1.1** part 2).

2. Formation of the acetic acid salt of the solid-supported amine:

a) DMF/MeOH wash: add 3 mL DMF/MeOH (sol. **2.2.1**), mix for 30 s and then remove solvent by filtration ( $\times 2$ ).

b) DMF/MeOH/AcOH wash: add 3 mL DMF/MeOH/AcOH (sol. **2.2.2**), mix for 5 min and then remove solvent by filtration.

c) DMF/MeOH wash: add 3 mL DMF/MeOH (sol. **2.2.1**), mix for 30 s and then remove solvent by filtration ( $\times 4$ ).

3. Formation of the iminium:

a) Add 2 mL 2-hydroxy-5-nitrobenzaldehyde solution (sol. **2.2.3**), mix for 1 h and then remove solvent by filtration.

b) DMF/MeOH wash: add 3 mL DMF/MeOH (sol. **2.2.1**), mix for 15 s and then remove solvent by filtration ( $\times 2$ ). (*see* **Note 6**).

4. Reduction:

a) Reduction: immediately add 2 mL of a freshly-prepared sodium cyanoborohydride solution (sol. **2.2.4**), mix for 1 h and then remove solvent by filtration.

b) DMF/MeOH wash: add 3 mL DMF/MeOH (sol. **2.2.1**), mix for 30 s and then remove solvent by filtration ( $\times 4$ ).

c) NMP wash: add 3 mL NMP, mix for 30 s and then remove solvent by filtration ( $\times 3$ ).

d) Piperidine wash: add 3 mL piperidine solution (sol. **2.1.6**), mix for 3 min and then remove solvent by filtration ( $\times 3$ ).

e) NMP wash: add 3 mL NMP, mix for 30 s and then remove solvent by filtration ( $\times 3$ ).

f) CH<sub>2</sub>Cl<sub>2</sub> wash: add 5 mL CH<sub>2</sub>Cl<sub>2</sub>, mix for 5 min, and then remove solvent by filtration ( $\times 3$ ).

g) NMP wash: add 3 mL NMP, mix for 30 s and then remove solvent by filtration ( $\times 2$ ).

5. Coupling of the C-terminal amino acid of the peptide sequence: couple the C-terminal amino acid of the desired sequence through standard protocol (*see* Subheading **3.1.1**, part **2a**). Repeat once the coupling step.

This should lead to 80-100 % coupling yield for most amino acid: Gly, Asp(*O*tBu), Ala, Glu(*O*tBu), Cys(*Trt*), Met, Asn(*Trt*), His(*Trt*), Phe, Gln(*Trt*) and Leu. In the case of Trp(*Boc*), Arg(*Pbf*), Tyr(*t*Bu), Ser(*t*Bu) and Lys(*Boc*), repeat once more the coupling step to achieve a >80 % coupling yield. For four more demanding residues, Thr(*t*Bu), Pro, Val, and Ile, prolonged coupling times or heating are needed to achieve an excellent yield. Note that the initial amount of resin can be adapted to take into account lower coupling yield. See reference 14 for the evaluation of the coupling efficiency of the 20 proteogenic amino acids.

6. Hydroxylamine treatment: add 3 mL hydroxylamine solution (sol. **2.3.**), mix for 20 min and then remove solvent by filtration ( $\times 3$ ). (*see* **Note 7**). Wash thoroughly with NMP, CH<sub>2</sub>Cl<sub>2</sub>, then NMP. This step is important for evaluating the coupling yield through UV titration of the fluorenylmethylpiperidine byproduct of the Fmoc deprotection, but also to ensure complete capping of any unreacted *N*-Hnb-Cys in case of uncomplete coupling.

7. Capping step: *see* Subheading **3.1.1**, part **2b**. Repeat once this capping step.

8. Fmoc removal: (*see* Subheading **3.1.1**, part **2c**). Eventually quantify the yield of the *N*-acylation by titration of the fluorenylmethylpiperidine byproduct. (*see* **Note 8**).

### 3.2.2 SPPS elongation of the crypto-thioester sequence

1. Perform a standard SPPS elongation of the target peptide sequence (*see* Subheading **3.1.1**, part **2**). Note that the final piperidine treatment has to be performed even when no Fmoc group is present on the N-terminal building block, for example when the N-terminal building block is a Boc-amino acid, acetic acid or pyroglutamic acid. As above mentioned, this treatment will readily cleave any *O*-Hnb ester that could contaminate the final product.

### 3.2.3 TFA treatment and precipitation

(*see* subheading **3.1.1**, part **3**). No specific protocol is needed to cleave and deprotect *N*-Hnb-Cys peptide crypto-thioesters.

### 3.2.4 Analysis and purification

(*see* subheading **3.1.2**). No specific protocols are needed for RP-HPLC analysis and purification of *N*-Hnb-Cys peptide crypto-thioesters. However, in rare cases, a broad or a double peak can be observed, that we attribute to slow *cis*-to-*trans* interconversion of the peptide-*N*-Hnb-Cys amide bond. Analysis and purification at high temperature (50-70°C) solves this problem. Note that the *St*Bu protecting group and the Hnb moiety of the *N*-Hnb device makes the peptide sequence more hydrophobic and thus more retained on RP-HPLC compared to the parent sequence. Also note that MALDI-TOF MS analysis will lead to laser-induced fragmentation of the nitro group while ESI-MS analysis does not.

### 3.3. Native chemical ligation (Fig. 6)

#### 3.3.1 Reaction preparation

To avoid oxidation of MPAA and peptides thiol groups, we recommend performing NCL under argon and using strictly deoxygenated solutions. This is not a special requirement for *N*-Hnb-Cys peptides, but rather a general recommendation for performing NCL reactions.

The representative experimental procedure described here uses a stoichiometric 2.5  $\mu\text{mol}$  amount of both cysteinyl peptides and *N*-Hnb-Cys crypto-thioester at a 5 mM concentration, but a slight excess of one of the partners (10-50%) can be desirable to help pushing up the reaction to completion within a few hours or overnight. Peptide concentrations of 1-10 mM usually work fine, the lower concentration leading to reactions typically completed overnight or after 48 h, depending on the C-terminal amino acid of the peptide thioester segment. (19).

Note that *N*-Hnb-Cys crypto-thioesters show slightly slower kinetics compared to those typically observed when using pre-formed thioesters due to their need to be converted in-situ into a reactive species.

1. Deoxygenate the phosphate and guanidine buffer (sol. 2.6.1). (see **Note 4**).
2. Freshly prepare the NCL buffer: 100 mM MPAA, 50 mM TCEP, 6 M Gu.HCl, 0.2 M phosphate buffer pH 6.5. In a plastic centrifuge tube, weigh accurately 16.8 mg MPAA, 14.3 mg TCEP.HCl and 10.3 mg NaOH. Seal with a rubber septum and gently flush with argon for 5 min. (see **Note 9**). Add 1 mL of deoxygenated phosphate and guanidine buffer.
3. In a plastic centrifuge tube weigh accurately 2.5  $\mu\text{mol}$  of cysteinyl peptide **1** and 2.5  $\mu\text{mol}$  of *N*-Hnb-Cys crypto-thioester peptide **2**. Seal the container with a rubber septum and gently flush with argon for 5 min. (see **Note 9**).
4. Under argon, transfer 500  $\mu\text{L}$  of the NCL buffer into the centrifuge tube containing the peptide segments, using a syringe fitted with a needle (see **Note 10**). The solution should immediately turn to yellow. Seal the rubber septum using a paraffin film, then sonicate or vortex the solution until complete dissolution of the peptides, eventually followed by centrifugation if the rubber septum has been contaminated by the yellow solution. Heat the plastic centrifuge tube at 37 °C.

#### 3.3.2 Reaction monitoring

Monitor the reaction using HPLC. Take aliquots at time intervals (typically after 3 h, 6 h, 24 h and 48 h) during the reaction following this procedure designed to strictly exclude oxygen contamination from the NCL reaction:

1. Connect the centrifuge tube to a gentle argon flow. (see **Note 10**).
2. In a HPLC vial fitted with a glass insert, add 108  $\mu\text{L}$  water + 1.5 % TFA (sol. 2.6.6).
3. Using a 5  $\mu\text{L}$  microsyringe, take a 2  $\mu\text{L}$  aliquot of the NCL solution through the rubber septum and dilute it into the acidic solution. (see **Note 11**).

4. Immediately analyze the sample *via* RP-HPLC. (*see* **Note 12**). Note that the *N*-Hnb-Cys(StBu) peptide will be converted into a faster eluting *N*-Hnb-Cys within minutes, an intermediate that can be identified through its typical Hnb UV signature (max UV absorption around 317 nm) if using a diode array detector or a 320 nm UV detector channel.

### 3.3.3 Reaction work-up and purification.

1. When the reaction is over (*i.e.* limiting reagent was consumed as judged by HPLC monitoring), remove the sample from 37 °C heating.
2. Dilute the reaction mixture with 1 mL of the water / acetonitrile / acetic acid quenching solution (sol. **2.6.7**).
3. Remove the rubber septum and transfer the solution into a 15 mL centrifugation conical tube containing 5 mL of the water / acetonitrile / acetic acid quenching solution (sol. **2.6.7**).
4. Wash the aqueous phase by adding diethyl ether (6 mL), gently shaking the conical tube and discarding the organic phase ( $\times$  3) in order to remove MPAA. Ensure to properly remove diethyl ether traces through a vigorous nitrogen or argon flow. (*see* **Note 14**).
5. In case of peptide precipitation during the diethyl ether washes, centrifuge the tube (3000 G, 15 min) to form a pellet, collect the aqueous phase then add 2 mL of 6 M Gu.HCl, and vortex or sonicate upon complete dissolution of the solid. Mix the resulted solution with the collected aqueous phase.
6. If needed add TCEP.HCl to the aqueous phase (final concentration 100 mM), raise the pH up to  $\sim$ 5 using 1M sodium hydroxyde, incubate during 30 min, then acidify the solution by addition of 2% final volume acetic acid. This step is aimed at reducing any unwanted disulfide bond arising from oxygen contamination during the reaction or work-up. Note that this step is not relevant if careful oxygen exclusion was performed as described in this chapter. Keeping the cysteine-containing peptide at low pH (2-3) usually prevents any unwanted disulfide bond formation, but the crude peptide in acidic solution should preferentially be stored at low temperature if purification cannot be performed immediately after the work-up.
7. Purify the ligation product (*see* Subheading **3.1.2**).

## 4. Notes

1. We prepare fresh solutions for the automated reductive amination reaction. We found that a one day-old sodium cyanoborohydride solution gave poorer results in terms of final peptide crypto-thioester purity.
2. The solubilization of guanidine hydrochloride is endothermic and rather slow to reach a 6 M final concentration in water. Stir, heat gently or sonicate the solution upon complete dissolution.

3. NaOH powder is highly hygroscopic, aliquot the commercial powder in several vials and store in a dessicator. Handle quickly during the weighting process.
4. Solvent deoxygenation protocol: add a magnetic stir bar into the flask containing the solution and seal it with a rubber septum, connect the flask to a three-way stopcock system using a needle. The two other ways are: an argon source (cylinder or balloon) and a 1-10 mbar medium vacuum pump. Stir the solution and turn on vacuum, bubbles should appear indicating removal of dissolved gases. If too vigorous stirring is applied, bubbling can be difficult to see. When no further bubbling occurs (avoid prolonged vacuum to prevent evaporation of the solvent), gently break the vacuum turning the stopcock to the argon position and vigorously stir for 1-2 min. Repeat these vacuum/argon cycles 3 times before disconnecting the solution by removing the needle (Fig. 7).
5. TFA is highly corrosive, volatile and toxic. Be careful while handling it. Prefer use in a fumehood.
6. To avoid imine hydrolysis and other side reactions, it is important to perform the reduction step immediately after iminium formation, and to follow our time recommendations for washings of the solid-supported iminium. We noticed once that an accidental 5-hour delay resulted in poorer final HPLC profile of the crypto-thioester peptide, including co-products not incorporating the Hnb group.
7. Hydroxylamine solution is particularly viscous. Ensure that the peptide synthesizer is able to deliver the solution to the reactor. We recently experienced problems when conducting this step at 16 °C: due to a higher viscosity, the Prelude peptide synthesizer we use failed to deliver the specified amount of this solution within the maximum time allowed by the manufacturer's software. Work is in progress to optimize an alternative reagent.
8. To quantify the amino acid coupling on the resin, it is possible to collect the Fmoc treatment solution and titrate the fluorenylmethylpiperidine adduct using an UV spectrophotometer at 301 nm ( $\epsilon = 7800 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ). Collect the Fmoc treatment and further resin wash solutions in a 200 mL volumetric flask, fill with  $\text{CH}_2\text{Cl}_2$  and measure the UV absorption at 301 nm against a  $\text{CH}_2\text{Cl}_2$  reference. 25  $\mu\text{mol}$  lead to a 0.975 optical density.
9. A gentle argon flush can conveniently be performed by connecting, through needles spanning the rubber septum, an argon input (either a balloon or direct connection to an argon cylinder) and a gas output, preferably an oil-filled bubbler which allows monitoring and regulating the argon flow, or if not available just a needle spanning the rubber septum.
10. A transfer under a controlled argon atmosphere can be performed by connecting the two plastic centrifuge tubes to a gentle argon flow. (*see Note 9*).
11. MPAA shows poor solubility in water below pH 6 and it can co-elute with peptides during HPLC analyses of purification. To avoid HPLC problems, perform a microextraction using diethyl

ether (for example, wash the 110  $\mu\text{L}$  aliquot with 100  $\mu\text{L}$  diethyl ether,  $\times 2$ ) and ensure to properly remove diethyl ether traces through vacuum or nitrogen flow before injecting the sample, in order to avoid HPLC problems (large or double peaks can be observed when traces of this solvent remains).

13. Delaying the HPLC analysis for more than one hour may afford more complex chromatograms (*N* to *S* acyl transfer can occur at acidic pH). Avoid delays or store the sample at low temperature (-80 °C storage is fine for a couple of days).

## 5. Acknowledgement

The Pierre Potier foundation for the chemistry of natural products is gratefully acknowledged for a PhD fellowship for VPT.

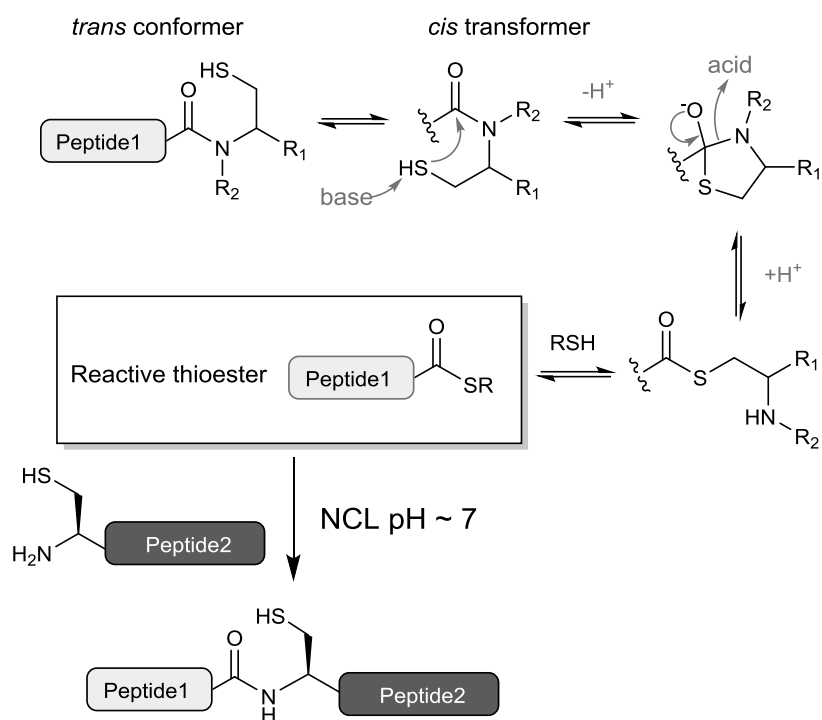
## 6. References

1. Dawson P E, Muir T W, Clark-Lewis I, Kent S B (1994) Synthesis of proteins by native chemical ligation. *Science* 266:776-779
2. Kawakami T, Sumida M, Nakamura K, Vorherr T, Aimoto S (2005) Peptide thioester preparation based on an *N-S* acyl shift reaction mediated by a thiol ligation auxiliary. *Tetrahedron Lett.* 46:8805-8807
3. Nagaike F, Onuma Y, Kanazawa C, Hojo H, Ueki A, Nakahara Y, Nakahara Y (2006) Efficient microwave-assisted tandem *N-* to *S*-acyl transfer and thioester exchange for the preparation of a glycosylated peptide thioester. *Org. Lett.* 8:4465-4468
4. Hojo H, Onuma Y, Akimoto Y, Nakahara Y, Nakahara Y (2007) *N*-Alkyl cysteine-assisted thioesterification of peptides. *Tetrahedron Lett.* 48:25-28
5. Kawakami T, Aimoto S (2007) Peptide ligation using a building block having a cysteinyl prolyl ester (CPE) autoactivating unit at the carboxy terminus. *Chem. Lett.* 36:76-77
6. Tsuda S, Shigenada A, Bando K, Otaka A (2009) *N*→*S* Acyl-transfer-mediated synthesis of peptide thioesters using anilide derivatives *Org Lett.* 11:823-826
7. Kang J, Richardson J P, Macmillan D (2009) 3-Mercaptopropionic acid-mediated synthesis of peptide and protein thioesters. *Chem Commun.* 407-409
8. Ollivier N, Dheur J, Mhidia R, Blanpain A, Melnyk O (2010) Bis(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.* 12:5238-5241
9. Hou W, Zhang X, Li F, Liu C-F (2010) Peptidyl *N,N*-bis(2-mercaptoethyl)-amides as thioester precursors for native chemical ligation. *Org. Lett.* 13:386-389
10. Zheng J-S, Chang H-N, Wang F-L, Liu L (2011) Fmoc synthesis of peptide thioesters without post-chain-assembly manipulation. *J. Am. Chem. Soc.* 133:11080-11083



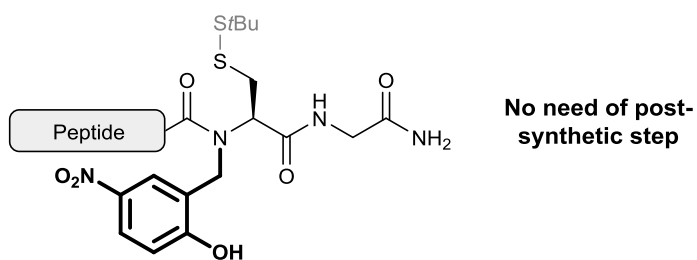
11. Burlina F, Papageorgiou G, Morris C, White P D, Offer J (2014) *In situ* thioester formation for protein ligation using  $\alpha$ -methylcysteine. *Chem. Sci.* 5:766-770
12. Asahina Y, Nabeshima K, Hojo H (2015) Peptidyl *N*-alkylcysteine as a peptide thioester surrogate in the native chemical ligation. *Tetrahedron Lett.* 56:1370-1373
13. Asahina Y, Komiya S, Ohagi A, Fujimoto R, Tamagaki H, Nakagawa K, Sato T, Akira S, Takao T, Ishii A, Nakahara Y, Hojo H (2015) Chemical synthesis of *O*-glycosylated human interleukin-2 by the reverse polarity protection strategy. *Angew. Chem. Int. Ed.* 127:8344-8348
14. Terrier V, Adihou H, Arnould M, Delmas A F, Aucagne V (2015) A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.*, DOI: 10.1039/C5SC02630J.
15. Johnson J, Quibell M, Owen D, Sheppard R C (1993) A reversible protecting group for the amide bond in peptides. Use in the synthesis of 'difficult sequences'. *Chem. Soc., Chem. Commun.* 369-372
16. L. P. Miranda L P, Meutermans W D, Smythe M L, Alewood P F (2000) An activated *O*  $\rightarrow$  *N* acyl transfer auxiliary: efficient amide-backbone substitution of hindered "difficult" peptides. *J. Org. Chem.* 65:5460-5468
17. Díaz-Mochón J. J, Bialy L, Bradley M (2004) Full orthogonality between Dde and Fmoc : the direct synthesis of PNA-peptide conjugates. *Org. Lett.* 6:1127-1129
18. Amblard M, Fehrentz J. A, Martinez J, Subra G (2005) Fundamentals of modern peptide synthesis. In: Howl J (ed) *Methods in Molecular Biology: Peptide synthesis and applications*, Volume 298. Humana Press, Totowa
19. Hackeng T M, Griffin J H, Dawson P E (1999) Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U. S. A.* 96:10068-10073

## 7. Figure captions

Fig. 1. Thioester formation mechanism through *N-S* acyl shift.

## (2-Hydroxy-5-nitrobenzyl)cysteine peptide crypto-thioesters

SiBu group prevents premature *N-S* shift and degradation  
Readily removed under NCL conditions



Automated reductive amination  
from inexpensive starting materials

Facile *N*-acylation  $\Leftrightarrow$  fully automated synthesis.  
Intramolecular catalysis of the *N-S* shift  $\Leftrightarrow$  direct use in NCL

Fig. 2. *N*-Hnb-Cys crypto-thioesters

### Annexe

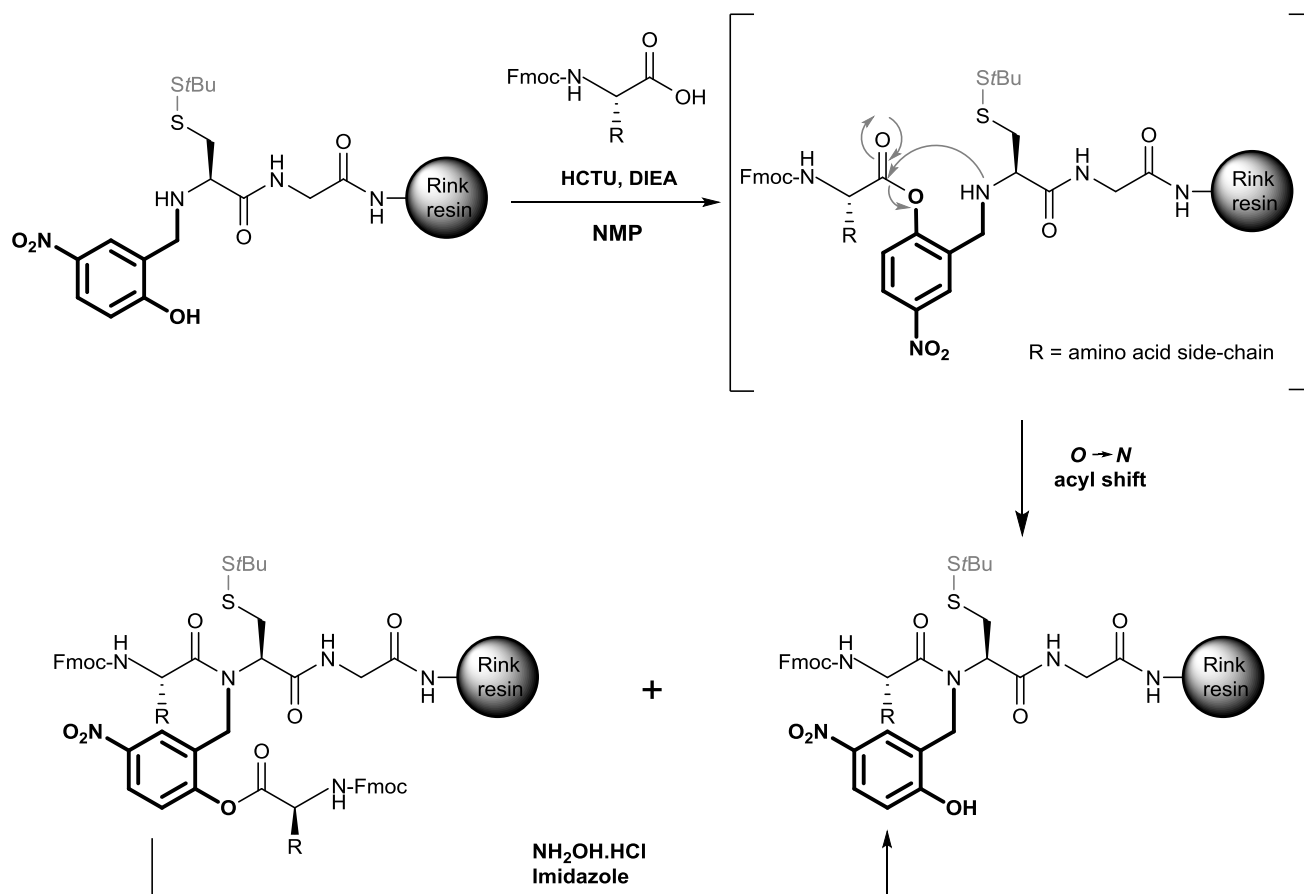


Fig. 3. Assistance of the *N*-acylation of the *N*-Hnb-Cys transesterification device by the Hnb group.

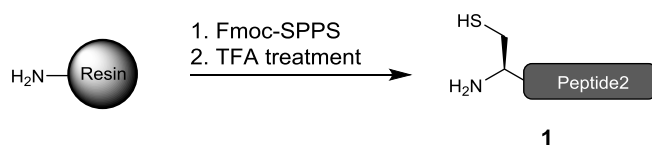


Fig. 4. Synthesis of the cysteinyl peptide 1.

Annexe

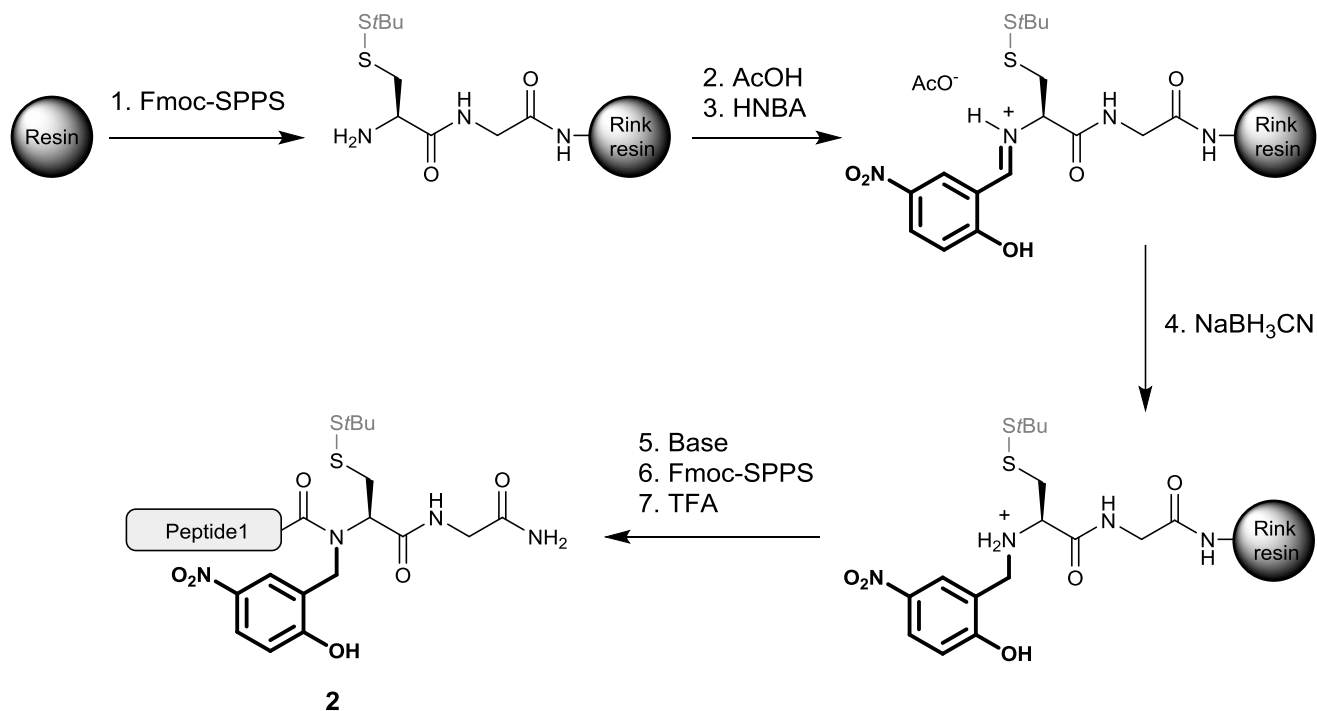


Fig. 5. Synthesis of the *N*-Hnb-Cys crypto-thioester peptide **2**.

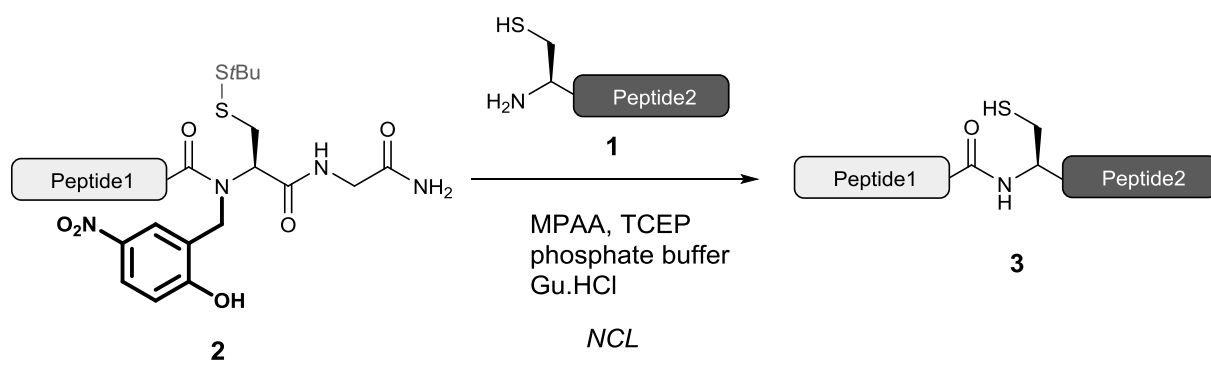


Fig. 6. NCL using *N*-Hnb-Cys crypto-thioester peptide **2**.

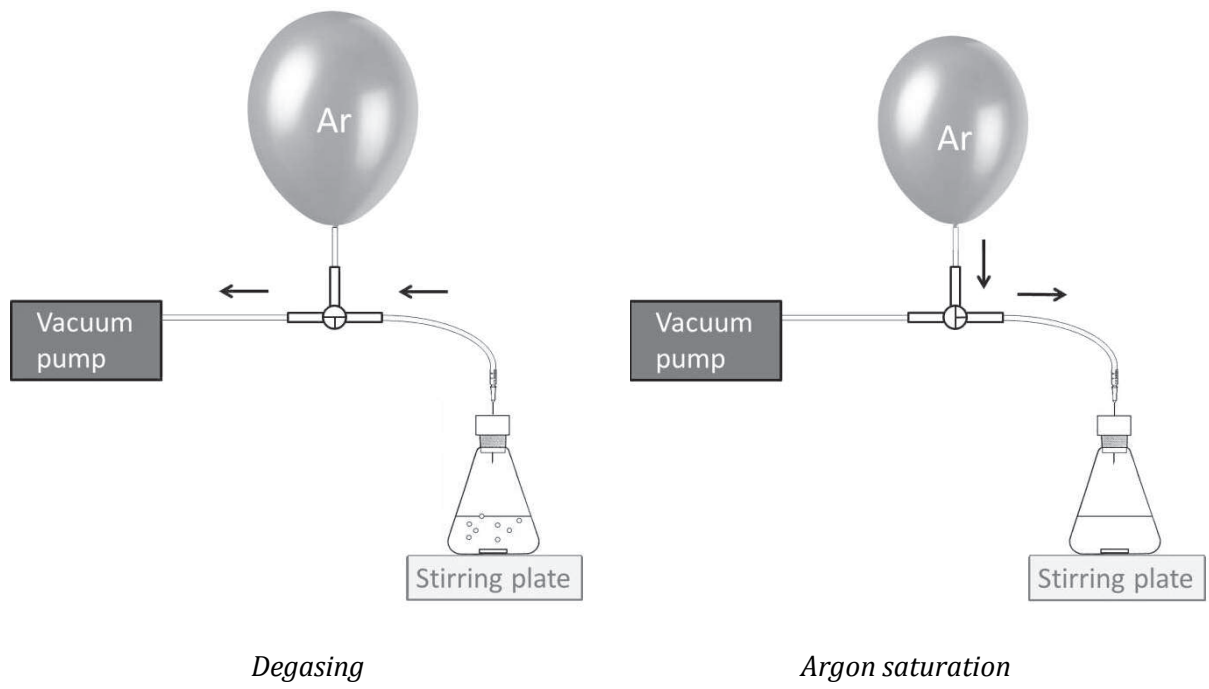


Fig. 7. Schematic representation of a simple deoxygenation system based on the use of a three-way stopcock to perform degassing/argon saturation cycles.

## Victor TERRIER

### **Synthèse biomimétique et automatisée de peptides crypto-thioester pour la ligation chimique native. Application à des peptides naturels riches en cystéine.**

Les peptides C $\alpha$ -thioester jouent un rôle central dans la synthèse de protéines par voie chimique : ils sont des partenaires clés dans la ligation chimique native (NCL), réaction qui a révolutionné ce domaine. L'accès à ces peptides par Fmoc-SPPS, méthode largement utilisée dans les laboratoires, est encore restreint à des experts. Cette limitation représente à l'heure actuelle un frein à la démocratisation de la synthèse de protéines par NCL.

Le premier volet de cette thèse décrit la conception et le développement d'une nouvelle méthode bio-inspirée de synthèse de précurseurs de peptides thioester, fondée sur un dispositif de thioestérification intramoléculaire de type *N*-(2-hydroxy-5-nitrobenzyl)-cystéine (*N*-Hnb-Cys). La synthèse des peptides portant ce dispositif a été totalement automatisée à partir de réactifs peu coûteux et ne requiert pas d'étapes post-SPPS avant la NCL. La conception biomimétique du dispositif –ce dernier opérant *via* un réarrangement par transfert d'acyle *N*→*S*–, résulte en des cinétiques de NCL rapides à pH neutre. Un large éventail de précurseurs de thioester a été synthétisé, ceci nous a permis d'explorer le potentiel et les limites de la méthodologie, tout en soulignant son efficacité même pour des « séquences difficiles » et de longs peptides.

Cette approche a été appliquée à la synthèse par ligation d'une gamme représentative de peptides naturels riches en cystéines, issus du venin d'un serpent, de mollusques, de primates ou encore de plantes. En particulier, nous avons décrit la première synthèse d'une Big-défensine (93 acides aminés), issue de l'huître et dont l'activité biologique est en cours d'évaluation. Par ailleurs, une voie synthétique originale pour accéder à des peptides comprenant une cystéine C-terminale a été proposée. Elle repose sur l'introduction de cet acide aminé par NCL, évitant les réactions secondaires inhérentes aux approches existantes. Nous avons également appliqué notre approche à la synthèse par NCL intramoléculaire du squelette peptidique cyclique de plusieurs produits naturels, avec des rendements remarquables. L'ensemble de nos résultats est extrêmement encourageant pour la généralisation de notre approche.

**Mots clés** : peptides, protéines, ligation chimique, NCL, thioesters, produits naturels.

### **Bio-inspired automated synthesis of peptides crypto-thioester for native chemical ligation. Application to naturally occurring cysteine-rich peptides.**

Peptide C $\alpha$ -thioesters play a prominent role in the chemical synthesis of proteins: they are key partners in native chemical ligation (NCL), a reaction that has revolutionized the field. Nonetheless, access to such peptides *via* the widely used Fmoc-SPPS is still limited to experts. This limitation is currently an obstacle to further popularization of NCL-based protein synthesis.

The first part of this thesis describes the design and optimization of a new bio-inspired methodology for the synthesis of peptide thioester precursors, based on an *N*-(2-hydroxy-5-nitrobenzyl)-cysteine intramolecular thioesterification device (*N*-Hnb-Cys). Synthesis of peptides bearing this device was fully automated, starting from inexpensive materials. Importantly, no post-SPPS steps are required prior to NCL. The biomimetic design of the device –that operates through an *N*→*S* acyl shift–, results in fast NCL kinetics at neutral pH. A broad range of thioester precursors has been synthesized; this allowed us to explore the scope and limitation of the methodology, while stressing its efficiency even for demanding sequences and long peptides.

This approach has been applied to the ligation-based synthesis of a representative variety of naturally occurring disulfide-rich peptides (DRP) sequences, from snake venom, molluscs, primates and plants. In particular, we have described the first synthesis of a Big-defensin (93 amino acids), discovered in oyster and whose biological activity is currently under evaluation. Furthermore, an original method for the synthesis of C-terminal cysteine-containing DRP has been proposed. It is based on the introduction of this amino acid through NCL, avoiding side reactions inherent to existing approaches. We have also applied our approach to the intramolecular NCL-based synthesis of the cyclic backbone of several DRP, with remarkable yields. Altogether, our results are extremely encouraging for the generalization of this methodology.

**Keywords**: peptides, proteins, chemical ligation, NCL, thioesters, natural products



Centre de Biophysique Moléculaire  
CNRS UPR 4301  
Rue Charles Sadron  
45071 Orléans Cedex 2

