HAL
open science

Native chemical ligation for the design of dynamic covalent peptides

Valentina Garavini

To cite this version:

Valentina Garavini. Native chemical ligation for the design of dynamic covalent peptides. Other. Université de Strasbourg, 2015. English. NNT : 2015STRAF053 . tel-01688394

HAL Id: tel-01688394

https://theses.hal.science/tel-01688394v1

Submitted on 19 Jan 2018

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.

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DE SCIENCES CHIMIQUES (ED 222)

THÈSE présentée par :

Valentina GARAVINI

soutenue le : **28 Septembre 2015**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie

Native Chemical Ligation for the Design of Dynamic Covalent Peptides

THÈSE dirigée par :

M. GIUSEPPONE Nicolas	Professeur, Université de Strasbourg - CNRS

RAPPORTEURS :

M. BARBOIU Mihail D.	Docteur, IEM-CNRS Montpellier
M. VINCENT Stéphane	Professeur, UNamur

AUTRES MEMBRES DU JURY :

M. WAGNER Alain	Docteur, Université de Strasbourg - CNRS

Table of Contents

Résumé en français **9**
ABSTRACT **9**
ACKNOWLEDGEMENTS **11**
ABBREVIATIONS AND SYMBOLS **13**
OBJECTIVES **17**
Chapter 1 – Background: Peptides and Proteins in Dynamic Combinatorial Chemistry **19**
1. Peptide and protein synthesis 19
a. Solid Phase Peptide Synthesis (SPPS) 20
i. From Boc/Bzl to Fmoc/tBu strategy 21
ii. Microwave (MW) heating in solid phase peptide synthesis 21
iii. Synthesis of large peptides and proteins 22
b. Peptide and protein synthesis by chemical ligation 23
i. Synthesis of peptides and proteins incorporating non-native covalent links 24
ii. Synthesis of native peptides and proteins 24
2. Dynamic combinatorial chemistry (DCC) 32
a. Design of a DCL 34
b. Selection mechanisms 35
c. Further considerations and applications 37
3. Peptides and proteins in dynamic combinatorial chemistry 37
a. Reversible chemistries leading to non-native peptide bonds 38
i. Amine-imine exchange reaction 38
ii. Thiol-thioester exchange reaction 39
iii. Thiol disulfide exchange reaction 40
iv. Non-covalent interactions 40
b. Reversible peptide bond: The enzyme approach 41
Chapter 2 – Peptides Incorporating N-(2-Thioethyl)-Cysteine **45**
1. Introduction: Design and objectives 45
2. Peptide synthesis 46
a. General procedure 47
b. Insertion of the 2-thioethyl chain 48
c. Glycine coupling: Acylation of N-alkyl cysteine 49

d. Simultaneous cleavage from the resin and side chain deprotection 49
e. Peptide oxidation 50
3. Exchange reaction between P1-Daa1-P2 and Daa1-P3 50
a. Exchange reaction setup 50
b. Exchange reaction mechanism 52
c. Exchange reaction analysis 53
d. Exchange reaction in detail 58
i. Complete oxidation of the reaction mixture prior to analysis 58
ii. Characterization of the thioester intermediates 60
iii. Characterization and quantification of the hydrolysis products .. 61
e. Exchange reaction with 1,4-butanedithiol replacing DTT 63
4. Exchange reaction between P1-Daa1-P3 and Daa1-P2 63
5. Control reaction between P1-Cys-P2 and Cys-P3 64
Chapter 3 – Peptides Incorporating N-Methyl-Cysteine 67
1. Introduction: N-methylation of peptides and proteins 67
a. N-methylation of peptides and proteins 67
b. N-methyl peptides as drugs 68
c. Design and objectives 70
2. Peptide synthesis 72
a. General procedure 73
b. Insertion of the methyl group 73
c. Glycine, lysine or valine coupling: Acylation of N-methyl cysteine 74
3. Exchange reaction between P1-Daa2-P2 and Daa2-P3 75
a. Exchange reaction setup 75
b. Exchange reaction mechanism 75
c. Exchange reaction analysis 76
d. Exchange reaction in detail 79
i. Characterization of the thioester intermediates 79
ii. Characterization and quantification of the hydrolysis products 81
iii. Exchange reaction in the presence of TCEP 83
4. Exchange reaction between P1'-Daa2-P2 and Daa2-P3 84
5. Exchange reaction between P1''-Daa2-P2 and P1-Daa2-P3 85
Chapter 4 – Affibody Molecules: Toward a First Biological Application of Reverse NCL 87
1. Introduction 87

a. Affinity proteins 87
b. Affibody molecules 88
c. Design and objectives 90
2. Peptide synthesis 93
a. General procedure 93
b. Fmoc group deprotection 94
c. Insertion of N-methyl-cysteine 95
i. First approach: In-situ N-methylation 95
ii. Second approach: Coupling of N-methylated cysteine 96
d. Glycine coupling: Acylation of N-methyl cysteine 97
e. Simultaneous cleavage from the resin and side-chain deprotection 97
3. Exchange reaction among H1-HER2, H2-HER2 and H3 or H1-IgG, H2-IgG and H3 98
a. Exchange reaction setup 99
b. Exchange reaction analysis 100
i. UPLC-ESI analysis 100
ii. MALDI-TOF analysis 101
iii. UPLC-ESI-TOF analysis 103
c. Influence of the filtration on the library 105
d. Influence of the exchange reaction medium on the library 106
e. StBu group deprotection and exchange reaction kinetics 106
4. Exchange reaction between H1-HER2 and H2a-HER2 107
5. Exchange reaction in the presence of the protein target: preliminary studies 108
Chapter 5 – Dynamic Glycopeptides: Toward Biological Applications of Reversible NCL 111
1. Introduction: Protein glycosylation 111
2. Dynamic glycopeptides with high affinity for Wheat Germ Agglutinine (WGA) 113
a. Design and objectives 113
b. Glycopeptide synthesis 114
3. Dynamic glycopeptides with high affinity for O-GlcNAcase (OGA) 117
a. Protein β-O-GlcNAcylation 117
b. Design and objectives 119
c. Glycopeptide synthesis 120
CONCLUSIONS AND PERSPECTIVES 123
EXPERIMENTAL SECTION 125

1. Solvents and reagents 127

2. Peptide synthesis 127

3. Chromatographic methods 127

a. Thin Layer Chromatography (TLC) 127

b. Preparative absorbance flash column chromatography 127

c. Analytical Ultra Performance Liquid Chromatography (UPLC) 128

i. UPLC-ESI 128

ii. UPLC-ESI-TOF 128

d. Preparative High Performance Liquid Chromatography (HPLC) 129

4. Structure determination 129

a. Nuclear Magnetic Resonance (NMR) 129

b. Mass Spectrometry (MS) 130

i. ESI-MS 130

ii. ESI-TOF 130

iii. MALDI-TOF 130

c. Infrared spectroscopy 130

d. Polarimetry measurements 130

5. Synthesis and characterization of organic compounds 131

a. Chapter 2 131

- Compound 1 131
- Peptides incorporating the dynamic unit N-(2-thioethyl)-cysteine (Daa1) 131
- Peptides incorporating cysteine (Cys) 135

b. Chapter 3 136

- Peptides incorporating the dynamic unit N-methyl-cysteine (Daa2) 136

c. Chapter 4 140

- Compound 2 140
- Compound 3 141
- Affibody helices 141
- Compound 4 144

d. Chapter 5 145

- Compound 5 145
- Compound 6 145
- Compound 7 146
- Compound 8 146

- Compound 9 147
- Compound 10 148
- Compound 11 148

6. Exchange reaction protocols 149

a. Chapters 2 and 3 149
- Exchange reaction between peptides incorporating the dynamic unit N-(2-thioethyl)-cysteine (Daa1) 149
- Exchange reaction between peptides incorporating cysteine (Cys) 149
- Exchange reaction between peptides incorporating the dynamic unit N-methyl-cysteine (Daa2) 150

b. Chapter 4 151
- Exchange reaction among the affibody helices H1-HER2, H2-HER2 and H3 or H1-IgG, H2-IgG and H3 or between H1-HER2 and H2a-HER2 151

ANNEXES 153

Annex 1 Structures and abbreviations of amino acids 155

Annex 2 Structures and abbreviations of specific amino acids and peptides used in Chapters 2 and 3 157

Annex 3 Structures and abbreviations of the affibody helices used in Chapter 4 158

Annex 4 Further characterization of the peptides from Chapter 2 160

Annex 5 Further characterization of the peptides from Chapter 3 172

Annex 6 Further characterization of the peptides from Chapter 4 182

Annex 7 Further characterization of the organic compounds from Chapter 5 189

Résumé en français

Une classe très importante de bio-molecules est celle des protéines. Elles sont impliqués dans pratiquement toutes les fonctions cellulaires (par exemple la catalyse, la liaison spécifique, la signalisation), mais leur utilisation n'est pas limitée aux sciences biologiques. Le stable squelette amidique et la grande diversité des possibles structures tridimensionnelles sont de grand intérêt dans d'autres domaines de la recherche, notamment dans la science des matériaux.

À différence des cellules, dans lesquelles différents systèmes sont responsables de la formation et de la rupture des liaisons peptidiques en réponse à différents stimuli, atteindre la réversibilité de cette liaison est un gros défi pour les chimistes organiques. La chimie combinatoire dynamique (DCC) est basée sur des interactions réversibles entre des unités de base. Toutes les possibles molécules formés par des interactions covalentes ou non-covalentes sont en équilibre thermodynamique et constituent une bibliothèque combinatoire dynamique (DCL). Cet équilibre est déplacé en présence d'une cible spécifique et les composants de la bibliothèque ayants une affinité majeure pour la cible sont stabilisés et amplifiés. Cette stratégie a ouvert un nouveau champ de recherche, allant de la découverte de nouveaux médicaments à la science des matériaux. Le potentiel de l'utilisation des peptides comme unités de base dans une DCL est énorme, mais aussi un gros défi en raison de l'intrinsèque stabilité et de la non-réversibilité de la liaison amide. La transamidation ou métathèse des amides constitue déjà un défi majeur. Dans le cas des peptides, la nécessité d'un contrôle serré sur la régiosélectivité au cours de la réaction d'échange et de conditions aqueuses douces (pH neutre, plage de température acceptable) complique ultérieurement les choses.

Jusqu'à aujourd'hui, DCLs à base de peptides se sont appuyés sur des réactions réversibles avec des faibles barrières d'activation. Les réactions chimiques les plus populaires étaient la réaction d'échange thiol-thioester, la réaction d'échange thiol-disulfure et la réaction d'échange amide-imine (qui comprend la liaison hydrazone). Récemment, Gellman et Ståhl ont fourni le premier exemple d'échange entre amides secondaires mais leur stratégie, impliquante un catalyseur métallique et un système non aqueux, n'est pas compatible avec les substrats peptidiques. L'approche la plus intéressante pour échanger fragments peptidiques dans des conditions douces était basée sur des biocatalyseurs tels que peptidases ou protéases. Bien que possible, déplacer l'équilibre de la réaction hydrolytique originaire vers la formation d'une liaison peptidique n'est pas évident. En outre, ces enzymes sont soit peu sélectives pour une certaine liaison peptidique (thermolysine), soit trop spécifiques pour une spécifique séquence

peptidique (sortase). Nous avons décrit une nouvelle méthodologie pour former et rompre la liaison peptidique dynamiquement, dans des conditions biocompatibles et sans compter sur l'utilisation d'enzymes.

La ligation chimique native (NCL) est une découverte assez récente dans le domaine de la chimie des peptides. Au cours des 20 dernières années, NCL a été l'outil le plus répandu pour la synthèse de longs peptides et de courtes protéines. Un peptide avec un thioester C-terminale et un peptide flanqué d'un résidu cystéine à son N-terminus réagissent dans des conditions aqueuses douces (pH neutre, température ambiante) pour former, en façon non-réversible, une liaison peptidique native. Cette réaction en deux étapes comprend l'attaque nucléophile du thiol de la cystéine au groupe carbonyle du thioester (transthioestérification), suivi par un réarrangement acyle S-N spontané, résultant en une liaison peptidique native (figure Ia). La ligation chimique native inverse a été utilisée pour synthétiser différents peptides ayants un group thioester à leur C-terminus. Par contre, en raison de la stabilité intrinsèque de la liaison peptidique, températures élevées (40-60 °C), un bas pH et un grand excès de thiol sont nécessaires afin de déplacer l'équilibre. Le résidu cystéine, de préférence dans la conformation *trans*, doit être déformée dans sa conformation *cis* pour permettre l'attaque intramoléculaire nucléophile de son groupe thiol (réarrangement acyle N-S, Figure Ib). En conclusion, les conditions nécessaires pour la NCL inverse ne sont pas compatibles avec le milieu aqueux doux de la NCL classique.

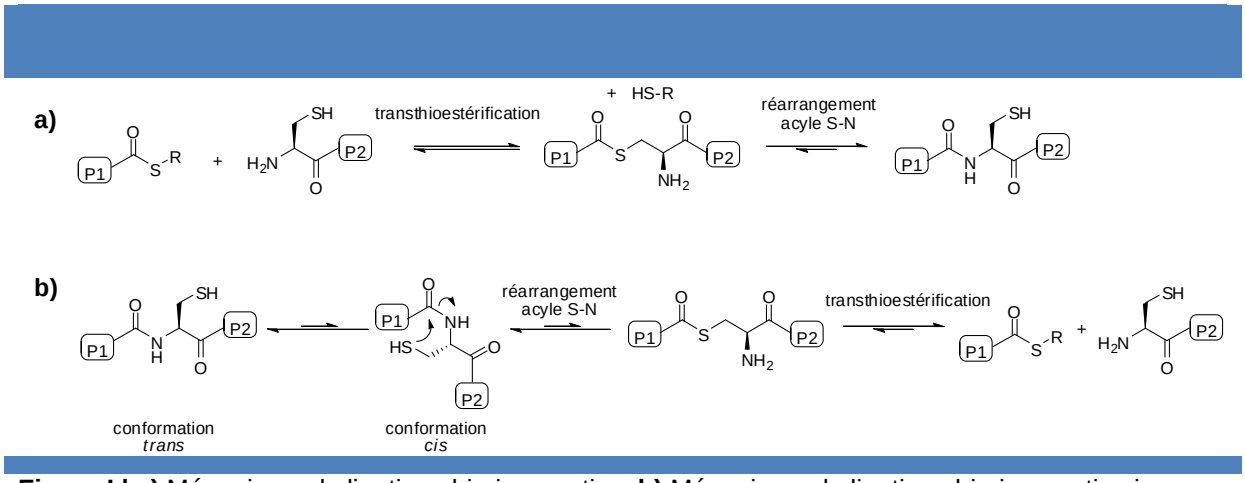


Figure I | a) Mécanisme de ligation chimique native. **b)** Mécanisme de ligation chimique native inverse.

Une étape importante vers l'échange de fragments peptidiques dans des systèmes aqueux a été la conception d'amides N-(2-sulfanyléthyle), qui peuvent agir comme thioesters latentes. Très récemment, les groupes de Melnyk et Liu ont décrit un procédé de ligation entre un peptide portant un group bis(2-sulfanyléthyl)amide (SEA) à son C-terminus et un peptide ayant une

cystéine N-terminale. Malgré la présence d'une amide tertiaire en position C-terminale, ils ont observé la réaction en conditions aqueuses neutres ou légèrement acides (pH 4 à 6). La forte réactivité de ce système est due à la présence constante d'un thiol dans la bonne position pour que le réarrangement intramoléculaire acyle N-S se produise (Figure IIa).

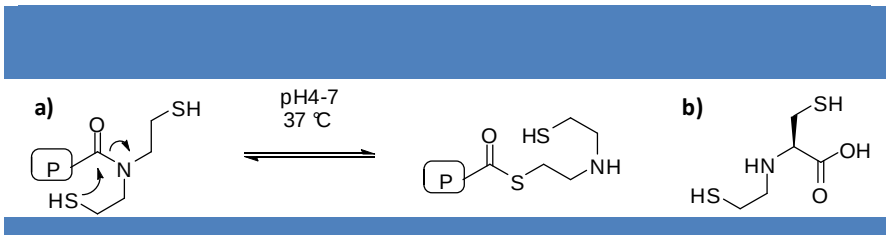


Figure II | a) Peptide ayant un groupe bis(2-sulfanyléthyl)amide (SEA) à son C-terminus. **b)** N-(2-thioethyl)-cystéine ou acide aminé dynamique 1 (**Daa1**).

Ce projet a étendu cette approche et a permis de développer une nouvelle méthodologie qui permet l'échange de fragments peptidiques rapidement et dans des conditions bio-compatibles. Nous avons conçu un nouveau acide aminé dynamique **Daa1** (Figure IIb) en ajoutant une chaîne 2-thioéthyle à l'extrémité N-terminale d'une cystéine. L'unité qui en résulte, de manière similaire au groupe SEA, porte deux groupes thiol, qui peuvent effectuer l'attaque nucléophile intramoléculaire résultant en un réarrangement acyle N-S. Nous avons préparé des courts peptides modèle par synthèse en phase solide assistée par micro-ondes et en utilisant la stratégie Fmoc/tBu. La chaîne N-(2-thioéthyle) a été ajouté sur un résidu de cystéine en phase solide, par un procédé déjà décrit dans la littérature. En raison de leur rapide oxydation intramoléculaire, les peptides ont été purifiés et stockés dans leur forme oxydée. La réaction d'échange (Figure IIIa, R = $-CH_2CH_2SH$) entre les deux peptides modèle **P1-Daa1-P2** (LYKG-**Daa1**-AKLL) et **Daa1-P3** (**Daa1**-AFKF) a été effectuée sous atmosphère d'argon, à température ambiante, dans un tampon phosphate (pH 6 à 9) et en présence de deux agents réducteurs. La tris(2-carboxyéthyl)phosphine (TCEP) a été ajoutée pour réduire la liaison disulfure intramoléculaire et le dithiothréitol (DTT) pour catalyser, grâce à la formation de l'intermédiaire P1-DTT, l'étape de transthioestérification de cette réaction. Ce dernier est le réactif de référence pour la réduction de protéines et a aussi été utilisé directement dans des cultures cellulaires. Aliquotes du mélange d'échange ont ensuite été dilués avec une solution contenante un témoin UV interne (acide 3,5-diméthoxybenzoïque) et du peroxyde d'hydrogène (pour oxyder entièrement tous les peptides et diminuer ainsi le nombre d'espèces dans le mélange). Ces solutions ont été analysés par chromatographie liquide à ultra performance couplée à unités UV et de spectroscopie de masse (UPLC-MS). En effet, nous avons pu détecter les deux produits d'échange **P1-Daa1-P3** (LYKG-**Daa1**-AFKF) et **Daa1-P2** (**Daa1**-AKLL).

Nous avons aussi étudié l'échange inverse à pH 7 entre **P1-Daa1-P3** et **Daa1-P2**, et nous avons obtenu la même distribution des peptides à l'équilibre. Cette expérience prouve, comme prévu, que le système est sous contrôle thermodynamique.

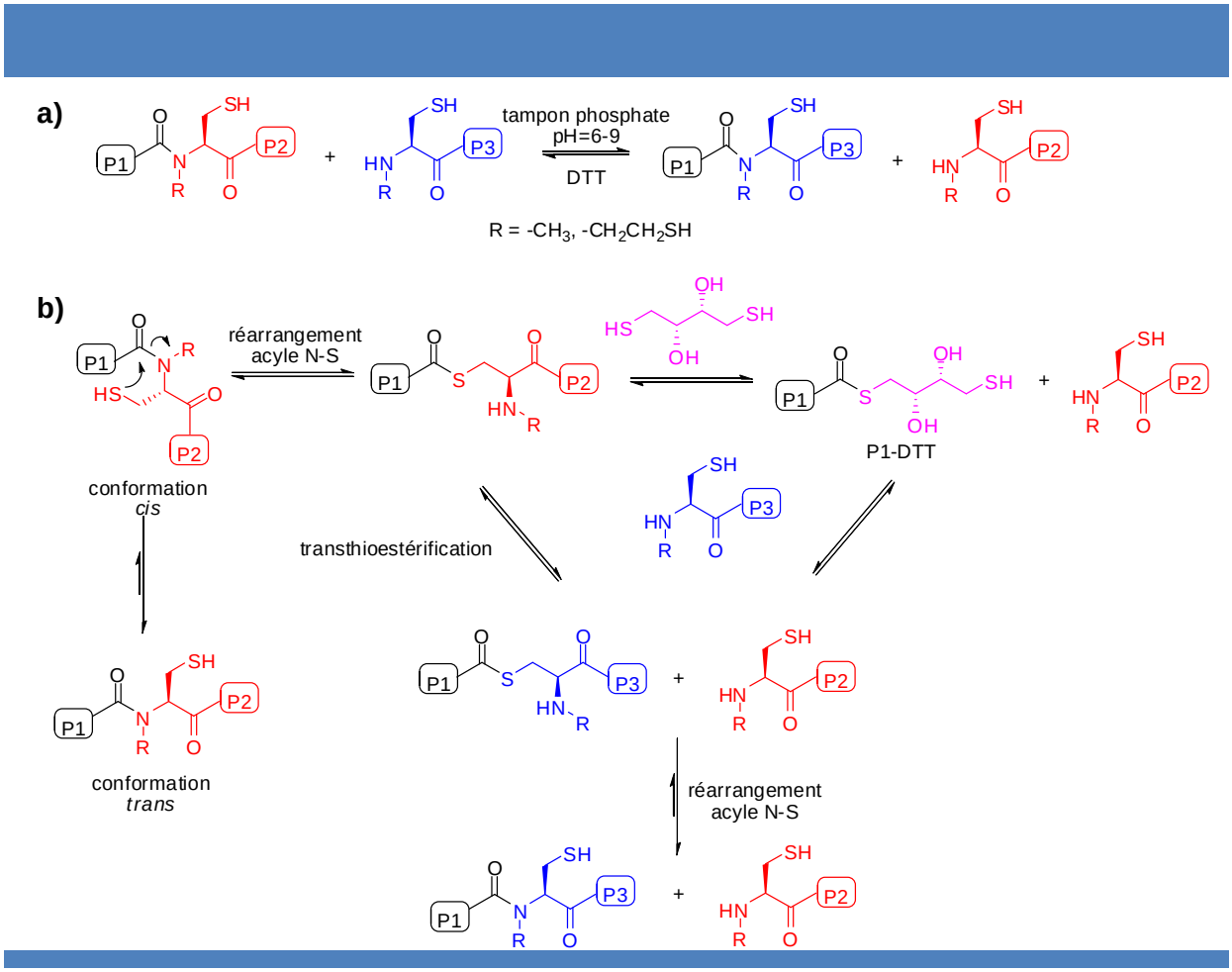


Figure III | a) Représentation schématique de la réaction d'échange. **b)** Mécanisme de réaction de la NCL réversible.

En suite, nous avons mené une expérience de contrôle à pH 7 sur les peptides «naturels» correspondants **P1-Cys-P2** (LYKG-C-AKLL) et **Cys-P3** (C-AFKF). Même après longues périodes d'équilibration (4 semaines) nous n'avons pas observé la formation de produits d'échange. Nous avons donc prouvé l'importance de la chaîne N-(2-thioéthyle) de **Daa1** dans le mécanisme de NCL dynamique.

Nous avons ensuite étudié les taux de dépendance de la réaction d'échange du pH (de 6 à 9). Comme décrit dans la littérature, le réarrangement de **P1-Daa1-P2** dans sa forme thioester est favorisée par une diminution du pH. Ceci est dû à la protonation des amines résultantes et à la conséquente inhibition du réarrangement acyle inverse S-N. En revanche, la transthioestérification est favorisée par une augmentation du pH en raison de la majeure

déprotonation des groupes thiol. Comme la vitesse initiale la plus rapide a été observé à pH 9, nous avons conclu que la transthioestérification était l'étape limitante la vitesse de la réaction globale d'échange. Ces résultats ont été confirmés par le comportement de l'intermédiaire **P1-DTT** (Figure IIIb), dont la concentration est élevée à pH 6, mais négligeable à pH 9. Malgré les taux initiaux plus rapides, nous avons observé l'hydrolyse résiduelle des intermédiaires thioester à pH élevé. Étant donné que ces produits ne sont pas impliqués dans l'équilibre et qu'ils sont nuisible pour la réaction globale, nous avons conclu que la valeur optimale de pH pour cette approche dynamique est proche de 7 (demi-temps d'équilibration de 2 heures).

Avec l'introduction de la N-(2-thioéthyl)-cystéine (**Daa1**) en peptides modèles, nous avons fourni le premier exemple de transamidation de peptides dans des conditions aqueuses douces. Même si c'est un résultant important, la biocompatibilité et la potentielle bioactivité des peptides portants cet acide aminé non naturel restent incertaines. Comme il était déjà connu dans la littérature, la présence de N-alkyle-cystéine en peptides favorise leur réarrangement en thioesters, probablement par des effets de distorsion du peptide dans sa conformation *cis*. Nous avons donc concentré notre investigation sur la N-méthyl-cystéine, un acide aminé N-alkylé présent dans nombreux produits naturels. Nous avons introduit cette acide aminé dynamique 2 (**Daa2**) dans nos peptides modèle (figure III, R = CH_3) **P1-Daa2-P2** (LYKG-**Daa2**-AKLL) et **Daa2-P3** (**Daa2**-AFKF) et étudié la réaction d'échange en présence de DTT. Aliquotes du mélange d'échange ont ensuite été diluées avec une solution contenant l'acide 3,5-diméthoxybenzoïque (témoin UV interne) et TCEP (pour maintenir tous les peptides dans leur forme réduite et diminuer le nombre d'espèces dans le mélange). Les solutions résultantes ont été analysés par UPLC-MS et les deux produits d'échange **P1-Daa2-P3** et **Daa2-P2** identifiés. En raison de la présence d'un seul groupe thiol, la réaction est plus lente que dans le cas des peptides modifiés avec l'unité **Daa1**. Néanmoins, le demi-temps d'équilibration de 10 heures à pH 7 et 37 °C est encore raisonnable et permet d'étendre l'approche combinatoire dynamique aux peptides dans des conditions biocompatibles. Enfin, nous avons étudié la réaction de métathèse entre deux peptides portants une unité **Daa2** interne: **P1-Daa2-P3** (LYKG-**Daa2**-AFKF) et **P1''-Daa2-P2** (LYKK-**Daa2**-AKLL). À pH 7 et 37 °C nous avons pu caractériser les produits d'échange **P1-Daa2-P2** et **P1''-Daa2-P3**, avec d'autres espèces impliquées dans la réaction.

Une autre étape importante dans le projet était de démontrer l'utilité de cette méthodologie pour la préparation ou la découverte de peptides bioactifs.

Comme preuve de principe, nous avons choisi de préparer des molécules affibody en

utilisant une approche de chimie combinatoire dynamique. Les affibodies sont peptides de 58 acide aminés, dérivés du domaine B de la région de la protéine staphylococcique A qui lie les immunoglobulines et sont utilisés comme alternatives synthétiques aux anticorps. Ces peptides sont très solubles dans l'eau, se replient dans une structure stable formée par trois faisceaux hélicoïdaux et sont extrêmement tolérants à la randomisation combinatoire des résidus dans les hélices 1 et 2. Fait intéressant, la ligation chimique native avait déjà été utilisée pour faciliter la préparation des affibodies et l'introduction d'un résidu cystéine dans la boucle flexible entre les hélices 2 et 3 n'avait pas affecté l'affinité des molécules de synthèse pour leurs cibles biologiques. À partir de la séquence de l'affibody (Figure IVa) avec une haute affinité pour le bio-marqueur du cancer du sein HER2 (K_D = 27 pM), nous avons conçu trois peptides correspondants aux trois hélices de l'affibody (H1, H2 et H3) et portants l'unité dynamique **Daa2** (Figure IVb). Nous avons prévu que le correspondant affibody dynamique H1-H2-H3 (Figure IVc) serait amplifié en présence de la protéine HER2 ou de son domaine extracellulaire, portant le site de liaison avec l'affibody. Pour exprimer ce domaine dans les cellules, nous avons démarré une collaboration avec le groupe du Dr. Marc Ruff à l'IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire).

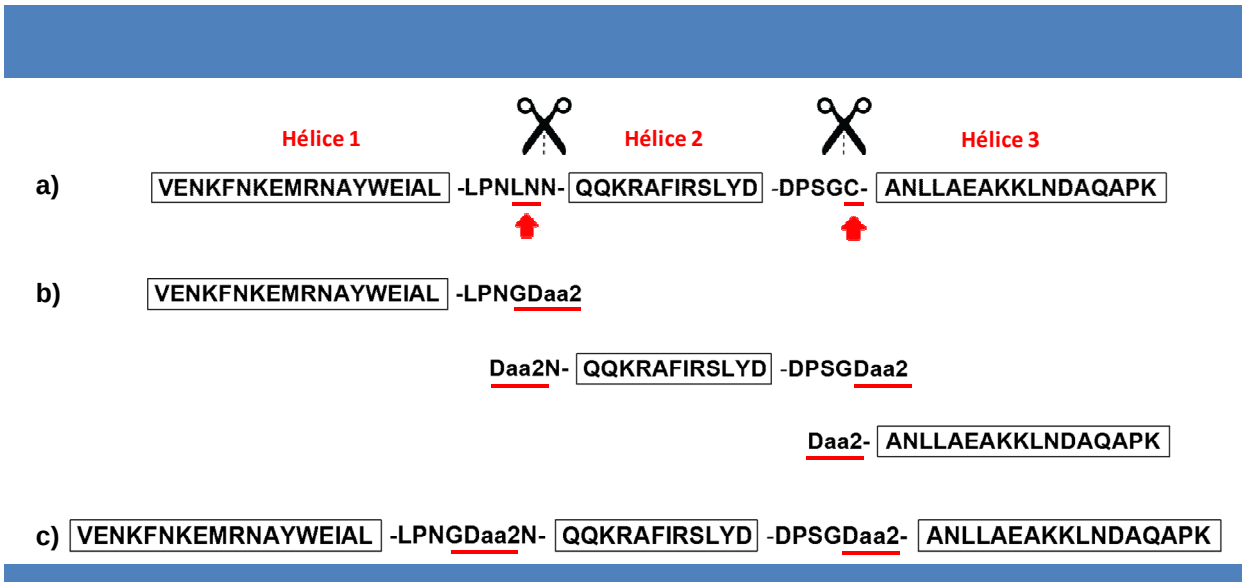


Figure IV | a) Séquence de l'affibody ayant une haute affinité pour HER2. **b)** Conception des trois hélices dynamiques H1, H2 et H3. **c)** Affibody dynamique H1-H2-H3.

Au même temps, nous avons décidé de poursuivre le projet avec l'affibody spécifique pour IgG (K_D = 22 nM). Nous avons synthétisé les trois hélices et effectué la réaction d'échange dans les mêmes conditions indiquées ci-dessus, sans IgG. Pour analyser le complexe mélange résultant, nous avons démarré une collaboration avec le Dr. Jean-Marc Strub au LSMBO

(Laboratoire de Spectrométrie de Masse BioOrganique). Grâce à la sensibilité des instruments présents dans ses laboratoires, nous avons pu démontrer la formation de l'affibody dynamique H1-H2-H3, même à faibles concentrations des hélices de départ. Le principal défi de ce projet est l'incompatibilité des cibles protéiques avec les conditions réductrices (DTT) utilisés pour effectuer la réaction d'échange. Pour surmonter cet obstacle, nous prévoyons d'utiliser une membrane de dialyse pour séparer la protéine (HER2 ou IgG) de l'agent réducteur (par exemple TCEP) lié à une résine. La membrane empêcherait la réduction de la protéine (dénaturation), tout en permettant la diffusion des produits d'échange.

En parallèle avec le projet des affibodies, nous avons démarré, en collaboration avec le groupe du Prof. Stéphane Vincent de l'Université de Namur (Belgique), deux projets visants à étendre notre approche combinatoire dynamique aux glycopeptides.

Un premier projet a été inspiré par le groupe de Renaudet, qui a récemment décrit un glycopeptide cyclique (Figure Va) avec une forte affinité (IC_{50} = 1,5 nM) pour l'agglutinine de germe de blé (WGA). WGA est une lectine spécifique pour la N-acétylglucosamine (GlcNAc) et contient 8 sites de liaison, 4 de chaque côté de la protéine. Nous avons conçu un glycopeptide cyclique similaire, où l'acide aminé dynamique 2 remplace les résidus d'alanine (Figure Vb). Une réaction d'échange entre deux unités du glycopeptide linéaire montré en Figure Vc, en présence de WGA, conduirait à la formation et à l'amplification du correspondant glycopeptide cyclique.

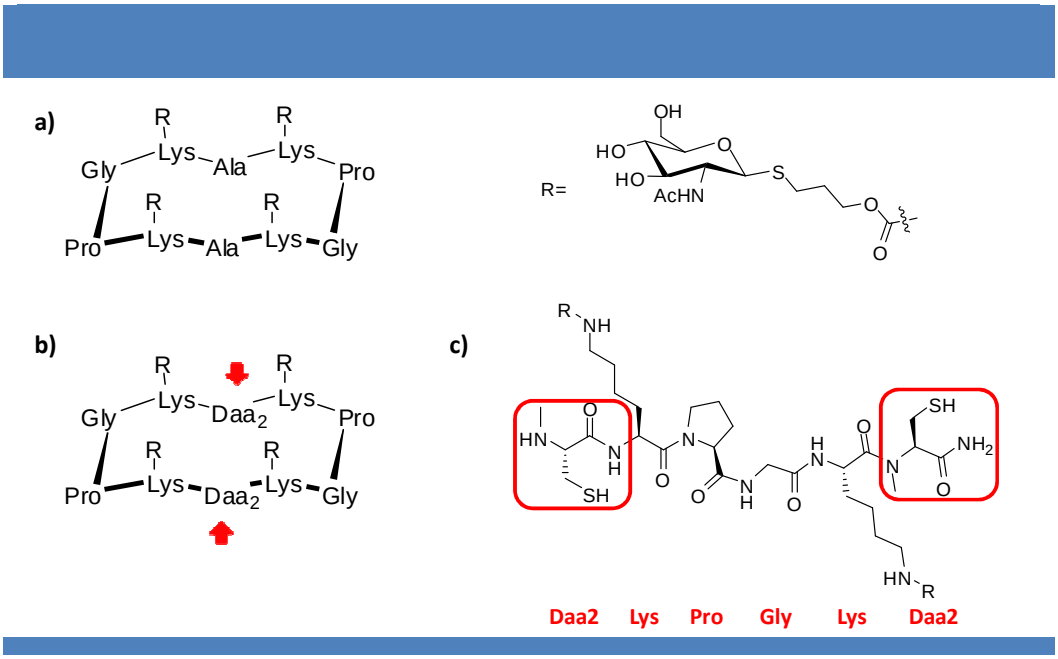


Figure V | a) Glycopeptide cyclique avec une forte affinité pour WGA. **b)** Conception du glycopeptide cyclique dynamique. **c)** Glycopeptide linéaire dynamique.

Le but du deuxième projet était la formation, suite à une réaction de NCL réversible, d'un glycopeptide avec une haute affinité pour l'enzyme O-N-acétylglucosaminidase (OGA). Nous avons conçu un glycopeptide sur la base du fragment de la protéine p53 qui lie OGA: nous avons remplacé le résidu Thr adjacent à Ser(O-GlcNAc) avec une unité de N-méthyl-cystéine (**Daa2**). En plus, nous avons remplacé Ser(O-GlcNAc) avec Cys(S-GlcNAc), plus stable à l'action hydrolytique de OGA (Figure VIa). La réaction d'échange montrée en Figure VIb entre deux fragments de ce glycopeptide devrait conduire à sa formation. L'ajout de OGA et de fragments randomisés de glycopeptide à la bibliothèque devrait conduire à l'amplification du glycopeptide NH_2-QLWVDC(S-GlcNAc)**Daa2**PPPG-$CONH_2$.

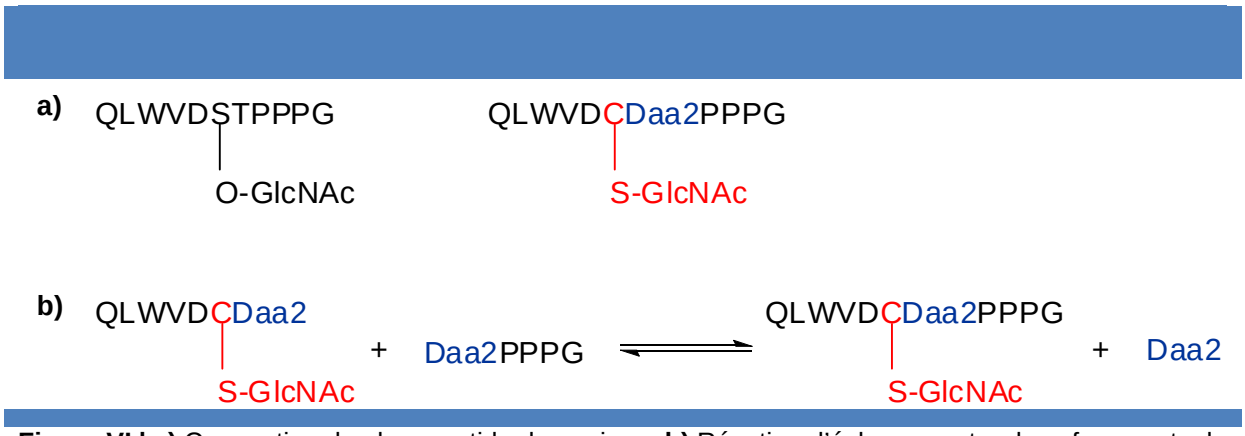


Figure VI | a) Conception du glycopeptide dynamique. **b)** Réaction d'échange entre deux fragments de ce glycopeptide.

En raison des complications apparues dans le projet avec les molécules affibody, nous avons préféré arrêter momentanément les projets des glycopeptides et nous concentrer sur l'étude de la réaction d'échange dans un système plus simple. Même si j'ai participé activement à la conception de ces expériences, ce travail a été principalement effectué par Cristian Reţe et n'est donc pas décrit en détail dans ce manuscrit.

ABSTRACT

Proteins are important biomolecules, involved in virtually all cell functions (e.g. catalysis, specific binding, signaling). Whereas different cellular systems are responsible for the continuous formation and disruption of peptide bonds in response to different stimuli, achieving the reversibility of such a bond is not a trivial challenge for organic chemists. Dynamic combinatorial chemistry (DCC) is based on reversible interactions among simple building blocks. All the possible molecules formed by these covalent or non-covalent interactions are in thermodynamic equilibrium and constitute a dynamic combinatorial library (DCL). Such equilibrium shifts in the presence of a given target and the components that best bind it are stabilized and amplified.

The potential of using peptides as building blocks in a DCL is enormous, yet challenging due to the intrinsic stability and non-reversibility of the amide bond. The transamidation or amide metathesis of simple amides already constitute a major challenge. In the case of peptides, the need for a tight control of the regioselectivity during the exchange reaction and for mild aqueous conditions (neutral pH, acceptable temperature range) further complicates matters. Hence, peptide-based DCLs have generally relied on reversible reactions with low activation barriers (i.e. not involving the native peptide bond) or, in a more sophisticated approach, on the use of peptidases. However, both these strategies have several limitations.

We have developed a new methodology to dynamically form and disrupt peptide bonds in biocompatible conditions without relying on the use of enzymes. It rests on native chemical ligation (NCL), a two-step reaction between a C-terminal thioester on one peptide and a N-terminal cysteine on a second peptide, that allows the formation of the native peptide bond in mild aqueous conditions. We started our investigation by designing a novel “dynamic amino acid” (Daa1), a N-(2-thioethyl)-cysteine moiety. Exchange reactions between model peptides incorporating this residue showed the fast formation of the exchange products in mild aqueous conditions and in the presence of reducing agents. Later, we improved this first design by introducing in the model peptides a natural occurring dynamic unit (Daa2), a N-methyl-cysteine moiety. Although with slower equilibration rates, we could identify the corresponding exchange peptides. Hence, we have developed a strategy to scramble natural occurring peptide fragments in mild conditions and without relying on the use of enzymes.

We have then started the investigation of potential biological applications of this new reversible NCL strategy. We have designed two different libraries based on peptide sequences

derived from affibody molecules specific for either HER2 or IgG and we plan to investigate the exchange reaction in the presence of the biological target.

In parallel, in collaboration with Prof. Stéphane Vincent (UNamur), we have designed two separate libraries based on peptide and glycopeptide fragments, whose components should rearrange in order to form the best binder to the given target (WGA or OGA).

ACKNOWLEDGEMENTS

First, I would like to acknowledge Prof. Nicolas Giuseppone, who gave me the opportunity to develop my scientific skills as well as to be part of two very international groups of young scientists. For the past three years I have been a member of the SAMS group and an early stage researcher in the Marie Curie Initial Training Network DYNANO (grant agreement 2011-289033). Second, I'd like to thank Dr. Yves Ruff for his guidance and constant support throughout this PhD (until the very end).

I would also like to acknowledge all the people from the DYNANO, who were always ready to discuss ideas and talk about possible collaborations. In particular, I'd like to thank Prof. Stéphane Vincent for the fruitful scientific discussions we had every six months and for the nice period I have spent in his laboratories in Namur (and thanks to the whole CBO group too, in particular to Marta and Olpi). Un remerciement spécial au Dr. Jean-Marc Strubb pour son aide précieux et sa constante disponibilité. I would also like to acknowledge the "peptide sub-group" (Cristian Reţe, Dr. Daniel Funeriu and Dr. Manickasundaram Samiappan) and all the other people of the SAMS group that have helped me in many different ways in these years: Emilie, Gad, Mounir, Mélodie, Julie, Odile, Yunjie, Jan, Qing, Thomas, Ting, Junjun, Tom, Quan, Eric, Antoine, Jean-Rémi (and I'm sure I've forgotten someone). I'd also like to thank Artem and Adrian for the good laughs we shared in the lab, and Justin and Simon for their support in these last few months. Special thanks to Sanne, for being by my side in all the stressful moments of my PhD (with an éclair à la vanille). And obviously thanks to Joe, who has been in charge of my mental health in this last year or so, and did a pretty good job.

I should also acknowledge other people from the ICS, in particular Iuliia, Akkiz, Rémi and Olga, and the little Italian community of Strasbourg. Grazie a Stefano per le continue discussioni su…tutto, e un grazie speciale a Cinzia, è incredibile che finiamo lo stesso giorno! Merci à mes amis de l'escalade, qui m'ont soutenu pendant ces années et qui m'ont vu disparaître ces derniers mois. Un gros merci à Mélanie, Mia, Alex et Matthieu. Et merci à mes colocs, surtout à Kévin, Lea, Antonin, Laurane et Alexis.

Grazie a Cristina&Rafel, i miei genitori all'estero, per avermi trascinato a Strasburgo e per essere sempre al mio fianco, nonostante la distanza. E grazie agli amici di sempre, Doda&Rusti in primis ma anche alle donnine ravennati, a Giova, Dario, Miki, alla Giuly, l'Ele, l'altra Ele e la Marty e a tutti quelli che mi fanno sentire a casa ogni volta che torno.

E infine grazie ai miei genitori per tutto, per il supporto incondizionato, per le "piccole"

deviazioni sulla tabella di marcia nel momento del bisogno e per essere un punto di riferimento, sempre e comunque. Grazie anche alla Milena, alle nonne e al resto della mia famiglia.

ABBREVIATIONS AND SYMBOLS

Å	Ångstrom
aa	amino acid
Ab	antibody
AIBN	azobisisobutyronitrile
Alloc	allyloxycarbonyl
ATR	attenuated total reflectance
Boc	*tert*-butyloxycarbonyl
Bpy	bipyridine
Bzl	benzyl
°C	Celsius degree
CBD	chitin binding domain
CHCA	α-cyano-4-hydroxycinnamic acid
CHO	chinese hamster ovary
COSY	correlation spectroscopy
*Cp*OGA	*Clostridium perfrigens OGA*
CSA	camphorsulfonic acid
Daa1	N-(2-thioethyl)-cysteine
Daa2	N-methyl-cysteine
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	Dynamic Combinatorial Chemistry
DCL	Dynamic Combinatorial Library
DCM	dichloromethane
DHB	2,5-dihydroxybenzoic acid
DIAD	diisopropyl azodicarboxylate
DIC	N,N'-diisopropylcarbodiimide
DIEA	N,N-diisopropylethylamine
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ecd	extracellular domain
EDTA	ethylenediaminetetraacetic acid
eq.	equivalents
ESI	electrospray ionization
Fmoc	9-fluorenylmethoxycarbonyl
g	gram
GlcNAc	N-acetyl-glucosamine
h	hour
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo [4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HF	hydrogen fluoride
HIV	human immunodeficiency virus
HMQC	multiple-quantum correlation
HOBt	hydroxybenzotriazole
*h*OGA	human OGA
HPLC	high performance liquid chromatography
Hz	hertz
ICP-AES	Inductively coupled plasma atomic emission spectroscopy
Ig	immunoglobulin
IGBMC	institut de génétique et de biologie moléculaire et cellulaire
IgG	immunoglobulin G
IR	infrared
J	coupling constant
K	equilibrium constant
kcal	kilocalory
K_D	equilibrium dissociation constant
kDa	kilodalton
K_M	Michaelis constant
LSMBO	laboratoire de spectrométrie de masse bioorganique
µL	microliter
µM	micromolar
MBHA	4-methylbenzhydrylamine
M	molar
MALDI	matrix-assisted laser desorption/ionization
Me	methyl
mg	milligram
MHz	megahertz
min	minutes
mL	milliliter
mM	millimolar
mmol	millimole
mol	mole
MPAA	4-mercaptophenylacetic acid
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MTBD	7-methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene
Mtt	4-methyltrityl
MW	microwave
m/z	mass-to-charge ratio
NBS	4-nitrobenzenesulfonyl
NCL	native chemical ligation

nM	nanomolar
NMP	N-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
OGA	O-GlcNAcase
O-GlcNAcase	O-GlcNAc hydrolase
O-GlcNAc	O-linked N-acetylglucosamine
OGT	O-GlcNAc transferase
Oxyma	ethylcyano(hydroxyimino)acetate
Pac	phenylacyl
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PEG	polyethylene glycol
PG	protecting group
pM	picomolar
PTFE	polytetrafluoroethylene
Py	pyridine
R^2	coefficient of determination
RAFT	regioselectively addressable functionalized template
RNase	ribonuclease
SEA	(2-sulfanylethyl)amino
SPPS	solid phase peptide synthesis
t	time
TBAF	tetra-N-butylammonium fluoride
TBAHS	tetrabutylammonium hydrogen sulfate
tBu	*tert*-butyl
TCEP	tris(2-carboxyethyl)phosphine
TES	triethylsilane
TFA	trifluoroacetic acid
TGS	target-guided synthesis
THF	tetrahydrofurane
TIC	total ion current
TIS	triisopropylsilane
TLC	thin layer chromatography
TOF	time of flight
t_r	retention time
Trt	trityl
UPLC	ultra performance liquid chromatography
UV	ultra violet
V_0	initial rate
v/v	volume over volume
w/v	weight over volume
W	watt
WGA	wheat germ agglutinine
Z	affibody molecule

Z_{HER2}	affibody molecule specific for HER2
Z_{IgG}	affibody molecule specific for IgG
$Z_{Insulin}$	affibody molecule specific for insulin

OBJECTIVES

Upon joining this research group in June 2012 I was assigned to work together with Dr. Yves Ruff toward the development of a reversible chemistry to scramble peptide fragments at the amide bond level. This new reaction could be used in dynamic combinatorial chemistry in a variety of applications and in different research fields, ranging from biology to materials science. For instance, one can envisage to create a dynamic combinatorial library of peptide fragments based on this reversible reaction and amplify a high-affinity ligand to a target template. Conversely, this reversible chemistry could be used to select the receptor for a guest molecule.

The main objective of my PhD was to develop an exchange reaction that could occur in mild aqueous conditions and be selective for a specific peptide bond. The intrinsic stability of the amide bond was the first big hurdle to overcome. Although one group had shown that transamidation and amide metathesis were possible, the reaction conditions reported were not compatible with peptide substrates. The necessity to operate in non-aqueous systems and to use harsh conditions (metal catalysts), together with the lack of selectivity for a specific junction, were serious limitations to this approach. Other groups circumvented the low reactivity of the amide bond by generating peptide-based libraries using different reversible chemistries. The thiol-disulfide, the amine-imine and the thiol-thioester exchange reactions were used, leading to non-native peptides.

The only strategy described in the literature to continuously form and disrupt peptide bonds was based on the use of peptidases. However, this approach had several limitations in terms of compatibility with protein targets and selectivity for a specific amide bond.

In order to develop a reversible chemistry at the peptide bond level we investigated the potential of native chemical ligation (NCL), the most popular approach for the synthesis of long peptides. This reaction was a perfect starting point for our project: it occurs in aqueous buffers, in mild conditions and specifically at the N-terminus of a cysteine residue. We thus focused our efforts in the development of a strategy to allow the reverse NCL to occur in the same mild conditions.

A second important objective of my PhD was to provide a biological application of this reversible chemistry. We wanted to generate peptide (or glycopeptide) dynamic combinatorial libraries based on this reversible reaction and amplify the best binder for a specific biological target (i.e. Her2, IgG, WGA and OGA).

Chapter 1 - Background: Peptides and Proteins in Dynamic Combinatorial Chemistry (DCC)

1. Peptide and protein synthesis

Proteins are an important class of biological macromolecules, involved in structural roles and virtually all cell functions (e.g. catalysis, specific binding, signaling).

The information necessary for the synthesis of a protein is encoded in DNA. Its transcription into mRNA, followed by mRNA translation into polypeptide chains, is the base of protein biosynthesis. However, protein diversity is enhanced by mRNA splicing prior to translation and by post-translational modifications. It has been estimated that a cell can express more than 100000 proteins[1], whose unique folded structures are strictly correlated to their functions.

Proteins have been routinely synthesized by means of molecular biology techniques. Importantly, the possibility of selectively modify one or more amino acid residues has provided useful insights into the mechanism of different enzymes and other proteins. However, this approach is limited to the 20 genetically encoded amino acids and insertion of post-translational modifications at specific sites is not possible with standard protocols. Although few methods[2] that overcome this limitation have been reported, they are generally too sophisticated to be routinely used.

In contrast to molecular biology techniques, synthetic chemistry enables the incorporation of virtually any functionality onto proteins. From the first synthesis of glycol glycine by Fischer[3] in 1901, peptide and protein chemistry has become a field of increasing importance.

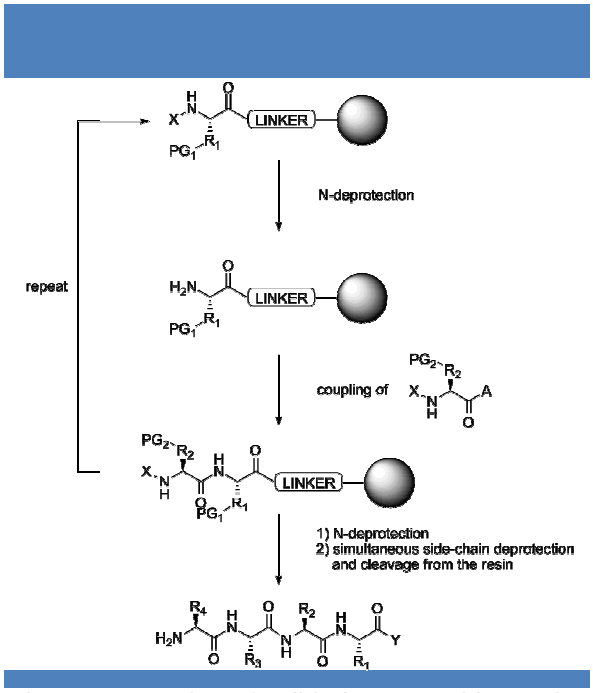
1. Unwin, R. D.; Gaskell, S. J.; Evans, C. A.; Whetton, A. D. The Potential for Proteomic Definition of Stem Cell Populations. *Exp. Hematol.* **2003**, *31*, 1147–1159.
2. a) Mendel, D.; Cornish, V. W.; Schultz, P. G. Site-Directed Mutagenesis with an Expanded Genetic Code. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 435–462. b) Hendrickson, T. L.; de Crécy-Lagard, V.; Schimmel, P. Incorporation of Nonnatural Amino Acids into Proteins. *Annu. Rev. Biochem.* **2004**, *73*, 147–176. c) Gubbens, J.; Kim, S. J.; Yang, Z.; Johnson, A. E.; Skach, W. R. In Vitro Incorporation of Nonnatural Amino Acids into Protein Using tRNA(Cys)-Derived Opal, Ochre, and Amber Suppressor tRNAs. *RNA* **2010**, *16*, 1660–1672.
3. Fischer, E.; Fourneau, E. Über Einige Derivate Des Glykocolls. *Berichte der Dtsch. Chem. Gesellschaft* **1901**, *34*, 2868–2877.

a. Solid Phase Peptide Synthesis (SPPS)

Initial attempts of performing peptide synthesis in solution were highly limited by the necessity of a tight control on the stereochemistry of the coupling reaction and by the poor solubility of the resulting peptides.

The introduction of solid phase peptide synthesis by Merrifield[4] in 1963 revolutionized the field enabling peptide synthesis to become a routine procedure, whereas before it was a daunting challenge. It has been estimated[5] that the synthesis of a target peptide is ~50 times more arduous in solution than in solid phase.

In this technique, the target peptide is grown on an insoluble polymeric support in the C→N direction, opposite to the naturally occurring synthesis in the ribosome. The C-terminal amino acid is covalently bound to the resin, its N-terminus deprotected and the conveniently protected following amino acid coupled. These steps are then repeated until the full peptide chain is obtained. Finally, after simultaneous side-chain deprotection and cleavage from the solid support, the target peptide is obtained (Scheme 1). Due to the number of reactions involved in a single synthesis, each step needs to occur with a near-quantitative yield. Importantly, since purification after each step simply involves filtration and washing, excess of reagents can be used for coupling.



Scheme 1 | Schematic representation of solid phase peptide synthesis. X = Boc or Fmoc; R_1, R_2, R_3, R_4 = amino acid side chains; PG_1 and PG_2 = side-chain protecting groups; A = C-terminal activating group; Y = OH or NH_2.

4. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
5. Kent, S. B. H. Total Chemical Synthesis of Proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.

i. From Boc/Bzl to Fmoc/tBu strategy

Protecting groups on the amino acid lateral chains and on their N-terminus play a fundamental role in peptide synthesis. Merrifield[4] initially introduced the so-called Boc/Bzl strategy, where the amino acid N-terminus is protected with a *tert*-butyloxycarbonyl (Boc) group and the side chains with a benzyl (Bzl) group. The Boc/Bzl strategy has been used almost exclusively for the first 15 years of solid phase peptide synthesis and it has permitted remarkable synthetic achievements[6]. On the other hand, the need for hydrogen fluoride (HF) to perform the final cleavage of the peptide from the resin is a serious drawback, as special equipment is required for the handling of this corrosive and toxic compound.

In the relatively recent Fmoc/*tert*-butyl[7] (tBu) strategy, the amino acid side chains are generally tBu-protected and the N-terminus 9-fluorenyl-methyloxycarbonyl (Fmoc) protected. The advantages of this approach on the original Boc/Bzl one are two fold. First, the N-terminal group is orthogonal to the side-chain groups: the Fmoc group is deprotected by base-induced β-elimination (e.g. piperidine in DMF), the removal of the *tert*-butyl occurs in acidic conditions, (e.g. TFA). Second, the final cleavage is simply achieved in the reaction vessel with TFA and scavengers, obviating the need for special equipment and dangerous handling.

ii. Microwave (MW) heating in solid phase peptide synthesis

In few examples conductive heating had been applied to SPPS of long peptides, especially during the coupling step[8]. The evolution in the use of microwave heating in organic synthesis in the 1990s led to its application to SPPS. While Wang and coworkers[9] used a slightly modified domestic microwave oven, Erdélyi and Gogoll[10] showed that special microwave reactors could indeed be used to dramatically improve speed and purity in SPPS. Unlike conventional heating, microwave energy directly activates any molecule (solvents and some reagents) with a dipole moment, allowing for rapid heating at the molecular level. In the synthesis of long and hydrophobic peptide sequences, microwaves have been shown to disrupt aggregation through dipole rotation of the polar peptide (Figure 1), thus improving the final yield of the synthesis. Importantly, DMF and NMP, the two most common solvents in SPPS, are polar molecules,

6. Ohno, M.; Eastlake, A.; Ontjes, D. A.; Anfinsen, C. B. Synthesis of the Fully Protected, Carboxyl-Terminal Tetradecapeptide Sequence of Staphylococcal Nuclease. *J. Am. Chem. Soc.* **1969**, *91*, 6842–6847.
7. Chang, C. D.; Meienhofer, J. Solid-Phase Peptide Synthesis Using Mild Base Cleavage of N Alpha-Fluorenylmethyloxycarbonylamino Acids, Exemplified by a Synthesis of Dihydrosomatostatin. *Int. J. Pept. Protein Res.* **1978**, *11*, 246–249.
8. a) Varanda, L. M.; Miranda, M. T. M. Solid-Phase Peptide Synthesis at Elevated Temperatures: A Search for an Optimized Synthesis Condition of Unsulfated Cholecystokinin-12. *J. Pept. Res.* **2009**, *50*, 102–108. b) Kaplan, B. E.; Hefta, L. J.; II, R. C. B.; Swiderek, K. M.; Shively, J. E. Solid-Phase Synthesis and Characterization of Carcinoembryonic Antigen (CEA) Domains. *J. Pept. Res.* **2009**, *52*, 249–260.
9. Yu, H. M.; Chen, S. T.; Wang, K. T. Enhanced Coupling Efficiency in Solid-Phase Peptide Synthesis by Microwave Irradiation. *J. Org. Chem.* **1992**, *57*, 4781–4784.
10. Erdélyi, M.; Gogoll, A. Rapid Microwave-Assisted Solid Phase Peptide Synthesis. *Synthesis* **2002**, *2002*, 1592–1596.

therefore they are easily heated by microwaves[11].

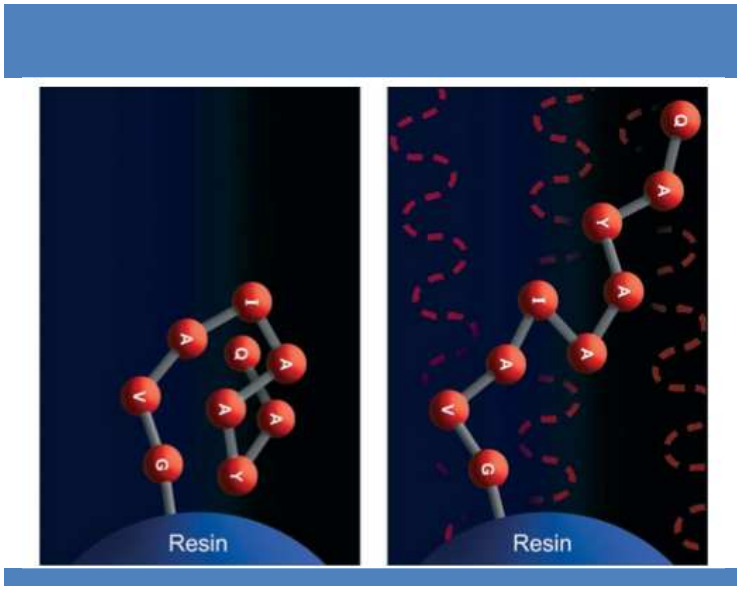


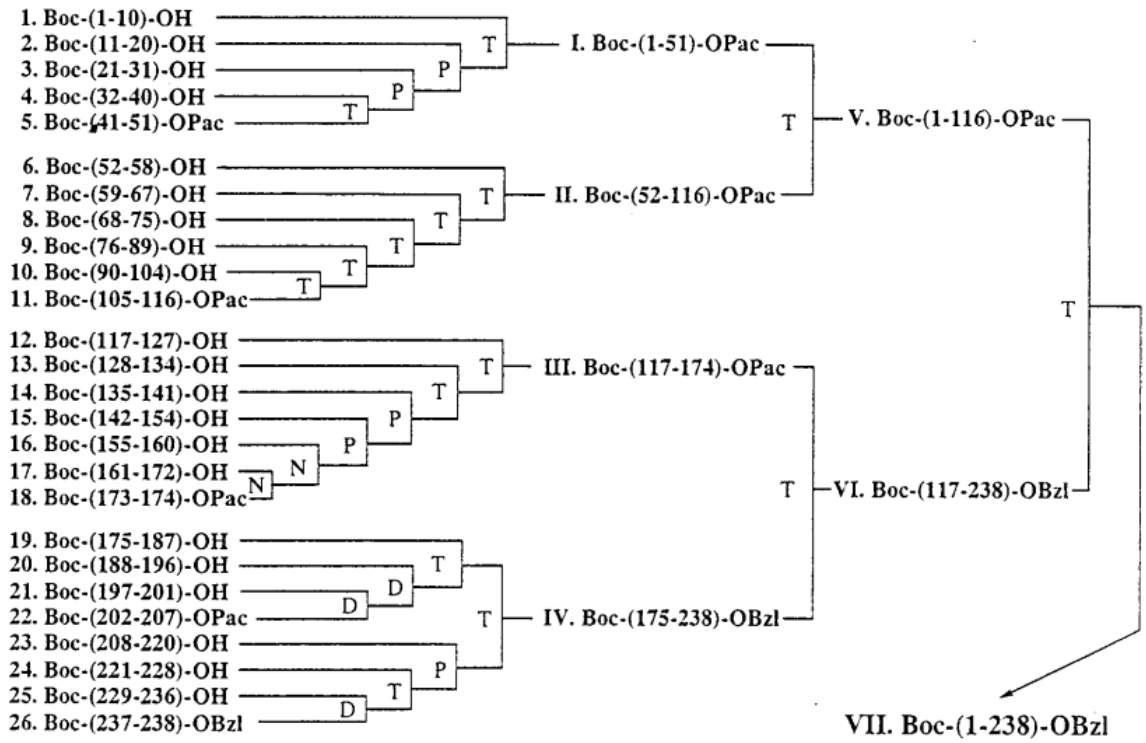
Figure 1 | Peptide chain attached to a solid support before and during microwave irradiation, showing disruption of aggregation through dipole rotation of the peptide (from ref. 12).

iii. Synthesis of large peptides and proteins

Countless peptides have been prepared by stepwise solid phase synthesis. From its introduction, new orthogonal protecting groups and new coupling reagents have been developed and purification and characterization techniques improved. However, and despite these achievements, peptides containing more than 50 amino acids are generally hard to obtain in good yield and purity. Considering that a typical natural protein molecule consists of ~300 amino acids[5], this is a serious limitation of solid phase peptide synthesis.

In order to overcome this problem, protected or partially protected peptide segments have been assembled, either in solution or in solid phase. An impressive example is the convergent synthesis of the 238-residue precursor molecule of the *Aequorea* green fluorescent protein by Nishiuchi *et al.*[13] by classical condensation of 26 different peptide segments (Scheme 2).

11. Pedersen, S. L.; Tofteng, A. P.; Malik, L.; Jensen, K. J. Microwave Heating in Solid-Phase Peptide Synthesis. *Chem. Soc. Rev.* **2012**, *41*, 1826–1844.
12. Palasek, S. A.; Cox, Z. J.; Collins, J. M. Limiting Racemization and Aspartimide Formation in Microwave-Enhanced Fmoc Solid Phase Peptide Synthesis. *J. Pept. Sci.* **2007**, *13*, 143–148.
13. Nishiuchi, Y.; Inui, T.; Nishio, H.; Bódi, 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. U. S. A.* **1998**, *95*, 13549–13554.



Scheme 2 | Convergent synthesis of the precursor molecule of the *Aequorea* green fluorescent protein. Bzl = benzyl; Pac = phenacyl; the following letters refer to the solvent used for the coupling reaction: D = DMF; N = NMP; P = phenol/$CHCl_3$ 1:3; T = TFE/ $CHCl_3$ 1:3 (adapted from ref. 1).

Another noteworthy strategy is the so-called "thioester method"[14], consisting in the reaction between a C-terminal alkyl thioester, activated with silver ions, and the N-terminus of a second peptide.

Although important synthetic goals have been achieved with these strategies, the need for protection of the peptide fragments is a serious drawback, especially in the case of long sequences. The reactivity between an activated carboxyl group and an amino group is often not high enough to overcome the relative low solubility of the peptide fragments.

b. Peptide and protein synthesis by chemical ligation

In order to overcome the solubility-related issues of the convergent strategies discussed above, the idea of selectively couple two or more unprotected peptide fragments has emerged. However, due to the variety of functional groups on the peptide side chains, the development of

14. 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_2. *Bull. Chem. Soc. Jpn. 64*, 111–117.

chemical ligation strategies to selectively form the target amide bond was indeed a significant challenge.

i. Synthesis of peptides and proteins incorporating non-native covalent links

The formation of a new amide bond in peptide synthesis is routinely achieved through reaction of an activated carboxyl group with a free amino group. However, this strategy cannot be applied in the case of non-protected peptides, as it would lead to the formation of multiple species.

In order to simplify this problem, the possibility of quickly and selectively forming a non-native covalent link between two unprotected peptides has been explored. These links include[15] disulfides, thioesters, thioethers, thiazolidines and oximes. An interesting example of this concept is the synthesis of the HIV-1 protease by Kent and Schnölzer[16]. They formed a thioester linkage between two peptides with a simple nucleophilic reaction between a thiocarboxylate and a bromoacetyl peptide, leading to a synthetic protein with full enzymatic activity.

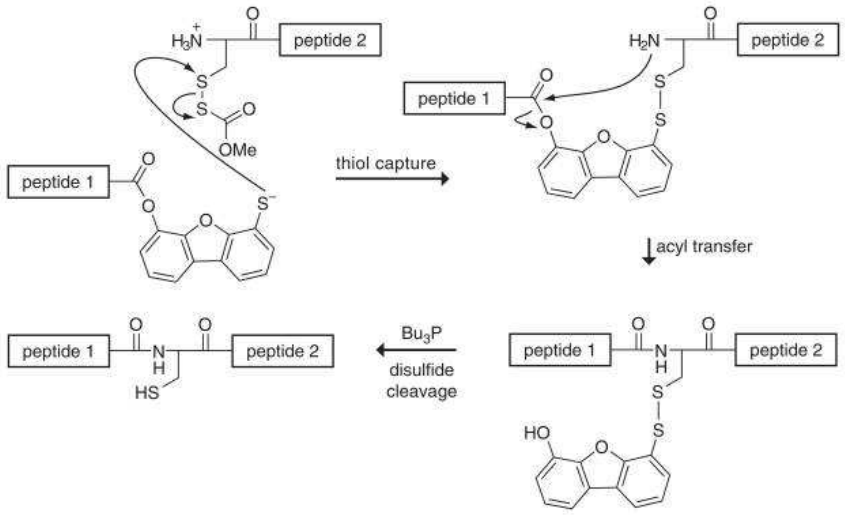
ii. Synthesis of native peptides and proteins

In parallel, new peptide ligation strategies emerged to selectively couple two unprotected peptides though the native amide bond, and generally consist of three steps. A rapid and selective reaction between two mutually reactive functional groups occurs, linking the two peptides in an initial capture step. The subsequent X→N acyl transfer (X = O, S) then occurs, favoured by proximity of the two reacting groups on the same molecule. Release of the capture moiety finally leads to a native polypeptide or protein.

<u>Prior thiol capture strategy</u>

In 1989 Kemp and coworkers[17] provided the first demonstration of chemoselective ligation of unprotected peptide fragments. In this approach (Scheme 3), the thiol-thioester exchange reaction was used in the capture step and a 4-hydroxy-6-mercaptodibenzofuran unit to bring the two coupling partners together, allowing the intramolecular O→N acyl transfer to occur. Finally, cleavage of the capture group afforded the native peptide.

15. Dawson, P. E.; Kent, S. B. Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
16. Schnolzer, M.; Kent, S. Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease. *Science* **1992**, *256*, 221–225.
17. Fotouhi, N.; Galakatos, N. G.; Kemp, D. S. Peptide Synthesis by Prior Thiol Capture. 6. Rates of the Disulfide-Bond-Forming Capture Reaction and Demonstration of the Overall Strategy by Synthesis of the C-Terminal 29-Peptide Sequence of BPTI. *J. Org. Chem.* **1989**, *54*, 2803–2817.



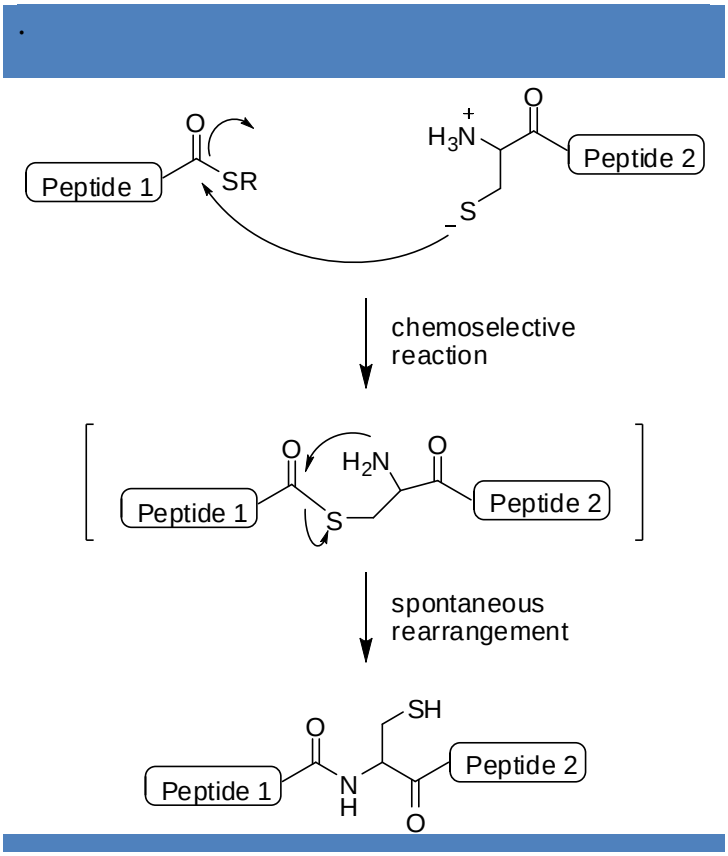
Scheme 3 | Prior thiol capture: first example of chemoselective ligation of unprotected peptides (from ref. 18).

Native chemical ligation (NCL)

Introduced by Dawson *et al.* in 1994, native chemical ligation[19] is to date the most commonly used method to synthesize long peptide sequences. It consists in the chemoselective reaction between a C-terminal thioester and an N-terminal cysteine (Scheme 4). After initial thiol-thioester exchange, the resulting cysteine thioester immediately rearranges to give a stable amide bond via S→N acyl shift. This reaction is generally performed in mild conditions (room temperature, neutral pH) in an aqueous buffer and in the presence of chaotropes (i.e guanidinium chloride or urea), to discourage peptide aggregation.

Although Dawson *et al.* had the merit of developing this reaction into a practical method to ligate peptide fragments, its chemical foundation was discovered by Wieland *et al.*[20] 40 years earlier. The authors showed that reaction between Val-SPh and Cys-OH in an aqueous buffer led to the formation of the dipeptide Val-Cys-OH, via a cysteine thioester intermediate. However, they did not further develop these first findings.

18. Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical Synthesis of Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91–118.
19. Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266*, 776–779.
20. Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. Über Peptidsynthesen. 8. Mitteilung Bildung von S-Haltigen Peptiden Durch Intramolekulare Wanderung von Aminoacylresten. *Justus Liebigs Ann. Chem.* **1953**, *583*, 129–149.



Scheme 4 | Mechanism of native chemical ligation, leading to the formation of a native peptide bond.

Importantly, NCL is stereo-, regio- and chemo-selective. Addition of an exogenous thiol as catalyst results in reversibility of the initial thiol-thioester exchange and enables the formation of the target peptide in high yields even if internal cysteine residues are present on either peptide fragment. The NCL efficacy and practicity is strictly dependent on the properties of thioesters, that are more resistant to hydrolysis, but more reactive toward thiolysis and aminolysis than the corresponding oxoesters.

In general, near quantitative yields can be obtained in few hours. However, the ligation rate can be modulated with the addition of different exogenous thiols[21] and it is strictly dependent on the nature of the C-terminal amino acid. Although it can occur between cysteine and any other amino acid, NCL is significantly slower[22] in the case of hindered residues such as valine, isoleucine and proline.

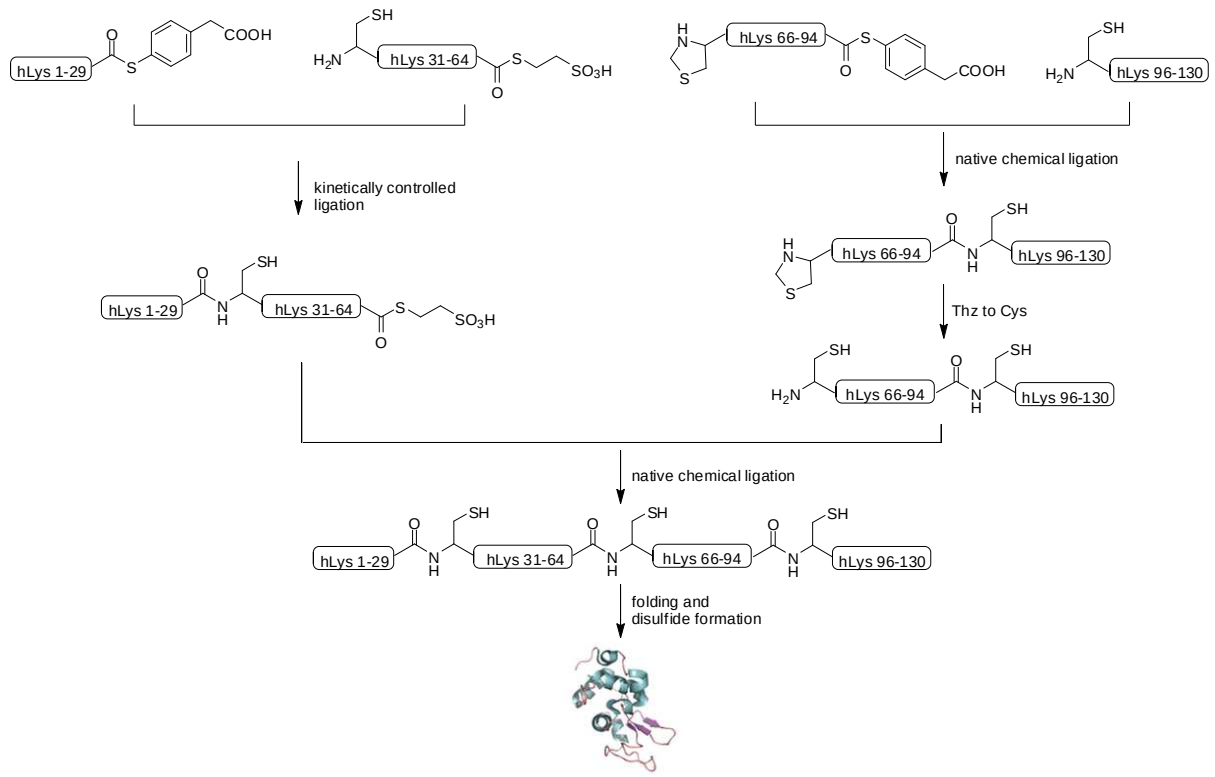
The versatility of native chemical ligation to enable the synthesis of a different peptide-based structures has been extensively investigated. Interestingly, NCL has also been

21. 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.
22. 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. U. S. A.* **1999**, *96*, 10068–10073.

used to couple unprotected peptides on a solid support[23] and to incorporate non-natural residues into polypepides. In a noteworthy example[24] a polyethylene glycol-modified erythropoyesis protein (166 residues) was synthesized with this strategy and showed a ~3-fold longer *in-vivo* lifetime than the natural protein.

Interestingly, native chemical ligation has been used to perform one-pot syntheses. For instance, crambin (46 residues)[25] was obtained by sequential coupling of three peptide fragments without purification of the intermediates. Another important advancement in NCL was the development of the kinetically controlled ligation[26], based on the slower reactivity of alkyl thioesters compared to aryl thioesters. As shown in Scheme 5, this strategy was used in combination with native chemical ligation for the synthesis[27] of the human lysozyme starting from four distinct peptides.

.



Scheme 5 | Convergent total synthesis of human lysozyme with full enzymatic activity using the kinetically controlled ligation approach.

23. Camarero, J. A.; Cotton, G. J.; Adeva, A.; Muir, T. W. Chemical Ligation of Unprotected Peptides Directly from a Solid Support. *J. Pept. Res.* **2009**, *51*, 303–316.
24. Kochendoerfer, G. G. Design and Chemical Synthesis of a Homogeneous Polymer-Modified Erythropoiesis Protein. *Science* **2003**, *299*, 884–887.
25. Bang, D.; Kent, S. B. H. A One-Pot Total Synthesis of Crambin. *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538.
26. Bang, D.; Pentelute, B. L.; Kent, S. B. H. Kinetically Controlled Ligation for the Convergent Chemical Synthesis of Proteins. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 3985–3988.
27. Durek, T.; Torbeev, V. Y.; Kent, S. B. H. Convergent Chemical Synthesis and High-Resolution X-Ray Structure of Human Lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 4846–4851.

Due to the fact that cysteine is the second least common amino acid found in proteins, the necessity of NCL for one of these residues at the ligation site is considered by many a severe limitation of this strategy. However, it has been shown that a cysteine can be introduced in a protein in various positions without causing misfolding nor loss of function. In some cases, alkylation of the non-native cysteine residue was performed[5].

In general, the necessity of a cysteine residue in the synthetic target is not a major issue and several strategies have been designed to circumvent it. For instance, if the native protein contains a methionine residue[28], the cysteine at the splice site can be replaced by a homocysteine residue. Methylation with methyl *p*-nitrobenzenesulfonate after completion of NCL then gives the native protein. A different approach consists in the removal[29] of the thiol group by means of catalytic desulfurization to afford the more common alanine residue. Taking advantage of these findings, different groups[30] have developed a plethora of NCL-like reactions using unnatural β- or γ-sulfanylamino acids followed by desulfurization to extend NCL to residues other than cysteine.

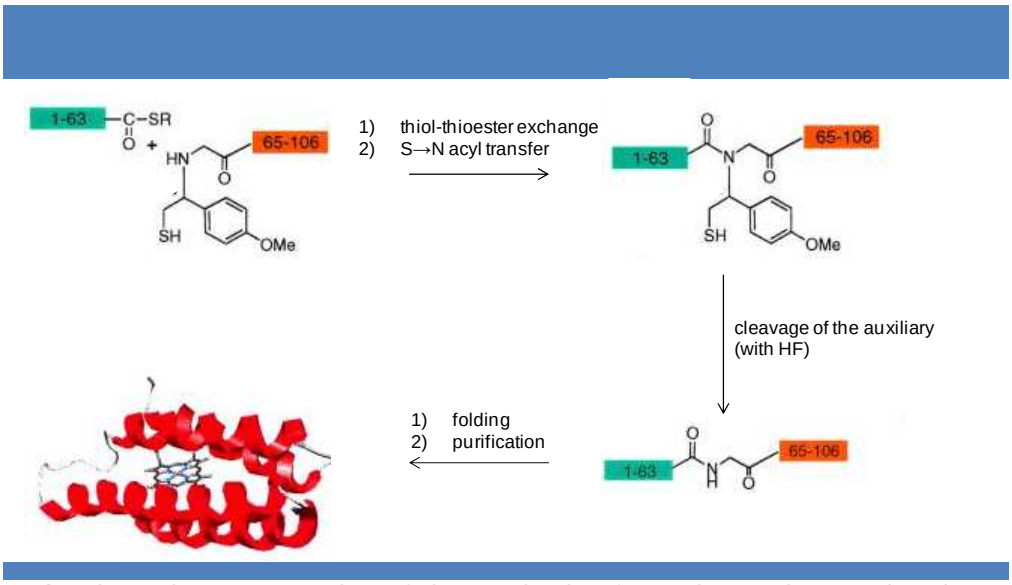
Importantly, all these synthetic strategies can be also applied to peptides containing selenocysteine. This amino acid (Sec or U) has low natural abundance but interesting properties. When used to replace cysteine in NCL, it has been shown[31] to accelerate the ligation rate. For instance, at pH 5 the reaction with selenocysteine is 1000-fold faster than with cysteine, due to the lower pKa of the selenol group. Moreover, replacement of cysteine with selenocysteine in proteins has led to their stabilization. This is due to the formation of selenodisulfide bonds, more resistant[32] to reducing conditions than the corresponding disulfide bridges.

Removable auxiliary strategy

The "removable auxiliary" strategy[33], introduced in 1996, is inspired by NCL and aims at overcoming the problems related to its need for a cysteine residue. It consists in the reaction between a C-terminal thioester and a thiol-containing group on the N-terminus of a peptide, that mimics the cysteine residue in NCL. After capture and S→N acyl shift, the auxiliary group is

28. Tam, J. P.; Yu, Q. Methionine Ligation Strategy in the Biomimetic Synthesis of Parathyroid Hormones. *Biopolymers* **1998**, *46*, 319–327.
29. 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.
30. Rohde, H.; Seitz, O. Ligation-Desulfurization: A Powerful Combination in the Synthesis of Peptides and Glycopeptides. *Biopolymers* **2010**, *94*, 551–559.
31. Hondal, R. J.; Nilsson, B. L.; Raines, R. T. Selenocysteine in Native Chemical Ligation and Expressed Protein Ligation. *J. Am. Chem. Soc.* **2001**, *123*, 5140–5141.
32. Besse, D.; Siedler, F.; Diercks, T.; Kessler, H.; Moroder, L. The Redox Potential of Selenocystine in Unconstrained Cyclic Peptides. *Angew. Chem. Int. Ed.* **1997**, *36*, 883–885.
33. Canne, L. E.; Bark, S. J.; Kent, S. B. H. Extending the Applicability of Native Chemical Ligation. *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896.

released to afford the target peptide. Starting from the initial N^{α}-ethanethiol and N^{α}-oxyethanethiol, other auxiliaries have been developed and this strategy has been successfully applied, for instance, to the total synthesis of cytochrome b562 (106 residues, Scheme 6)[34]. Despite these results, the ligation mechanism is highly sensitive to the nature of the two amino acids at the ligation site and this approach is generally less effective than NCL.



Scheme 6 | Schematic representation of the synthesis of cytochrome b562 using the removable auxiliary strategy.

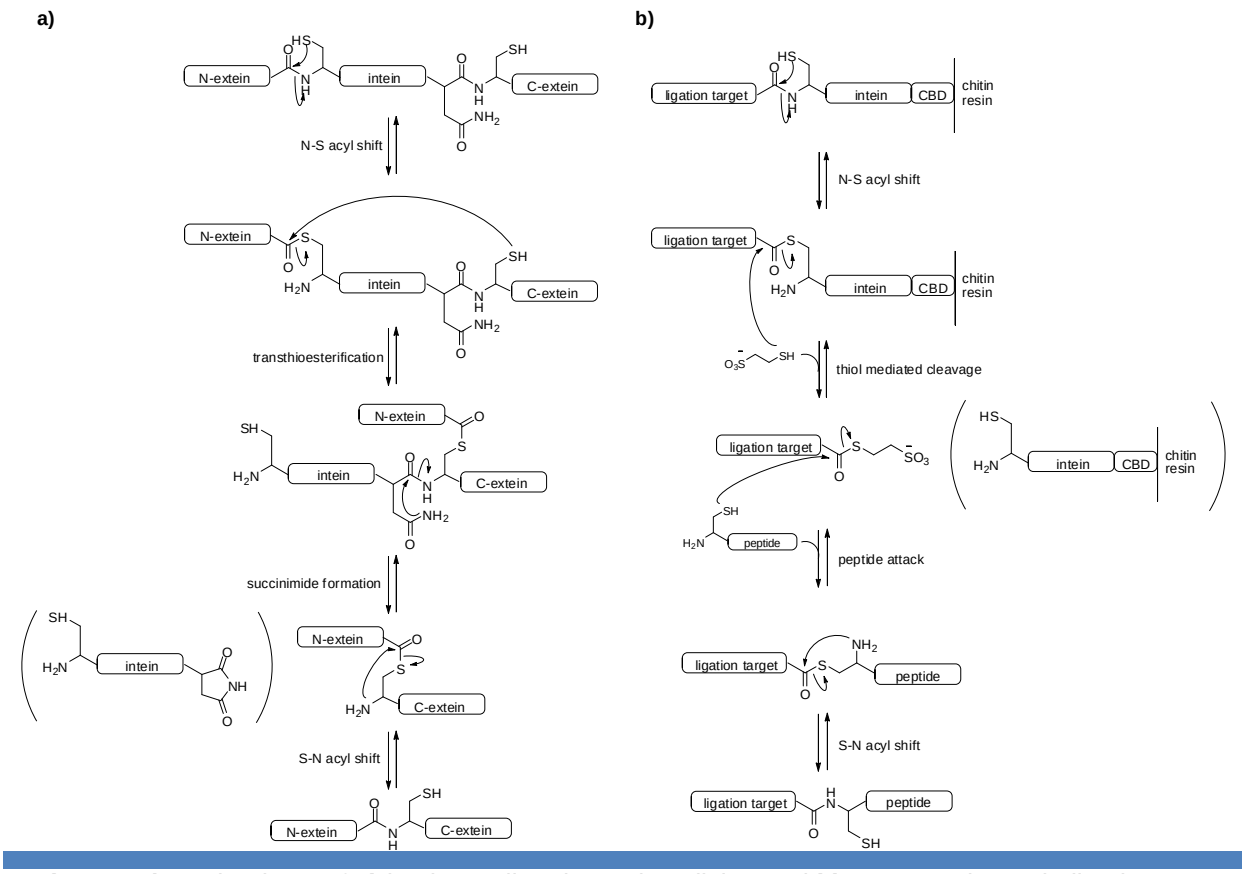
Expressed protein ligation

First described in 1998, this strategy[35] enables the synthesis of large proteins as well as the insertion of chemical modifications (e.g. glycosides, fluorophores, labeled isotopes) and takes advantage of the intein[36] auto-excision mechanism (Scheme 7a). Inteins are peptide fragments able to excise themselves from longer peptides or proteins while joining the remaining portions (called exteins) through a native peptide bond. Interestingly, the first step of this splicing mechanism is the formation of a thioester (or ester) due to intramolecular N→S acyl shift (or N→O if Ser or Thr replace Cys).

Expressed protein ligation consists in introducing a thioester at the target protein C-terminus by fusing it with a modified intein, attached to an affinity column. This resin-bound protein thioester can then be cleaved with a thiol reagent and the resulting thioester undergo native chemical ligation with an N-terminal cysteine peptide (Scheme 7b). Since the peptide is

34. Low, D. W.; Hill, M. G.; Carrasco, M. R.; Kent, S. B.; Botti, P. Total Synthesis of Cytochrome b562 by Native Chemical Ligation Using a Removable Auxiliary. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6554–6559.
35. a) Severinov, K.; Muir, T. W. Expressed Protein Ligation, a Novel Method for Studying Protein-Protein Interactions in Transcription. *J. Biol. Chem.* **1998**, *273*, 16205–16209. b) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed Protein Ligation: A General Method for Protein Engineering. *Proc. Natl. Acad. Sci.* **1998**, *95*, 6705–6710. c) Evans, T. C.; Benner, J.; Xu, M. Q. Semisynthesis of Cytotoxic Proteins Using a Modified Protein Splicing Element. *Protein Sci.* **1998**, *7*, 2256–2264.
36. Paulus, H. Protein Splicing and Related Forms of Protein Autoprocessing. *Annu. Rev. Biochem.* **2000**, *69*, 447–496.

prepared by SPPS, it is virtually possible to add any chemical modification at this level. This strategy merges the capability of biology to afford large recombinant proteins with the versatility of chemistry.



Scheme 7 | Mechanisms of **a)** intein-mediated protein splicing and **b)** expressed protein ligation. CBD = chitin binding domain.

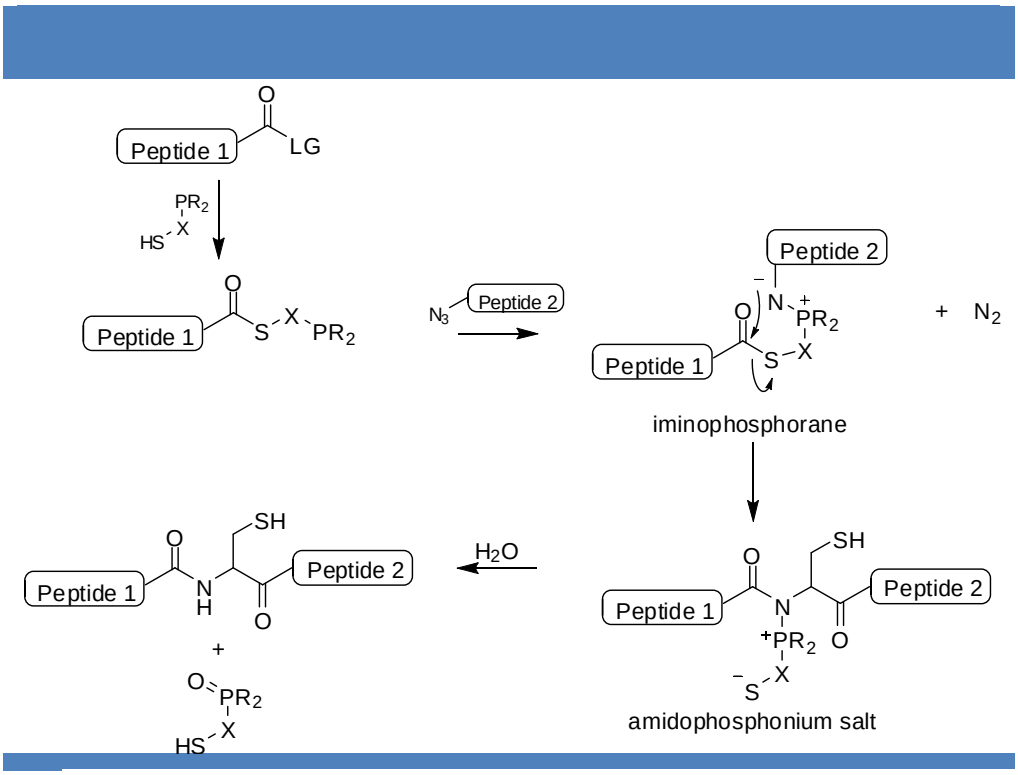
Conformationally assisted ligation

In case of peptides and proteins that have a high propensity for self-association, ligation can occur between a C-terminal thioester on one peptide and the N-terminal amino group of a second peptide, obviating the need for a cysteine residue. Despite its limited applicability, it has been used for the synthesis of different peptides and proteins, including RNAase A (124 residues)[37] and site-specifically labeled chymotrypsin inhibitor 2 (64 residues)[38].

37. Beligere, G. S.; Dawson, P. E. Conformationally Assisted Protein Ligation Using C-Terminal Thioester Peptides. *J. Am. Chem. Soc.* **1999**, *121*, 6332–6333.
38. Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S. Single-Molecule Protein Folding: Diffusion Fluorescence Resonance Energy Transfer Studies of the Denaturation of Chymotrypsin Inhibitor 2. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5179–5184.

Staudinger ligation

The Staudinger reaction[39] consists in the reduction of an azide into an amine by a phosphine, via the formation of a stable iminophosphorane intermediate. Further investigation[40] of the potential of this reaction has led to the development of a "traceless" Staudinger ligation[41] for protein synthesis, first reported by Raines and coworkers in 2000 (Scheme 8). A C-terminal phosphinorthoester on one peptide reacts with the N-terminal azide present on a second peptide to give an iminophosphorane intermediate. Subsequent intramolecular S→N acyl shift leads to an amidophosphonium salt that can be readily hydrolized to afford the target peptide.



Scheme 8 | Mechanism of the Staudinger ligation of peptides mediated by a phosphinothiol. LG = leaving group.

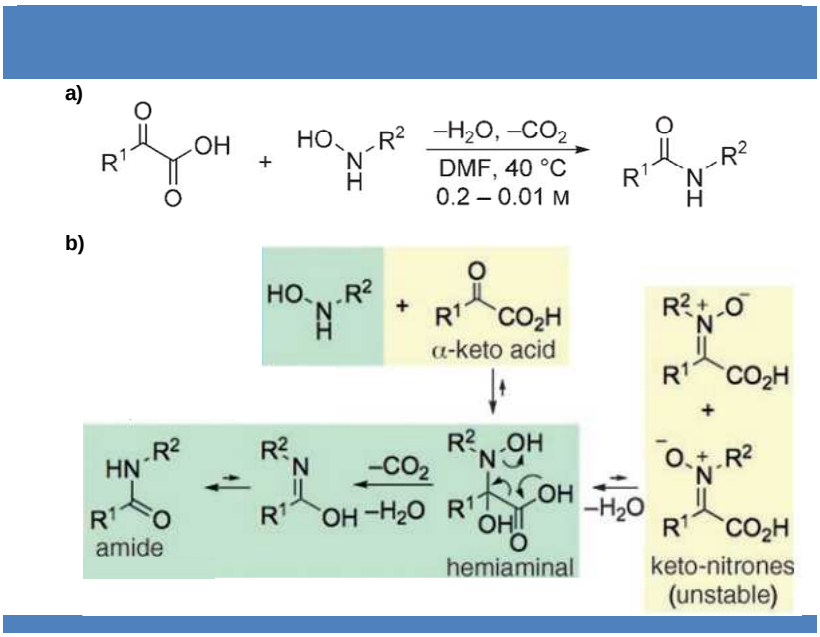
New phosphinothiols have been developed to improve the ligation rate and the reaction used not only in protein synthesis, but also in other areas of chemical biology[42]. Although the need for a cysteine residue is obviated, one of the limitations of this strategy is the need for at least one glycine at the ligation junction. In the case of two non-glycyl residues low yields

39. Staudinger, H.; Meyer, J. Über Neue Organische Phosphorverbindungen III. Phosphinmethylenderivate Und Phosphinimine. *Helv. Chim. Acta* **1919**, *2*, 635–646.
40. a) Bosch, I.; Romea, P.; Urpí, F.; Vilarrasa, J. Alternative Procedures for the Macrolactamisation of Ω-Azido Acids. *Tetrahedron Lett.* **1993**, *34*, 4671–4674. b) Saxon, E.; Bertozzi, C. E. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000**, *287*, 2007–2010.
41. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. Staudinger Ligation: A Peptide from a Thioester and Azide. *Org. Lett.* **2000**, *2*, 1939–1941.
42. Köhn, M.; Breinbauer, R. The Staudinger Ligation-a Gift to Chemical Biology. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 3106–3116.

(20-50%) were obtained[43].

Decarboxylative condensation of α-ketoacids and N-alkylhydroxylamines

In this strategy, recently reported by Bode *et al.*[44], a peptide bond is formed by chemoselective reaction between a C-terminal α-ketoacid on one peptide and an N-terminal alkylhydroxylamine on a second peptide (Scheme 9). This approach does not involve a cysteine residue and has the advantage of producing only carbon dioxide and water as byproducts. However, the yields are strictly dependent on the nature of the two amino acids at the ligation site and are generally lower than those obtained with native chemical ligation.



Scheme 9 | a) Schematic representation of the reaction between an α-ketoacid and a N-alkylhydroxylamine, leading to the formation of an amide bond and **b)** insight into the mechanism of this reaction (adapted from ref. 44).

2. Dynamic combinatorial chemistry

Dynamic combinatorial chemistry rests on bond reversibility. Basic components are reversibly connected through covalent or non-covalent interactions (i.e. at the molecular or supramolecular level) and all the possible combinations of this initial set of building blocks constitute a dynamic combinatorial library (DCL). Importantly, all the members of a DCL are in

43. 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.
44. 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.

thermodynamic equilibrium. The possibility for the reversible interactions to be continuously formed and disrupted allows the system to adapt in the presence of external stimuli[45].

This property of DCLs has been exploited in many different research fields[46]. In particular, it is used in drug discovery[47] for the selection of species with high affinity for a given target (Figure 2c). The thermodynamic gain upon binding of some library members to the target induces a bias in the system. Hence, the DCL adapts in order to favour the formation of the best binders at equilibrium. Other strategies for the identification of lead compounds include the combinatorial approach (Figure 2a) and the kinetic target-guided synthesis (TGS, Figure 2b)[48].

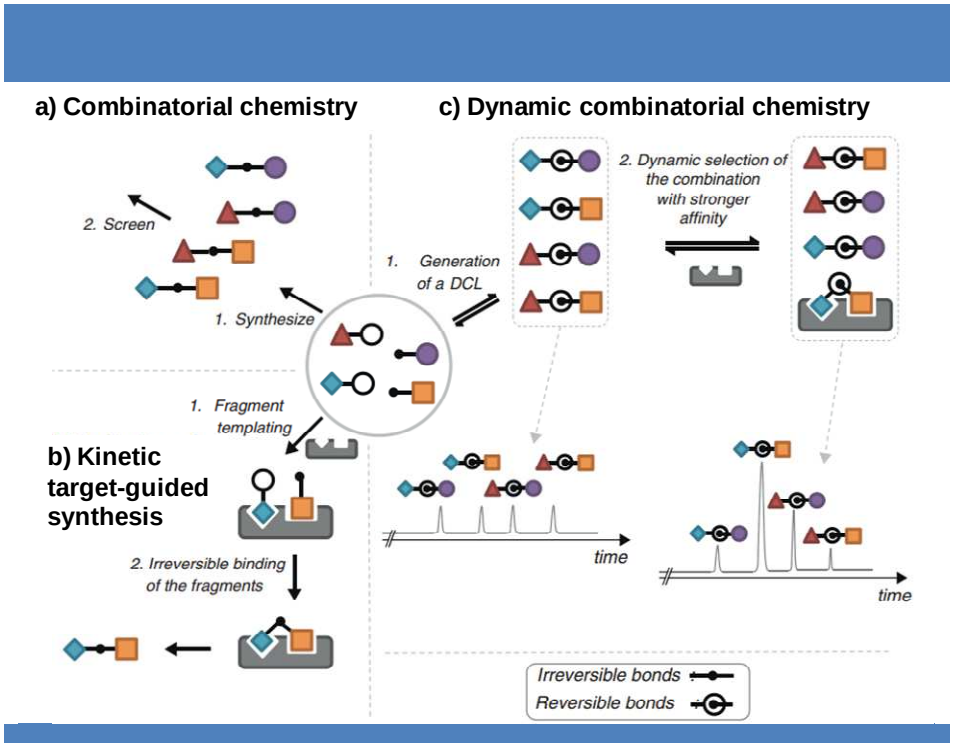


Figure 2 | Strategies used for identifying lead compounds from a chemical library (adapted from ref. 49).

In the classical combinatorial approach a chemical library is synthesized first and then screened in a subsequent step. This strategy has been very actively pursued by pharmaceutical companies. The repetitive application of specific chemical reactions to one (or more) starting unit has led to the generation of combinatorial libraries whose components have high structural diversity. However, the need for synthesis and purification of an enormous number of molecules prior to screening is a serious limitation of this approach. Similarly to the DCC

45. Lehn, J.-M. Dynamic Combinatorial Chemistry and Virtual Combinatorial Libraries. *Chem. - A Eur. J.* **1999**, *5*, 2455–2463.
46. a) Moulin, E.; Cormos, G.; Giuseppone, N. Dynamic Combinatorial Chemistry as a Tool for the Design of Functional Materials and Devices. *Chem. Soc. Rev.* **2012**, *41*, 1031–1049. b) Zhang, Y.; Barboiu, M. Constitutional Dynamic Materials—Toward Natural Selection of Function. *Chem. Rev.* **2015**, 150716084026002.
47. Ramström, O.; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Discov.* **2002**, *1*, 26–36.
48. Hu, X.; Manetsch, R. Kinetic Target-Guided Synthesis. *Chem. Soc. Rev.* **2010**, *39*, 1316–1324.
49. Melnyk, O.; Agouridas, V. From Protein Total Synthesis to Peptide Transamidation and Metathesis: Playing with the Reversibility of N,S-Acyl or N,Se-Acyl Migration Reactions. *Curr. Opin. Chem. Biol.* **2014**, *22*, 137–145.

strategy, kinetic TGS combines the synthesis and selection processes by synthesizing the compounds in the presence of the target molecule (e.g. a receptor). Starting from two sets of fragments, those with high affinity to the target are initially bound to it separately and then connected through a chemoselective and irreversible reaction, often a Huisgen's cycloaddition[50] between an azide and an alkyne (i.e. click chemistry).

a. Design of a DCL

In dynamic combinatorial chemistry the design of the initial building blocks and of the exchange experiments is not a trivial issue. For what concerns the initial set of components, it is important that the functional groups involved in the exchange reaction are conveniently exposed (i.e. on different scaffolds, on linkers of appropriate length)[47]. However, the choice of the dynamic chemistry certainly plays a key role in the design of the experiments.

The exchange reaction has to meet several requirements[51]. Reversibility on a reasonable time scale is one of the most important ones and can eventually be obtained by addition of a catalyst. Moreover, as equilibration and selection ideally occur simultaneously, the reaction conditions should be mild enough not to interfere with the recognition process. Also, all the library members should in theory be soluble in the reaction medium and isoenergetic to avoid the introduction of bias in the system. Importantly, a desirable feature of a DCL is the possibility to freeze the exchange reaction (i.e. by changing the conditions, by adding quenching reagents) to facilitate the analysis of the system.

As already mentioned above, many reversible covalent and non-covalent chemistries have been extensively explored (Table 1). Among them, addition-elimination reactions at carbonyl groups are by far the most important class, especially the imine exchange one. Interestingly, intramolecular configurational and conformational changes, such as *cis-trans* isomerization or internal rotation, can also be used to produce DCLs.

50. Millward, S. W.; Agnew, H. D.; Lai, B.; Lee, S. S.; Lim, J.; Nag, A.; Pitram, S.; Rohde, R.; Heath, J. R. In Situ Click Chemistry: From Small Molecule Discovery to Synthetic Antibodies. *Integr. Biol.* **2013**, *5*, 87–95.
51. Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. Dynamic Combinatorial Chemistry. *Chem. Rev.* **2006**, *106*, 3652–3711.

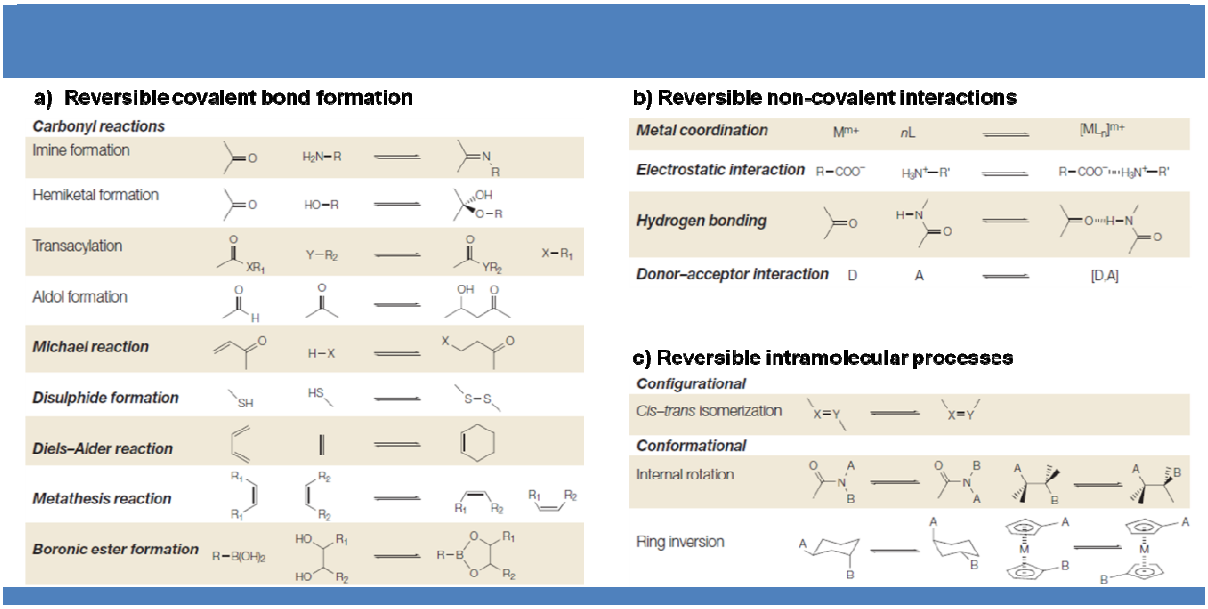


Table 1 | Selection of reversible reactions or intramolecular processes used for dynamic combinatorial chemistry (adapted from ref. 47).

b. Selection mechanisms

In a dynamic combinatorial library, the distribution of its components depends on a number of factors and can thus be controlled in different fashions[52].

External templating has been by far the most extensively used approach to direct a library's distribution. In this strategy, a DCL equilibrates in the presence of a given template molecule to select and amplify the component that best binds it. Importantly, this template may either be a guest or a host molecule. While in the first case the building blocks are combined to form a receptor for the given molecule (molding, Figure 3a), in the second instance the initial components of the DCL form a species that binds or stabilizes a macromolecule (casting, Figure 3b). Different ligands for biomolecules have been identified with this approach. However, when dealing with small quantities of sensitive biological targets, it is often convenient to separate the generation of the library from the screening phase[47]. This approach consists in pre-equilibrating the DCL and then quenching the reversible interactions to identify the compounds under static conditions. Since no amplification takes place, the best binders can be identified with classical deconvolution techniques.

52. Beeren, S. R.; Sanders, J. K. M. History and Principles of Dynamic Combinatorial Chemistry. In *Dynamic Combinatorial Chemistry*; Reek, J. N. H.; Otto, S., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2010; pp. 1–22.

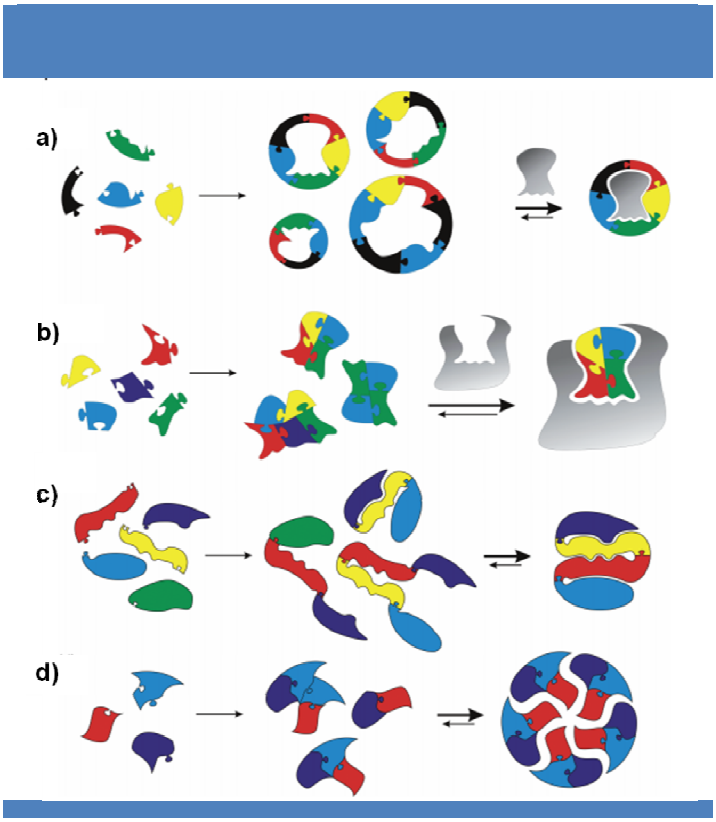


Figure 3 | Selection of a library member through by addition of **a)** a guest or **b)** a host molecule or through **c)** intramolecular or **d)** intermolecular self-templating (adapted from ref. 52).

Selection can also occur through internal templating: members of a DCL are stabilized by non-covalent intramolecular or intermolecular interactions and are thus amplified. Intramolecular self-templating (Figure 3c) occurs in those species that can fold upon themselves. This approach can thus be used to promote the formation of stable foldamers[53] or to investigate the folding of natural and synthetic polymers. In intermolecular self-templating (Figure 3d), non-covalent interactions among the members of the library promote the stabilization and amplification of a new molecule[54].

Selection of a library member can also occur in response to a physical stimulus such as a variation of temperature or pH[55], modulation of the electric field[56] or irradiation[57]. This is particularly important in dynamic materials that can adapt and reorganize in response to these stimuli. Finally, a phase change, either gelation or crystallization, can drive the formation and amplification of one library member. An important class of molecules selected with this

53. Oh, K.; Jeong, K.-S.; Moore, J. S. M-Phenylene Ethynylene Sequences Joined by Imine Linkages: Dynamic Covalent Oligomers. *J. Org. Chem.* **2003**, *68*, 8397–8403.
54. Xu, S.; Giuseppone, N. Self-Duplicating Amplification in a Dynamic Combinatorial Library. *J. Am. Chem. Soc.* **2008**, *130*, 1826–1827.
55. Giuseppone, N.; Lehn, J.-M. Protonic and Temperature Modulation of Constituent Expression by Component Selection in a Dynamic Combinatorial Library of Imines. *Chemistry* **2006**, *12*, 1715–1722.
56. Giuseppone, N.; Lehn, J.-M. Electric-Field Modulation of Component Exchange in Constitutional Dynamic Liquid Crystals. *Angew. Chem. Int. Ed.* **2006**, *45*, 4619–4624.
57. Ingerman, L. A.; Waters, M. L. Photoswitchable Dynamic Combinatorial Libraries: Coupling Azobenzene Photoisomerization with Hydrazone Exchange. *J. Org. Chem.* **2009**, *74*, 111–117.

strategy is that of dynamic polymers, "dynamers"[58], generated by reversible linkage of dynamic monomers.

c. Further considerations and applications

The dynamic combinatorial approach has been used in a number of different research fields. If its most obvious application is probably drug discovery[59], catalysts[60] and sensors[61] have also been developed, as well as dynamic polymers, whose adaptive nature confers them interesting materials properties[62]. Interestingly, some dynamers have been commercially developed as biocompatible materials for surgery.

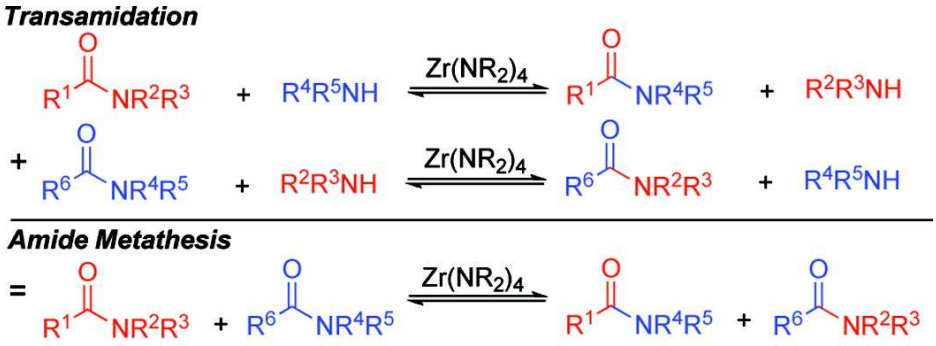
Dynamic processes are involved in all supramolecular assemblies and the list of possible applications for DCC is in continuous expansion. Importantly, dynamic self-assembly properties can be exploited to build increasingly small devices through a bottom-up approach.

3. Peptides and proteins in dynamic combinatorial chemistry

As discussed above, dynamic combinatorial chemistry is one of the most intriguing strategies for the identification of novel molecules with specific binding or physico-chemical properties. Although many different reversible chemistries have been extensively investigated, the stable amide bond has barely been considered. Its intrinsic stability is due to the C-N partial double bond character[63]: the nitrogen can donate its ion pair to the neighbouring carbon and the electrons of the carbonyl group can be pushed toward the electronegative oxygen, resulting in a stable resonance structure. Compared to its amine counterpart, the amide C-N bond is shorter (1.32 Å)[64] and has a larger rotation barrier (15-20 kcal/mol)[65], resulting in higher stability.

58. Chow, C.-F.; Fujii, S.; Lehn, J.-M. Crystallization-Driven Constitutional Changes of Dynamic Polymers in Response to Neat/solution Conditions. *Chem. Commun.* **2007**, 4363–4365.
59. Nicolaou, K. C.; Hughes, R.; Cho, S. Y.; Winssinger, N.; Smethurst, C.; Labischinski, H.; Endermann, R. Target-Accelerated Combinatorial Synthesis and Discovery of Highly Potent Antibiotics Effective Against Vancomycin-Resistant Bacteria. *Angew. Chem. Int. Ed.* **2000**, *39*, 3823–3828.
60. Vial, L.; Sanders, J. K. M.; Otto, S. A Catalyst for an Acetal Hydrolysis Reaction from a Dynamic Combinatorial Library. *New J. Chem.* **2005**, *29*, 1001.
61. Buryak, A.; Severin, K. Dynamic Combinatorial Libraries of Dye Complexes as Sensors. *Angew. Chem. Int. Ed.* **2005**, *44*, 7935–7938.
62. Roy, N.; Bruchmann, B.; Lehn, J.-M. DYNAMERS: Dynamic Polymers as Self-Healing Materials. *Chem. Soc. Rev.* **2015**, *44*, 3786–3807.
63. Corey, R. B. X-Ray Diffraction Studies of Crystalline Amino Acids and Peptides. *Progress in the Chemistry of Organic Natural Products*; Springer Vienna: Vienna, 1951; pp. 310–340.
64. Corey, R. B.; Pauling, L. Fundamental Dimensions of Polypeptide Chains. *Proc. R. Soc. B Biol. Sci.* **1953**, *141*, 10–20.
65. Kemnitz, C. R.; Loewen, M. J. "Amide Resonance" Correlates with a Breadth of C-N Rotation Barriers. *J. Am. Chem. Soc.* **2007**, *129*, 2521–2528.

The only examples of reversible amide bonds in the literature come from the Ståhl group, that improved initial findings[66] and achieved transamidation and amide metathesis[67] at room temperature and in the presence of a zirconium catalyst (Scheme 10).



Scheme 10 | Transamidation and amide metathesis at room temperature (adapted from ref. 67).

Unfortunately, these remarkable results can not be applied to the peptide bond. The necessity to operate in organic solvents and to use metal catalysts together with the lack of selectivity for a specific bond are serious limitations of this approach.

a. Reversible chemistries leading to non-native peptide bonds

Amino acids and small peptides have been among the first molecules to be incorporated into dynamic combinatorial libraries. However, achieving reversibility at the peptide bond level was a major challenge. Other chemistries have thus been investigated, leading to non-native peptide molecules. These reversible reactions have low activation barriers, are chemoselective and are generally compatible with mild aqueous conditions (neutral pH, acceptable temperature range). Interestingly, of the many that fulfill these criteria, only three covalent chemistries have been applied to peptide-based DCLs.

i. Amine-imine exchange reaction

This type of reversible chemistry includes the hydrazone linkage that was used, for

66. Eldred, S. E.; Stone, D. A.; Gellman, S. H.; Stahl, S. S. Catalytic Transamidation under Moderate Conditions. *J. Am. Chem. Soc.* **2003**, *125*, 3422–3423. b) Hoerter, J. M.; Otte, K. M.; Gellman, S. H.; Cui, Q.; Stahl, S. S. Discovery and Mechanistic Study of Al(III)-Catalyzed Transamidation of Tertiary Amides. *J. Am. Chem. Soc.* **2008**, *130*, 647–654.
67. Stephenson, N. A.; Zhu, J.; Gellman, S. H.; Stahl, S. S. Catalytic Transamidation Reactions Compatible with Tertiary Amide Metathesis under Ambient Conditions. *J. Am. Chem. Soc.* **2009**, *131*, 10003–10008.

instance, by Sanders and coworkers[68]. They provided an interesting example of how DCC can lead to unpredictable structures. From a library of dipeptide hydrazones, of all the possible exchange products, a [2]-catenane consisting of two interlocked macrocyclic trimers was selected as receptor for the neurotransmitter acetylcholine (Figure 4).

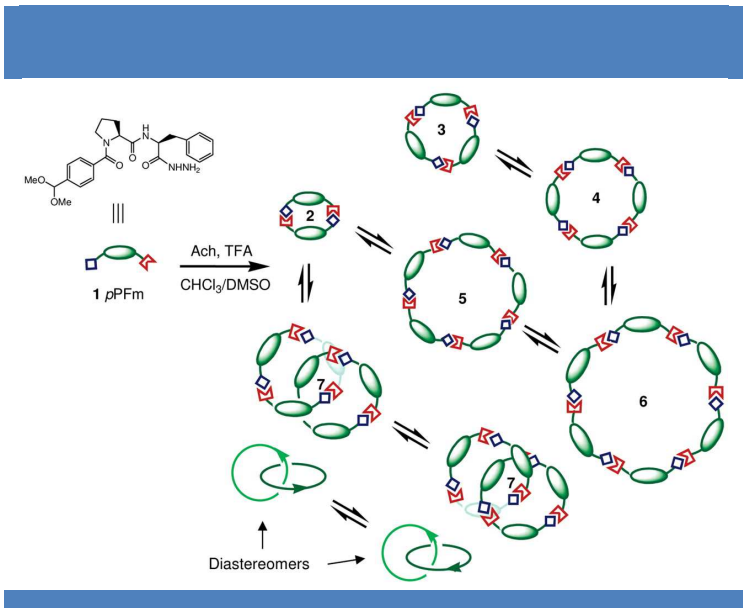


Figure 4 | Dynamic combinatorial library based on dipeptide hydrazones (from ref. 68).

ii. Thiol-thioester exchange reaction

Peptides with one or more amide bonds replaced with a thioester are called thiodepsipeptides. Ghosh *et al.*[69] reported the generation of a thiodepsipeptide macrocycle library from a short peptide incorporating both functionalities: a cysteine residue, hence a thiol group, and a C-terminal thioester (Figure 5).

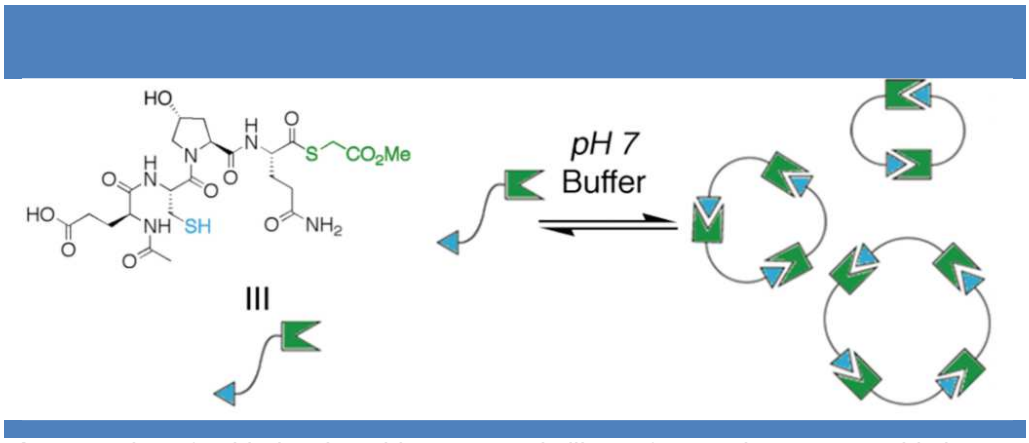


Figure 5 | Generation of a thiodepsipeptide macrocycle library from a short tetrapeptide incorporating a thiol and a thioester group (from ref. 69).

68. Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. Amplification of Acetylcholine-Binding Catenanes from Dynamic Combinatorial Libraries. *Science* **2005**, *308*, 667–669.
69. Ghosh, S.; Ingerman, L. A.; Frye, A. G.; Lee, S. J.; Gagné, M. R.; Waters, M. L. Dynamic Cyclic Thiodepsipeptide Libraries from Thiol-Thioester Exchange. *Org. Lett.* **2010**, *12*, 1860–1863.

In a more refined example, Ashkenasy and coworkers[70] investigated the behavior of thiodepsipeptides forming coiled-coil structures, in the presence of thiol molecules and in response to the introduction of changes in the folding conditions or of an external template.

iii. Thiol-disulfide exchange reaction

Using this reversible chemistry, Rauschenberg *et al.*[71] selected cyclic receptors for different carbohydrates from simple Cys-aa-Cys tripeptides (Figure 6). In particular, one member of the library was selected in the presence of N-acetyl neuraminic acid (NANA), a major component of glycoconjugates.

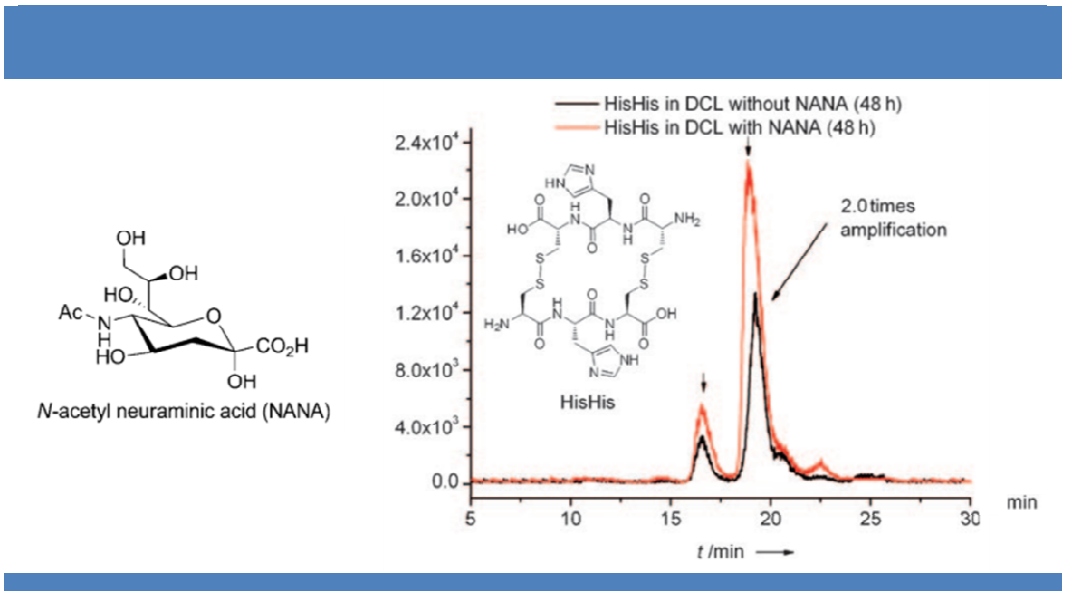


Figure 6 | Selection of the cyclic receptor for NANA, formed by two Cys-His-Cys tripeptides (adapted from ref. 71).

iv. Non-covalent interactions

In the Case group[72] hydrophobic and metal-ion interactions have been exploited to select stable trimeric proteins from a library of peptide helices. After an initial recognition step due to the formation of an hydrophobic core, the selected species were further stabilized by addition of a substoichiometrical amount of Fe^{2+}. This was accomplished by appending metal-binding bipyridine ligands (Bpy) to the peptide subunits (Figure 7).

70. Dadon, Z.; Samiappan, M.; Shahar, A.; Zarivach, R.; Ashkenasy, G. A High-Resolution Structure That Provides Insight into Coiled-Coil Thiodepsipeptide Dynamic Chemistry. *Angew. Chem. Int. Ed.* **2013**, *52*, 9944–9947.
71. Rauschenberg, M.; Bomke, S.; Karst, U.; Ravoo, B. J. Dynamic Peptides as Biomimetic Carbohydrate Receptors. *Angew. Chem. Int. Ed.* **2010**, *49*, 7340–7345.
72. Roy, L.; Case, M. A. Protein Core Packing by Dynamic Combinatorial Chemistry. *J. Am. Chem. Soc.* **2010**, *132*, 8894–8896.

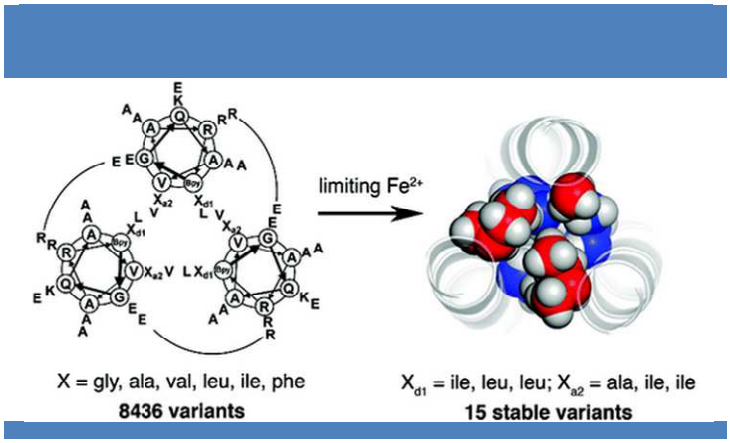


Figure 7 | Stabilization of the trihelical structure by interaction of the bibypidine units (Bpy) with the metal ion (from ref. 72).

An outstanding example of the potential of supramolecular interactions has been recently reported by Song and Tezcan [73]. They obtained an artificial metallo-β-lactamase by metal-directed self-assembly of a monomeric redox protein. The resulting tetrameric structure was tested in *Escherichia coli* and showed full enzymatic activity (Figure 8).

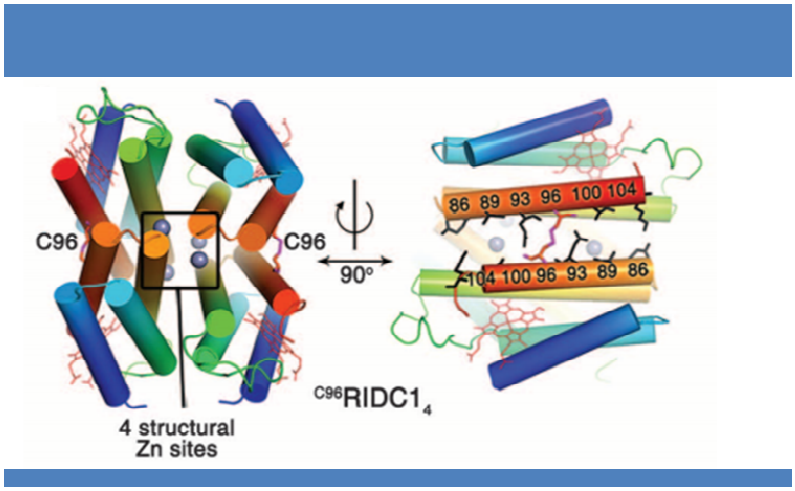


Figure 8 | Structure of the tetrameric Zn complex with enzymatic activity (adapted from ref. 73).

b. Reversible peptide bond: The enzyme approach

An interesting approach to scramble peptide fragments at the native amide bond consists in the use of proteases or peptidases. It had already been shown[74] that these enzymes could be engineered in order to form, as opposed to break, peptide bonds. The synthesis of several proteins was achieved with enzyme-assisted synthesis, including the large RNase A (124

73. Song, W. J.; Tezcan, F. A. A Designed Supramolecular Protein Assembly with in Vivo Enzymatic Activity. *Science* **2014**, *346*, 1525–1528.
74. Kullmann, W. Protease-Catalyzed Peptide Bond Formation: Application to Synthesis of the COOH-Terminal Octapeptide of Cholecystokinin. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 2840–2844.

residues)[75]. Since these enzymes can catalyse both the formation and the disruption of the peptide bond, they can be used to generate DCLs.

Although protein engineering is a powerful tool, simpler methods exist to shift the equilibrium of native proteases from hydrolysis to synthesis. These include the use of non-aqueous cosolvents and the addition of a large excess of one of the reactants. Moreover, if the synthetic product precipitates or is captured by a molecular trap, the equilibrium of the system will be shifted toward its formation. A very elegant application of this last approach was described by Swann *et al.*[76] and consists in the generation of a peptide fragment library by thermolysine, a soluble protease with broad specificity. When this experiment was carried out in the presence of an antibody specific for the N-terminus of β-endorphin, amplification was detected. Importantly, the antibody was protected from proteolysis by a semipermeable membrane (Figure 9) and the peptide fragments could migrate between the two compartments.

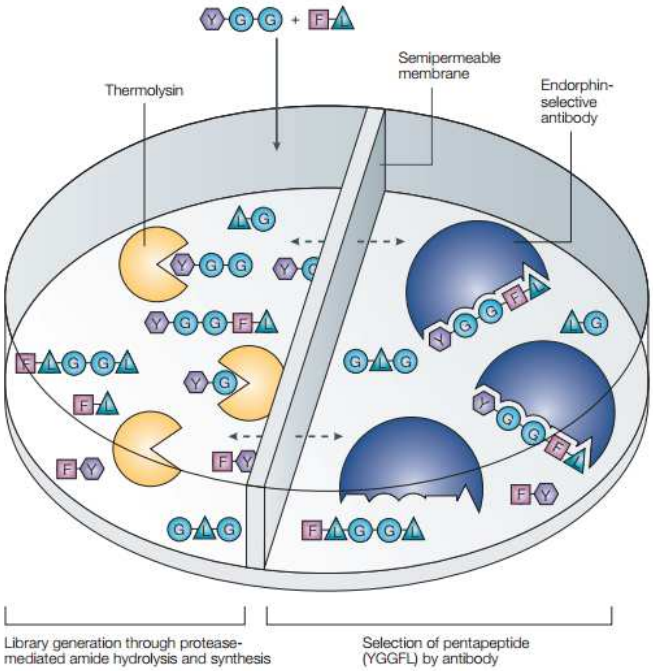


Figure 9 | Generation and screening of a peptide library from Tyr-Gly-Gly (YGG) and Phe-Leu (FL) fragments. Thermolysine catalysis both peptide hydrolysis and synthesis and is kept in a separate compartment from the antibody by a semipermeable membrane. Amplification of the specific segment YGGFL was observed (from ref. 47).

75. Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. A Designed Peptide Ligase for Total Synthesis of Ribonuclease A with Unnatural Catalytic Residues. *Science* **1994**, *266*, 243–247.
76. Swann, P. G.; Casanova, R. A.; Desai, A.; Frauenhoff, M. M.; Urbancic, M.; Slomczynska, U.; Hopfinger, A. J.; Le Breton, G. C.; Venton, D. L. Nonspecific Protease-Catalyzed Hydrolysis/synthesis of a Mixture of Peptides: Product Diversity and Ligand Amplification by a Molecular Trap. *Biopolymers* **1996**, *40*, 617–625.

Thermolysine was also used by Ulijn and coworkers, together with other enzymes[77], to promote supramolecular gelation. For instance[78], they created a dynamic combinatorial library of Fmoc-protected tripeptides by using thermolysine to catalyse peptide bond formation, hydrolysis and transamidation. In the aqueous buffer some of these library members were able to self-assemble into fibers via hydrophobic interactions and/or π-π stacking, thus triggering hydrogel formation.

In another example, the use of sortase A (SrtA) was investigated[79] for N-terminal labeling of proteins. This enzyme normally catalyzes the cleavage of the peptide bond between the threonine and the glycine residues of a LPXTG motif. However, it can also be instructed to promote the synthesis of the same bond.

Although interesting, the enzyme-based approach has several limitations. First, shifting the equilibrium from the native hydrolytic reaction toward the formation of peptide bonds is not trivial. Second, enzymes are generally sensitive biomolecules and instability issues might emerge over the long equilibration times that are often needed in DCC. In addition to this, proteases are by definition incompatible with proteic targets, which further complicates the reaction setup. Finally, they are either poorly selective for a given peptide bond (i.e. thermolysine) or too specific for a given sequence (i.e. sortase). Hence, specific conditions and setups need to be defined on a case-by-case basis. Unfortunately, this enzyme-based approach can not be considered a general methodology for scrambling peptide fragments.

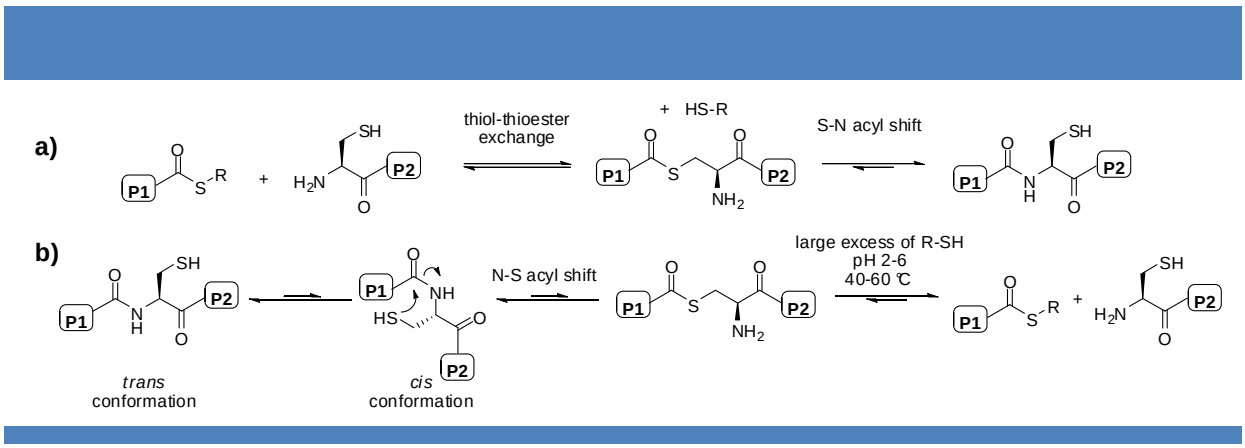
77. Fleming, S.; Ulijn, R. V. Design of Nanostructures Based on Aromatic Peptide Amphiphiles. *Chem. Soc. Rev.* **2014**, *43*, 8150–8177.
78. Toledano, S.; Williams, R. J.; Jayawarna, V.; Ulijn, R. V. Enzyme-Triggered Self-Assembly of Peptide Hydrogels via Reversed Hydrolysis. *J. Am. Chem. Soc.* **2006**, *128*, 1070–1071.
79. Williamson, D. J.; Fascione, M. A.; Webb, M. E.; Turnbull, W. B. Efficient N-Terminal Labeling of Proteins by Use of Sortase. *Angew. Chem. Int. Ed.* **2012**, *51*, 9377–9380.

Chapter 2 – Peptides Incorporating N-(2-Thioethyl)-Cysteine

1. Introduction: Design and objectives

As already discussed in Chapter 1, scrambling peptide fragments at the amide bond level is not trivial. Although the enzyme approach has been successfully used in some specific cases, it has several limitations.

We wanted to develop an abiotic methodology to achieve reversibility of this strong covalent bond in mild aqueous conditions. To do so, we investigated the potential of native chemical ligation. As already discussed in Chapter 1, NCL (Scheme 11a) is the most popular method to synthesize long peptides and consists in the selective reaction of a C-terminal thioester with an N-terminal cysteine to form a native peptide bond in mild aqueous conditions. In contrast, the reverse NCL (Scheme 11b) for the synthesis of peptide thioesters from peptides containing a C-terminal cysteine, requires harsh conditions (excess of thiols, acidic conditions and heating)[80]. This is due to the need for the peptide to assume the thermodynamically disfavored *cis* conformation in order for the initial N→S acyl transfer to occur. Hence, the biggest challenge in the development of a reversible native chemical ligation was to allow the first step of the reverse-NCL, the N→S acyl shift, to occur at low temperature (20-40 °C) and under neutral to mildly acidic conditions.



Scheme 11 | a) Forward and **b)** reverse native chemical ligation.

The large diffusion of native chemical ligation has led to the development of different strategies to widen its versatility. In particular, considerable efforts have been addressed to developing synthetic protocols for the synthesis of peptide thioesters compatible with Fmoc

80. a) Kang, J.; Richardson, J. P.; Macmillan, D. 3-Mercaptopropionic Acid-Mediated Synthesis of Peptide and Protein Thioesters. *Chem. Commun.* **2009**, 407–409. b) Macmillan, D.; Adams, A.; Premdjee, B. Shifting Native Chemical Ligation into Reverse through N→S Acyl Transfer. *Isr. J. Chem.* **2011**, *51*, 885–899.

SPPS. Among others, C-terminal N-sulfanylethylanilides (SEAlides)[81] have been shown to rearrange into transient thioesters by spontaneous intramolecular N→S acyl transfer. These first findings led to the development of the C-terminal bis(2-sulfanylethyl)amino (SEA) group. Two groups[82] independently reported that peptides containing this unit can undergo intramolecular N→S acyl shift at 37 °C and in neutral to mildly acidic conditions (Figure 10a). The exceptional reactivity of this tertiary amide was explained with the presence of two thiol groups, one of which is always correctly positioned for the N→S acyl shift to occur.

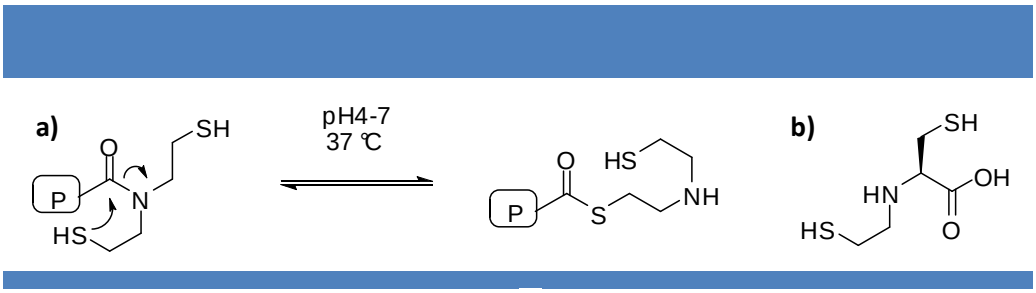


Figure 10 | a) Schematic representation of the N→S intramolecular acyl shift occurring in peptides bearing a C-terminal bis(2-sulfanylethyl)amino group. **b)** The non-natural dynamic amino acid N-(2-thioethyl)-cysteine (**Daa1**).

Taking advantage of this exceptional reactivity, we designed the modified non-natural amino acid N-(2-thioethyl)-cysteine (Figure 10b). In analogy to the bis(2-sulfanylethyl)amino group, this dynamic amino acid (**Daa1**) bears two thiol groups that can undergo N→S acyl shift. We introduced this dynamic unit in model peptides, either at the N-terminus or internally, and investigated the exchange reaction between them in mild aqueous conditions. We expected peptide fragments to be scrambled specifically at the junction between **Daa1** and the amino acid in α-position to this residue, according to a reversible NCL mechanism extensively described later in this Chapter. In order to facilitate the exchange reaction, initial investigations were carried out on peptides containing a glycine in that position, an amino acid that had been correlated[22] to fast NCL rates. All the other amino acid residues were randomly chosen.

2. Peptide synthesis

In order to investigate the dynamic nature of our system, we synthesized short model

81. Otaka, A.; Sato, K.; Ding, H.; Shigenaga, A. One-Pot/sequential Native Chemical Ligation Using N-Sulfanylethylanilide Peptide. *Chem. Rec.* **2012**, *12*, 479–490.
82. a) 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. b) Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-Sulfanylethyl)amino Native Peptide Ligation. *Org. Lett.* **2010**, *12*, 5238–5241. c) Dheur, J.; Ollivier, N.; Vallin, A.; Melnyk, O. Synthesis of Peptide Alkylthioesters Using the Intramolecular N,S-Acyl Shift Properties of bis(2-Sulfanylethyl)amido Peptides. *J. Org. Chem.* **2011**, *76*, 3194–3202.

peptides containing a **Daa1** unit either in N-terminal position or internally. Since the C-terminus of these peptides was not involved in the exchange reactions, we prepared the peptides as C-terminal amides, more stable and easier to synthesize than their C-terminal carboxylic acid counterparts. The two starting peptides **P1-Daa1-P2** (H_2N-Leu-Tyr-Lys-Gly-**Daa1**-Ala-Lys-Leu-Leu-$CONH_2$) and **Daa1-P3** (**Daa1**-Ala-Phe-Lys-Phe-$CONH_2$) and the exchange products **P1-Daa1-P3** (H_2N-Leu-Tyr-Lys-Gly-**Daa1**-Ala-Phe-Lys-Phe-$CONH_2$) and **Daa1-P2** (**Daa1**-Ala-Lys-Leu-Leu-$CONH_2$) were isolated in their oxidized form to facilitate their purification and are shown in Figure 11.

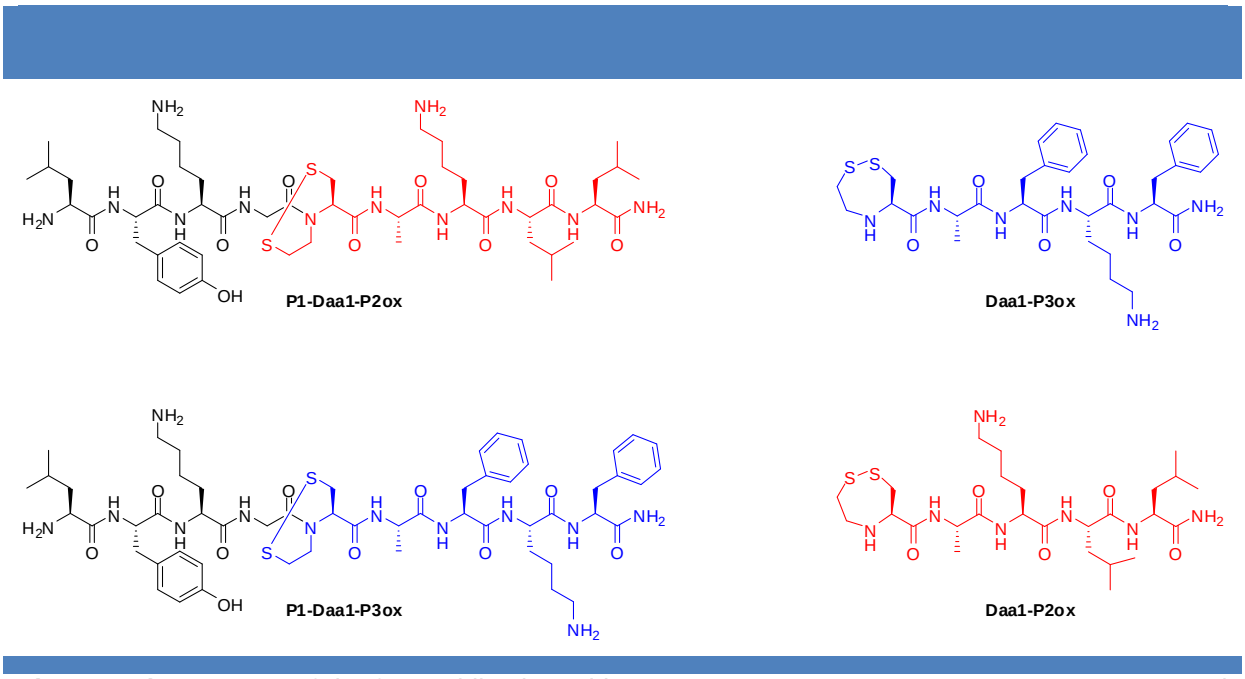


Figure 11 | Structures of the four oxidized peptides **P1-Daa1-P2ox**, **Daa1-P3ox**, **P1-Daa1-P3ox** and **Daa1-P2ox**.

a. General procedure

Reactions were typically performed on a Fmoc-Rink-amide-4-methylbenzhydrylamine (MBHA) resin crosslinked with 1% divinylbenzene at a 0.5 mmol scale (loading 0.74 mmol/g). This resin, derived from the Rink[83] resin, allows the synthesis of peptide amides and is typically cleaved by treatment with 95% TFA.

Unless stated otherwise, our model peptides were synthesized on a CEM Liberty 1 microwave synthesizer, using the Fmoc/*tert*-butyl strategy. Coupling were performed at 70 ºC

83. Rink, H. Solid-Phase Synthesis of Protected Peptide Fragments Using a Trialkoxy-Diphenyl-Methylester Resin. *Tetrahedron Lett.* **1987**, *28*, 3787–3790.

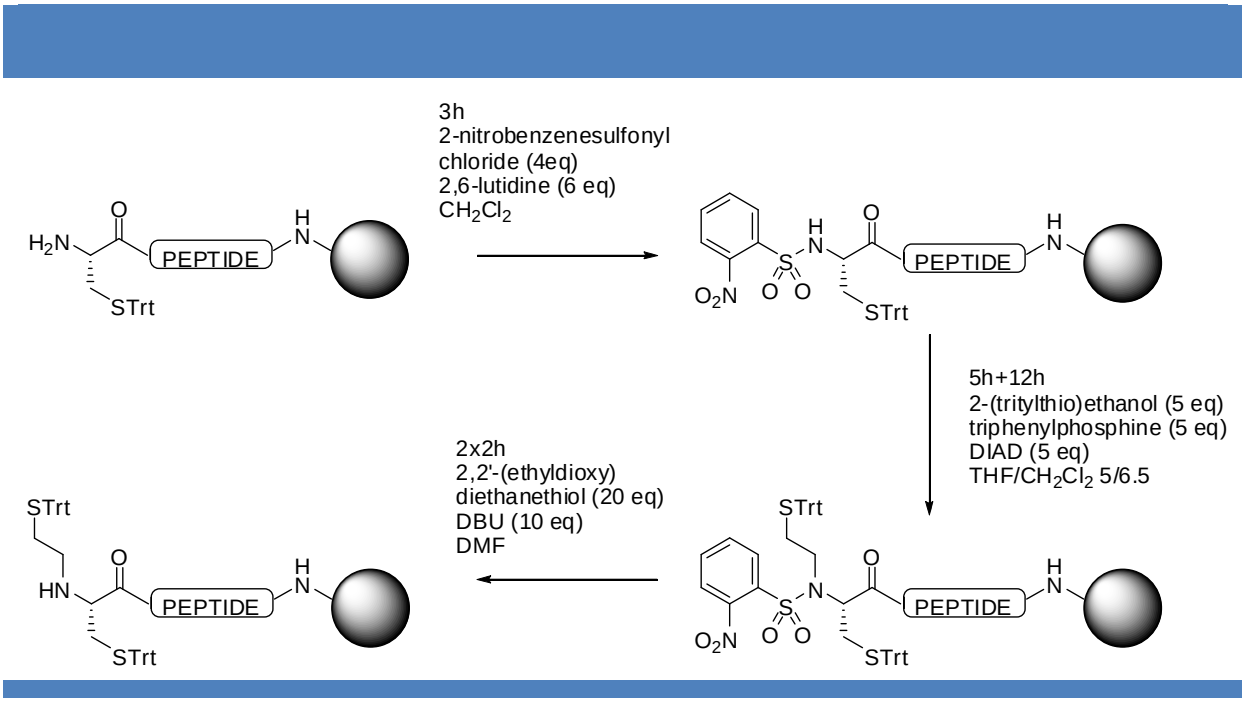
and 35 W (microwave power) for 5 min using 4-fold molar excess of each Fmoc-aa-OH (10 mL of a 0.2 M solution in DMF), HBTU (4 mL of a 0.5 M solution in DMF) and DIEA (2 mL of a 2 M solution in NMP). Being sensitive to racemization, Fmoc-Cys(Trt)-OH was coupled for 2 min without microwave heating and then at 50 ºC with 25 W for 4 min. Fmoc groups were deprotected with 2 successive treatments with a 20 %v/v piperidine solution in DMF (15 mL) at 70 ºC and 55 W for 3 min.

b. Insertion of the 2-thioethyl chain

In order to introduce the thioethyl chain on the cysteine N-terminus, we adapted a protocol described by Liu and coworkers[82a] for the modification of primary amines on solid support.

After standard deprotection of the Fmoc group at the terminal cysteine, we transferred the resin from the microwave synthesizer to a parallel synthesizer. As shown in Scheme 12, we first activated the terminal amino group by converting it into a sulfonamide with 2-nitrobenzenesulfonylchloride and 2,6-lutidine in DCM. We then performed two successive Mitsunobu reactions [84] with 2-(tritylthio)ethanol, triphenylphosphine and diisopropyl azodicarboxylate (DIAD) in a dry THF/DCM mixture under argon atmosphere. Finally, we deprotected the amino group by treating the resin with a 2,2'-(ethyldioxy)diethanethiol and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) solution in DMF. 2-(trithio)ethanol[85] **1**, was easily synthesized in one step from 2-mercaptoethanol and tritylchloride in THF.

84. Mitsunobu, O.; Yamada, M. Preparation of Esters of Carboxylic and Phosphoric Acid via Quaternary Phosphonium Salts. *Bull. Chem. Soc. Jpn.* **1967**, *40*, 2380–2382.
85. Yeo, W.-S.; Min, D.-H.; Hsieh, R. W.; Greene, G. L.; Mrksich, M. Label-Free Detection of Protein-Protein Interactions on Biochips. *Angew. Chem. Int. Ed.* **2005**, *44*, 5480–5483.



Scheme 12 | Schematic representation of the three-step procedure for the incorporation of the 2-tritylthioethyl chain on the cysteine residue.

c. Glycine coupling: Acylation of N-alkyl cysteine

Due to steric hindrance, the acylation of N-alkylated amino acids generally proceeds in low yields under standard conditions. In the particular case of the synthesis on solid support, the reactivity of resin-bound secondary amines towards activated amino acids is even lower than in solution[86].

We initially performed the coupling reaction using a 8-fold molar excess of Fmoc-Gly-OH, N,N'-diisopropylcarbodiimide (DIC) and ethylcyano(hydroxyimino)acetate (Oxyma)[87]. The reaction was placed in an orbital shaker and kept under argon for three days to achieve the complete coupling of the glycine. In parallel, in the attempt of improving the coupling efficiency, we investigated a second approach. Reasoning that the coupling was hampered by both the N-(2-trityl-thioethyl) chain and the protecting group on the cysteine residue, we considered performing removal of the bulky trityl group and simultaneous disulphide bond formation with iodine in DMF[88]. Although rapid and efficient, the strong coloration associated with iodine was very persistent. Even after several washings with different solvents, including a

86. Angell, Y. M.; García-Echeverría, C.; Rich, D. H. Comparative Studies of the Coupling of N-Methylated, Sterically Hindered Amino Acids during Solid-Phase Peptide Synthesis. *Tetrahedron Lett.* **1994**, *35*, 5981–5984.
87. Subirós-Funosas, R.; Prohens, R.; Barbas, R.; El-Faham, A.; Albericio, F. Oxyma: An Efficient Additive for Peptide Synthesis to Replace the Benzotriazole-Based HOBt and HOAt with a Lower Risk of Explosion. *Chemistry* **2009**, *15*, 9394–9403.
88. Kamber, B.; Rittel, W. Eine Neue, Einfache Methode Zur Synthese von Cystinpeptiden. *Helv. Chim. Acta* **1968**, *51*, 2061–2069.

buffered ascorbic acid[89] solution in DMF, residual iodine was still present in the resin. However, we performed the following coupling of Fmoc-Gly-OH to test this approach. We tested different conditions (i.e. different reagents, room temperature or microwave heating) and obtained complete coupling when using microwave irradiation. However, UPLC-ESI analysis of the crude peptides showed a variety of impurities, probably due to side-reactions associated with iodine oxidation or to the persistency of iodine in the resin. Overall, this second strategy was unsuccessful in improving the glycine coupling and was therefore abandoned.

d. Simultaneous cleavage from the resin and side-chain deprotection

For the final cleavage of the peptide from the resin and the deprotection of the side chain protecting groups, aqueous concentrated TFA (95 %v/v) is generally the reagent of choice. Nevertheless, during this reaction, highly reactive carbocations are generated and it is necessary to trap them to avoid undesired reactions with sensitive amino acids. Silanes are routinely used as carbocation scavangers and, in the case of sequences containing Cys, Met or Arg, a sulfur containing scavenger is also generally recommended.

In our protocol the resin was suspended in a freshly prepared 10 mL solution of TFA/TIS/2-2'(ethyldioxy)diethanethiol/H_2O: 92.5/2.5/2.5/2.5 vol%. After 1 h, the resin was filtered and washed with TFA (5 mL). The combined filtrates were concentrated in vacuo and then added to cold diethylether (200 mL). The resulting precipitate was then separated by centrifugation, washed with cold diethylether (2 x 40 mL) and dried under argon to afford the crude peptides.

During the synthesis, a mini-cleavage was performed at regular intervals to verify that the reactions had gone to completion. A minimum amount of the resin was cleaved with a 1 mL solution of TFA/TIS/2-2'(ethyldioxy)diethanethiol/H_2O: 92.5/2.5/2.5/2.5 vol%. After 30 min the solvent was then reduced under vacuo and the crude peptide precipitated in diethyl ether. The precipitate was then isolated by centrifugation, dissolved in acidic water (0.1 %v/v formic acid) and/or methanol and analyzed via UPLC-ESI.

89. Pohls, M.; Ambrosius, D.; Grtözinger, J.; Kretzschmar, T.; Saunders, D.; Wollmer, A.; Brandenburg, D.; Bitter-Suermann, D.; Höcker, H. Cyclic Disulfide Analogues of the Complement Component C3a Synthesis and Conformational Investigations. *Int. J. Pept. Protein Res.* **2009**, *41*, 362–375.

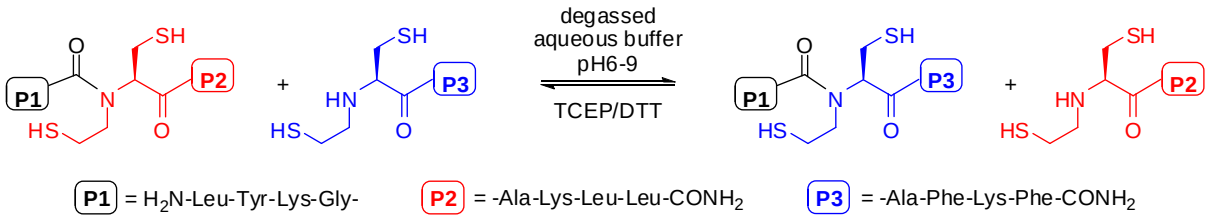
e. Peptide oxidation

Bearing two vicinal thiol groups, these peptides have a high tendency to oxidize. Since purification of these peptides in their reduced form might have been problematic, we added an oxidation step. The crude peptides were thus dissolved in milliQ water, the pH adjusted to 7-8 with NH_4CO_3 and compressed air[90] was bubbled in this solution prior to purification.

3. Exchange reaction between P1-Daa1-P2 and Daa1-P3

a. Exchange reaction setup

The exchange reaction described in Scheme 13 between **P1-Daa1-P2** and **Daa1-P3** was investigated at different pHs (6, 7, 8 and 9).



Scheme 13 | Schematic representation of a reversible NCL based on the use of **Daa1** N-(2-thioethyl)-cysteine.

In order for the reaction to occur, **P1-Daa1-P2ox** and **Daa1-P3ox** need to be reduced to the corresponding **P1-Daa1-P2** and **Daa1-P3**. Among the possible reducing agents commercially available, we focused our attention on dithiothreitol (DTT)[91]. This additive is routinely used for protein assays[92] to avoid disulfide formation from thiols and it has also been used directly in cells[93]. It thus seemed the ideal reagent in our investigation of a reversible chemistry occurring in bio-compatible conditions. One of the main drawbacks related to the use of DTT is its quick oxidation in solution. Once the quantity of reduced DTT in solution is low, the peptides are

90. Kellenberger, C.; Hietter, H.; Luu, B. Regioselective Formation of the Three Disulfide Bonds of a 35-Residue Insect Peptide. *Pept. Res.* **1995**, *8*, 321–327.
91. Cleland, W. W. Dithiothreitol, a New Protective Reagent for SH Groups. *Biochemistry* **1964**, *3*, 480–482.
92. Jocelyn, P. C. *Sulfur and Sulfur Amino Acids*; Methods in Enzymology; Elsevier, 1987; Vol. 143.
93. Tatu, U.; Braakman, I.; Helenius, A. Membrane Glycoprotein Folding, Oligomerization and Intracellular Transport: Effects of Dithiothreitol in Living Cells. *EMBO J.* **1993**, *12*, 2151–2157.

reconverted in their oxidized forms **P1-Daa1-P2ox** and **Daa1-P3ox**, thus preventing the exchange reaction to proceed. Performing the reactions in degassed buffers and under argon atmosphere was not sufficient to avoid the rapid oxidation of DTT and the consequent formation of oxidized peptides. For this reason we decided to add TCEP[94]. This powerful reducing agent is less sensitive to air oxidation than DTT and its addition to the reaction mixture allows the exchange reaction to equilibrate before the peptide oxidation takes place. Although being frequently used in biochemistry, TCEP is known to cause desulfurization[95] as side reaction, i.e. of cysteine into alanine. For this reason we added to the exchange reaction just the stechiometric amount of TCEP necessary to reduce the intramolecular disulfide bonds present in the two starting peptides.

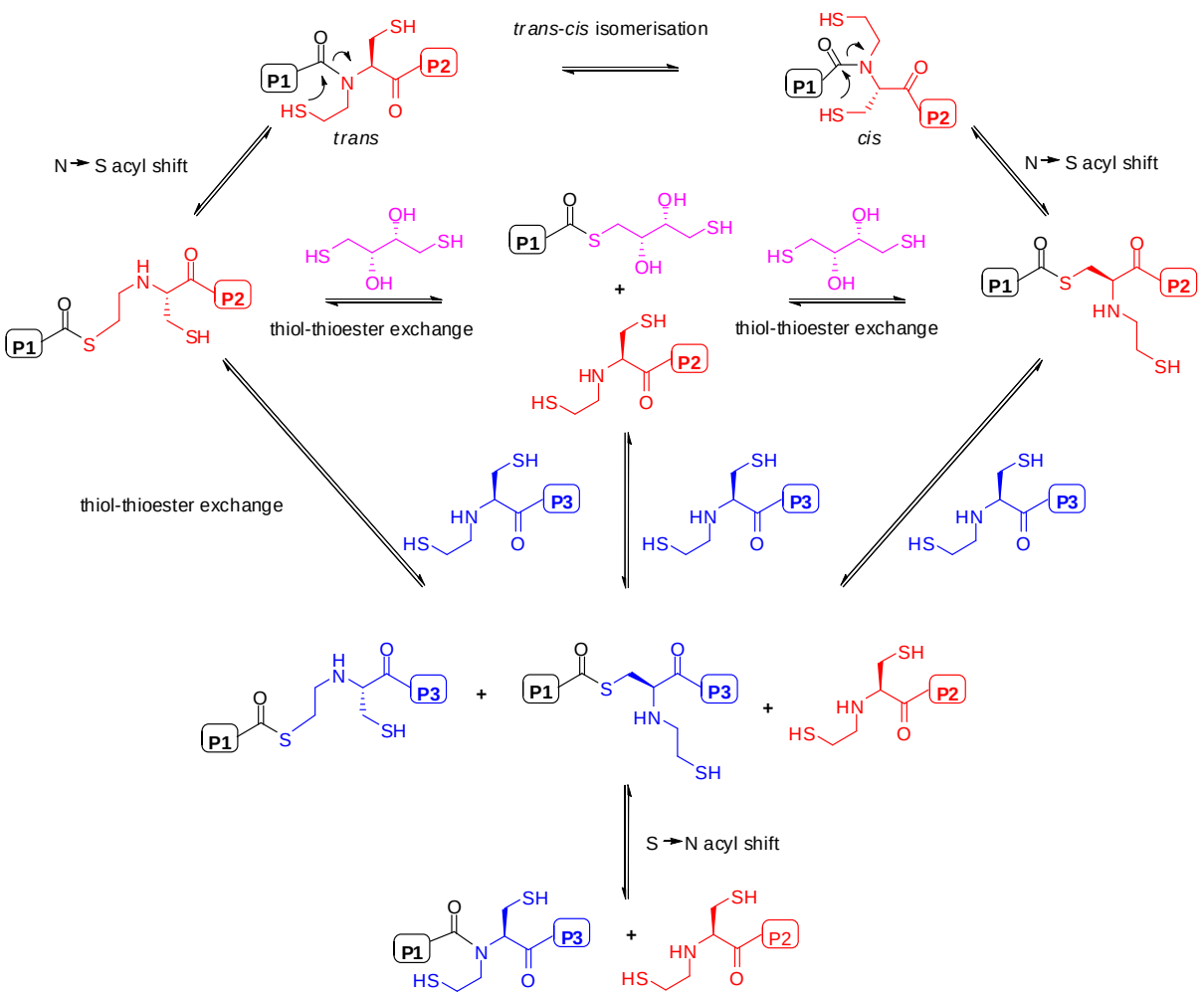
The exchange reactions were run in 0.2 M phosphate buffers prepared at pH 6, 7, 8 and 9. In order to prevent oxidation, these buffers were degassed using freeze/thaw cycles and placed under argon prior to use. For each pH, three solutions of **Daa1-P3** (formate salt), **P1-Daa1-P2** (diformate salt) and TCEP were mixed in an HPLC vial and immediately frozen and lyophilized. This lyophilized powder containing the peptides (final concentration 2.5 mM) and TCEP (final concentration 5 mM) was then diluted under argon with 500 µL of phosphate buffer (pH 6, 7, 8 or 9) containing DTT (25 mM). The exchange reactions were then left stirring under positive argon pressure at room temperature in an HPLC vial inserted in a Schlenk tube (setup shown in Figure 12).

Figure 12 | Picture of the setup: the HPLC vial containing the stirring bar and the exchange reaction solution is inserted in a Schlenk tube.

94. 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.
95. Wang, Z.; Rejtar, T.; Zhou, Z. S.; Karger, B. L. Desulfurization of Cysteine-Containing Peptides Resulting from Sample Preparation for Protein Characterization by Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 267–275.

b. Exchange reaction mechanism

For each conformation of peptide **P1-Daa1-P2** (*trans* or *cis*), one thiol group is always correctly positioned for the intramolecular nucleophilic attack to occur. The two possible resulting **P1-Daa1-P2** thioesters can then undergo transthioesterification to afford the two possible **P1-Daa1-P3** thioesters and **Daa1-P2**. This reaction can occur directly with **Daa1-P3** or through the formation of the **P1-DTT** intermediate that can in turn undergo thiol-thioester exchange with **Daa1-P3**. Finally, the two **P1-Daa1-P3** thioesters can rearrange, via S→N acyl shift, to afford peptide **P1-Daa1-P3**. All the reactions shown in Scheme 14 are fully reversible.



Scheme 14 | Schematic representation of the *trans-cis* isomerisation, N→S acyl shift, thiol-thioester exchange and S → N acyl shift equilibriums involved in a reversible NCL based on the use of **Daa1**.

c. Exchange reaction analysis

Quantitative analysis of the exchange reaction kinetics at pHs ranging from 6 to 9 were determined by UPLC-ESI analysis. At regular time points, aliquots (25 μL) of the exchange reaction mixture were diluted with a solution (175 μL) containing H_2O_2 and the internal UV standard 3,5-dimethoxybenzoic acid; 2 μL of the resulting solution were then submitted to UPLC-ESI analysis and a second 2 μL injection was performed. The UV standard was added to account for potential errors during the sample preparation and injection process. Different molecules were tested as potential UV standards. The choice fell on 3,5-dimethoxybenzoic acid as, during UPLC-ESI analysis, it elutes after the different species involved in the exchange reaction. In the resulting chromatograms the peak corresponding to this molecule did not overlap to other peaks and its area value could thus be used as an internal reference. The reaction mixture was completely oxidized in order to avoid complications in the analysis. Indeed, these peptides have a high tendency to form intramolecular disulfide bonds and direct analysis of the exchange reaction mixture, without an oxidation step, leads to complex chromatograms where peaks corresponding to the reduced and oxidized forms of the peptides overlap (see page 66).

Typical chromatograms to monitor the exchange reactions are shown in Figure 13. At t = 0 h only the two starting peptides **P1-Daa1-P2ox** and **Daa1-P3ox** are present in solution (together with oxidized DTT and the UV standard). Although we could observe the oxidized forms of the two exchange products **P1-Daa1-P3** and **Daa1-P2**, together with **P1-DTTox**, after only few hours, Figure 13 shows the peptide distribution at equilibrium, after 46 h of equilibration.

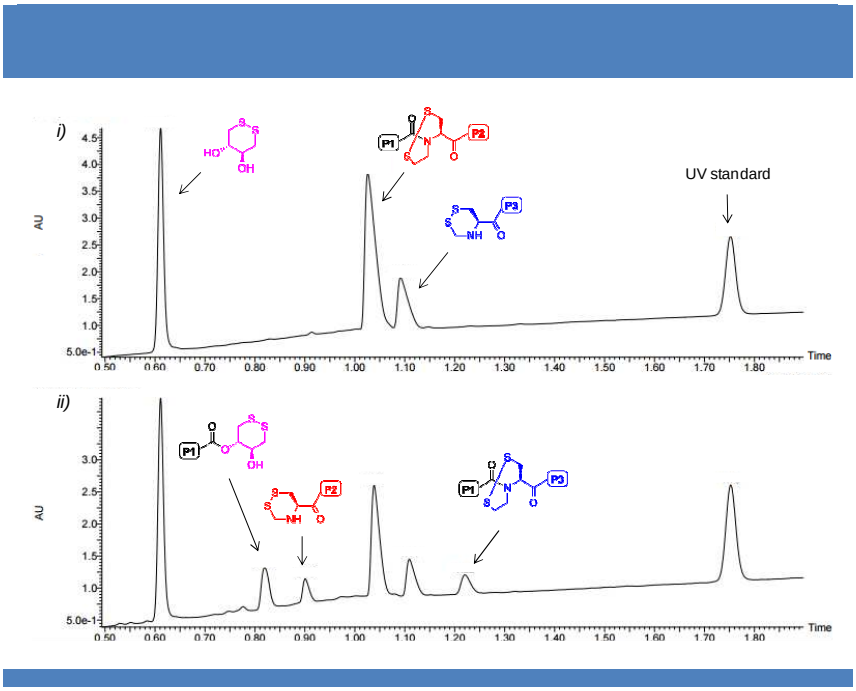


Figure 13 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the exchange reaction between **P1-Daa1-P2** and **Daa1-P3** at pH 6 at i) t = 0 h and ii) t = 46 h.

Calibration curves (Figure 14) were established for peptides **P1-Daa1-P2** and **P1-Daa1-P3** by injecting 2 µL of the peptides at known concentrations (62.50, 125.00, 156.25, 312.50 and 625.00 µM) under the same conditions used for diluting the aliquots of the exchange reactions. Each standard point corresponds to an average of 3 replicates. After a manual baseline correction with the peak analyzer tool of OriginPro 9.0, the area of the peak corresponding to the peptide was corrected according to the area of the internal UV standard and plotted against the concentration. A linear regression analysis gave 2 linear calibration curves with R^2 of 0.992 and 0.996 that where then used to calculate the concentration of the peptides in the exchange mixtures.

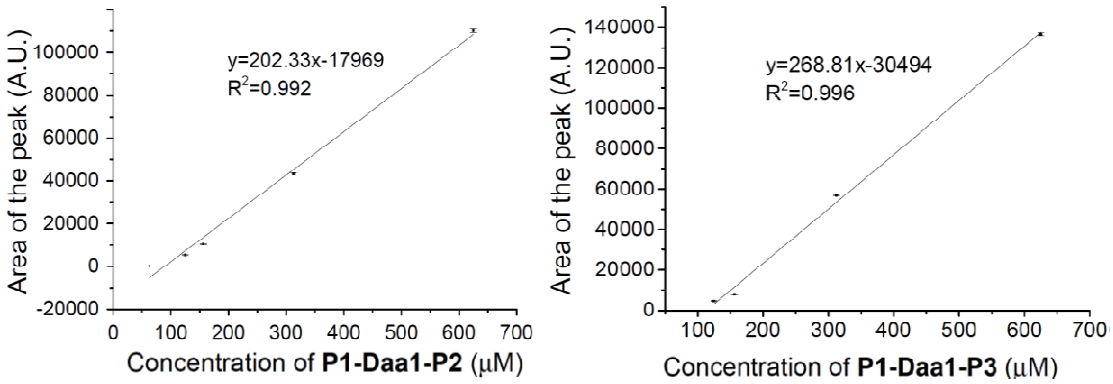


Figure 14 | Calibration curves obtained by linear regression for **P1-Daa1-P2** and **P1-Daa1-P3**.

For each exchange reaction, each timepoint was determined as a mean of two separate injections in the UPLC. In order to calculate the kinetics parameters of the exchange reaction at each of the four pHs investigated, the UPLC chromatograms recorded with UV detection were processed as follows. After a manual baseline correction with the peak analyzer tool of OriginPro 9.0, the area values of the peaks corresponding to **P1-Daa1-P2** and **P1-Daa1-P3** were extracted and corrected according to the area value of the UV standard for each injection. For each timepoint, the mean between the two area values obtained was then expressed as concentration [**P1-Daa1-P2**] and [**P1-Daa1-P3**] of the two peptides, according to the calibration curves reported above. Finally, we expressed the % of the exchange product [**P1-Daa1-P3**] as [**P1-Daa1-P3**]/([**P1-Daa1-P2**]+[**P1-Daa1-P3**]) (Figure 15) and calculated fitting curves for each set of values (i.e. each pH).

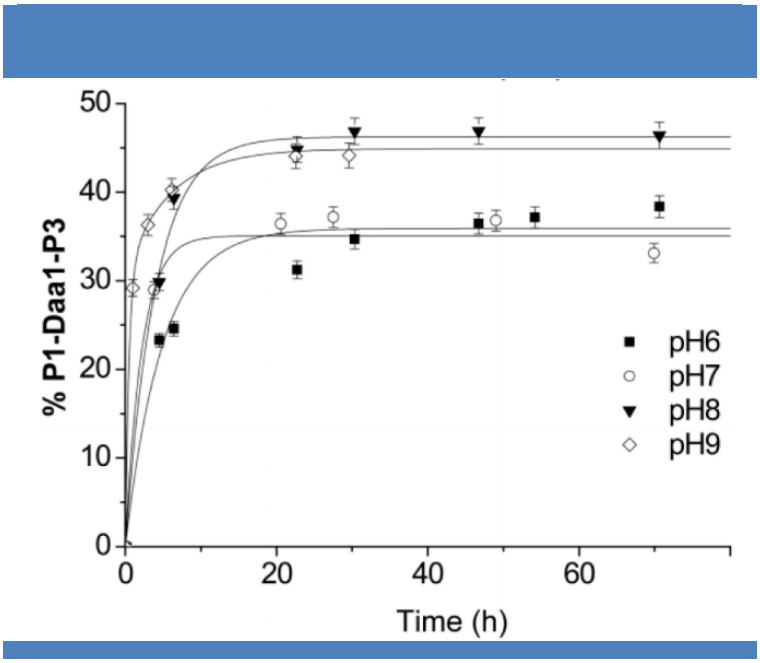


Figure 15 | % of exchange product [**P1-Daa1-P3**] (=[**P1-Daa1-P3**]/([**P1-Daa1-P2**]+[**P1-Daa1-P3**])) as a function of time starting from 2.5 mM solutions of **P1-Daa1-P2** and **P1-Daa1-P3** at room temperature at pHs 6, 7, 8 and 9 in a phosphate buffer (200 mM, DTT 25 mM).

At all the studied pHs we could indeed observe the rapid formation of the exchange products. From the graph reported in Figure 15 we could extrapolate the half-times of equilibration ($t_{1/2}$) 5.5, 2.0, 4.0 and 1.5 h and initial rates (V_0) 6.94×10^{-8}, 8.33×10^{-8}, 1.11×10^{-7} and 1.53×10^{-7} $M \cdot s^{-1}$ respectively at pH 6, 7, 8 and 9. The V_0 values were calculated using the "tangent" tool of the Origin 9.0 software at t = 0 h.

This rate dependency on pH can be explained according to the literature on both the reactions required to yield the exchange products **Daa1-P2** and **P1-Daa1-P3**. Indeed, the

rearrangement of amide-based bis(2-sulfanylethyl)amino units into their thioester form is favored by lowering the pH[82] suggesting that, in addition to the increase of the *cis* amide bond population, inhibition of the S→N acyl shift by protonation of the amide could shift the equilibrium toward thioester formation. In the particular case of our dynamic unit, the formation of two distinct thioesters isomers of **P1-Daa1-P2** was confirmed at low pH (see page 68). Conversely, the second determining step of the dynamic NCL pathway, i.e. the thiol-thioester exchange, is accelerated by a mild increase in the pH (i.e. 8-9). This can be simply explained with the partial deprotonation of the **Daa1** thiols at higher pHs; the increased nucleophilicity of the resulting thiolates gives higher reaction rates. Interestingly, the fastest initial rate at pH 9 of this dynamic pathway, suggests the thiol-thioester exchange as the rate limiting step of the overall exchange reaction.

Moreover, we could observe that the equilibrium constant (K) was strongly affected by the medium, leading to differential expressions of the products at thermodynamic equilibrium for each set of pHs. This behavior of dynamic covalent libraries involving pH-dependent mechanisms had already been described in the literature[55].

When performing the reaction at lower pHs, we could observe the presence of **P1-DTT** together with the exchange peptides. As shown in Figure 16, **P1-DTT** results from the thiol-thioester exchange between one of the two possible thioester isomers of **P1-Daa1-P2** and DTT. Being in equilibrium with all the other components of the mixture and constantly regenerated during the exchange process, **P1-DTT** is not detrimental to the dynamics of the system. Interestingly, at pH 8, the proportion of **P1-DTT** at equilibrium is only residual while at pH 9 it is not even detectable.

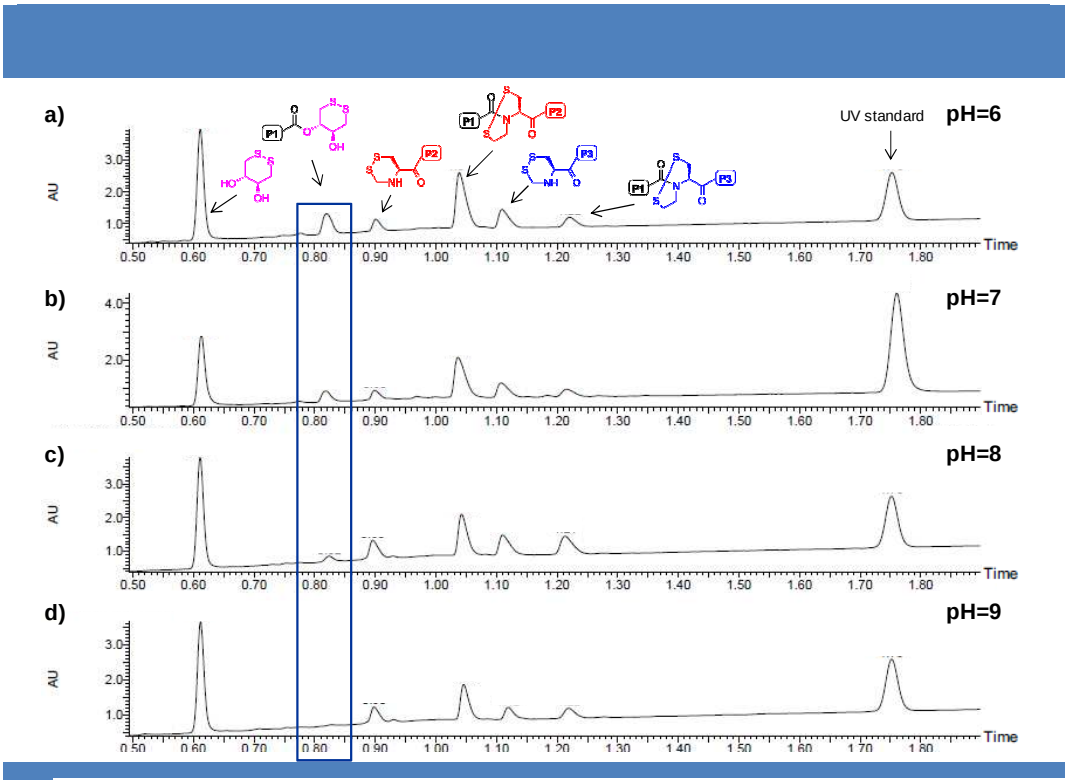


Figure 16 | Reversed phase chromatograms (UPLC-ESI) recorded with UV and MS detection of exchange reactions between **P1-Daa1-P2** and **Daa1-P3**, leading to **P1-Daa1-P3** and **Daa1-P2**, and showing the **P1-DTT** thioester intermediates (peaks in the rectangle) as a function of the pH (from 6 to 9). All chromatograms are given at thermodynamic equilibrium.

One could thus speculate that the optimal pH for this reaction to occur is 8 or 9. However, at these pHs and after long time of equilibration, residual hydrolysis of the thioester intermediates was observed, leading to the formation of small amounts of the C-terminal free carboxylic peptide Leu-Tyr-Lys-Gly (see page 70). Thus, in order to keep the fastest exchange rate and to avoid long-term hydrolysis, one can consider that the optimal pH value for this dynamic process is close to 7.

d. Exchange reaction in detail

i. Complete oxidation of the reaction mixture prior to analysis

Figure 17 shows the effect of diluting an aliquot of the exchange reaction with the solution containing H_2O_2 and the internal UV standard, in comparison to diluting an aliquot of the same exchange reaction with only H_2O. While in the first case the seven peaks are well separated, in the second case overlapping of the many oxidized and reduced species present in solution complicates the analysis of the exchange mixture.

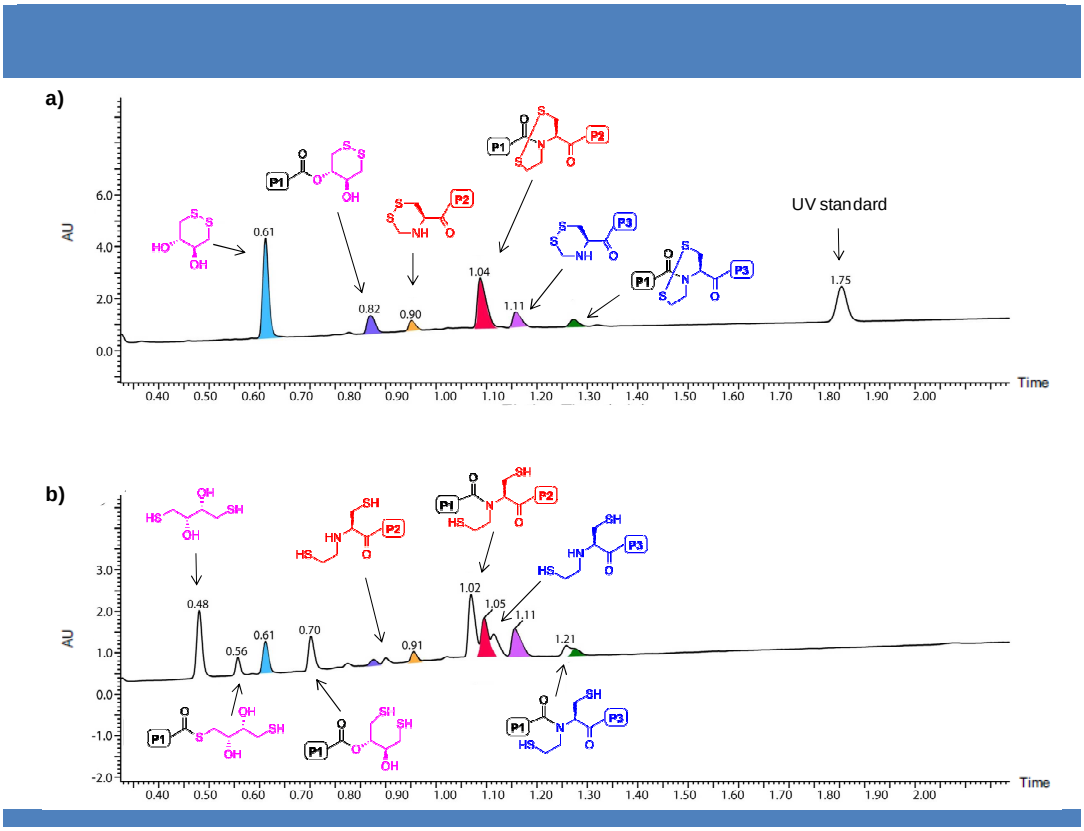


Figure 17 | a) Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of an aliquot (2 µL) of the exchange reaction (25 µL) at pH 6 and after 23 h of equilibration diluted with a solution (175 µL) containing H_2O_2 and the internal UV standard; attribution of each peak, corresponding to the oxidized species. **b)** Reverse phase chromatograms of an aliquot (2 µL) of the same exchange reaction (25 µL) diluted with H_2O (175 µL); attribution of the peaks corresponding to the reduced species, partly overlapping the peaks corresponding to the oxidized species and shown in **a)**.

Peaks were assigned using the corresponding mass spectrum data. Nevertheless, the precise attribution of three of these peaks required further experiments. In Figure 17a the peak at 0.82 min could be attributed to either the $[M+H]^+$ ion of **P1-DTTox** ester or to the $[M+2H]^{2+}$ ion of the **P1-DTTox** thioester dimer. Similarly, in Figure 17b the peaks at 0.56 and 0.70 min could be assigned to the $[M+H]^+$ ion of the two possible **P1-DTT** isomers (thioester or ester). In order to discriminate between thioesters and esters, we performed some experiments in the presence of hydroxylamine. This was based on the premise[96] that thioester bonds are more susceptible to cleavage by weak nucleophiles than oxygen esters. A 2 mL solution of **P1-Daa1-P2** (5 mM) in a pH 7 buffer containing DTT (250 mM) was let equilibrate overnight to allow the quick formation of the possible **P1-DTT** esters and/or thioesters. Aliquots from half of this solution were then analyzed by UPLC-ESI as such and after a 2 min treatment with an equivalent volume of a 2 M solution of hydroxylamine buffered at pH 5 (Figures 18a and 18b). The signal remaining after treatment with hydroxylamine was thus attributed to the ester

96. Xu, M. Q.; Perler, F. B. The Mechanism of Protein Splicing and Its Modulation by Mutation. *EMBO J.* **1996**, *15*, 5146–5153.

form of **P1-DTT**. The other half was vigorously stirred in air for 3 h to allow complete oxidation and then analyzed by UPLC-ESI as such and after a 2 min treatment with an equivalent volume of a 2 M solution of hydroxylamine buffered at pH 5 (Figures 18c and 18d). The resistance of **P1-DTTox** to cleavage allowed the clear attribution of the peak to the **P1-DTTox** ester with an intramolecular disulfide bond.

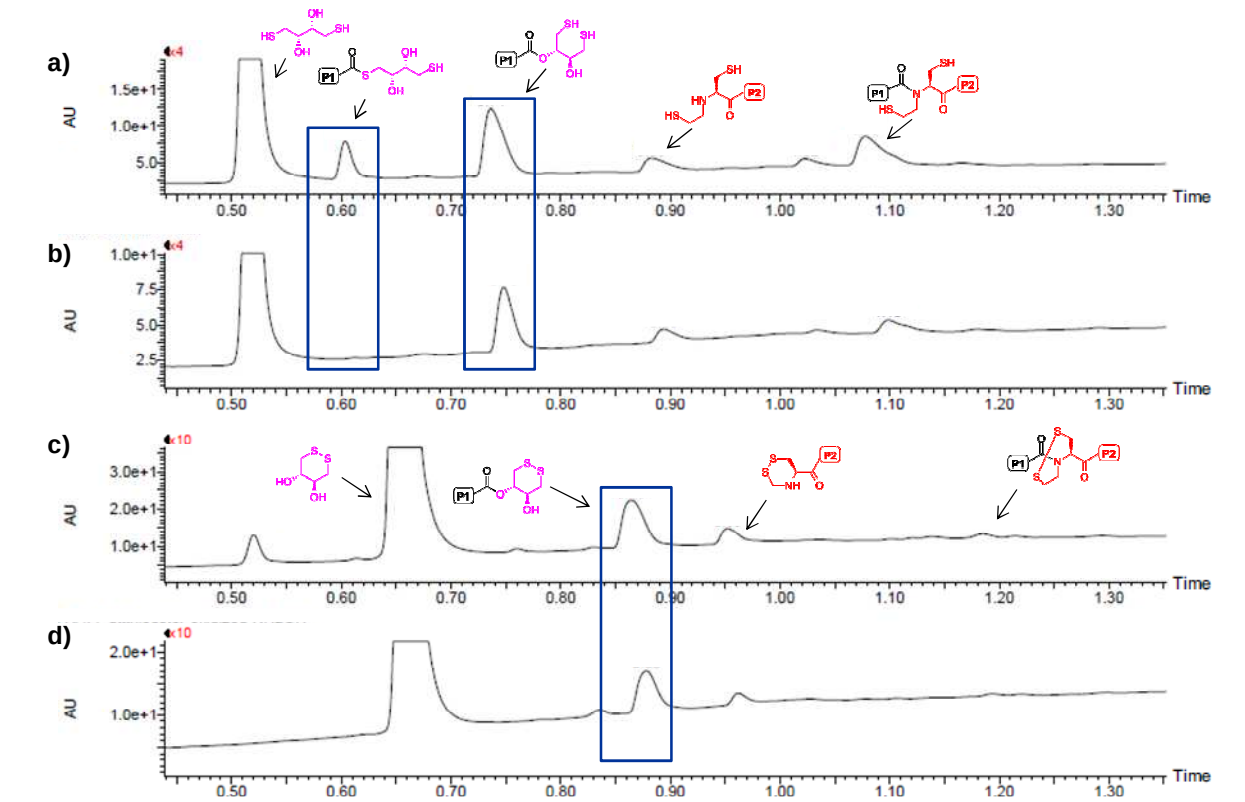


Figure 18 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection after overnight equilibration of a 5 mM solution of **P1-Daa1-P2** (5 mM) in a pH 7 buffer containing DTT (250 mM) **a)** before and **b)** after treatment (50:50 vol%) with a 2 M solution of hydroxylamine buffered at pH 5. Reverse phase chromatograms obtained as described above but after complete oxidation of the mixture, **c)** before and **d)** after treatment with hydroxylamine.

ii. Characterization of the thioester intermediates

Previous work on peptides incorporating the bis(2-sulfanylethyl)amino moiety showed that acidic conditions favored the thioester isomer, whereas the amide configuration was predominant at neutral pH. We were thus confident that at the pHs investigated, the last step of the exchange mechanism was a rapid intramolecular S→N acyl shift leading to the amide isomers of the exchange peptides. In addition, the four (starting and exchange) peptides have been synthesized and characterized. They have thus been used as references to establish retention times for UPLC-ESI experiments. Moreover, NMR analysis of these four compounds

was consistent with an amide form due to the absence in the ^{13}C NMR spectra of any peak in the region corresponding to the carbonyl of aliphatic thioesters (around 200 ppm).

A first experiment was run on peptide **P1-Daa1-P2** in order to promote the formation of the corresponding thioesters (Figure 19). A solution of the peptide in a pH 6 phosphate buffer containing TCEP (28 mM) was acidified with 2.5 %v/v TFA and let stir for 20 min. As a result we could observe two new peaks corresponding to the two possible peptide thioesters. Moreover, we observed an unexpected peak, whose molecular weight might correspond to the thiazoline derivatives formed *in situ* by intramolecular nucleophilic addition and dehydration (Figure 19 *ii)*). Neutralization of the solution with $(NH_4)_2CO_3$ led to the reconversion of the different species into the more stable amide peptide **P1-Daa1-P2**.

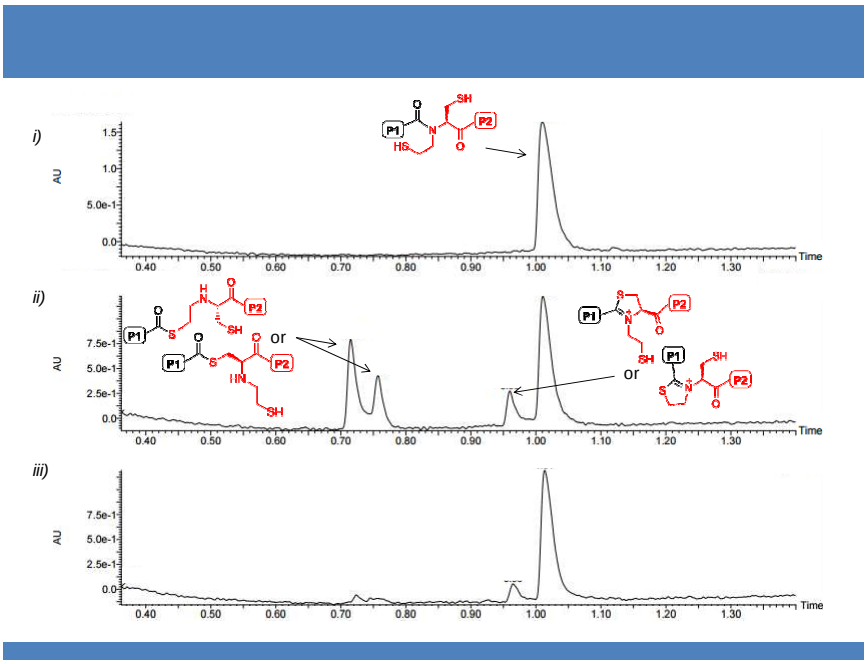


Figure 19 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of **P1-Daa1-P2** (320 µM) *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 20 min in the same solution + 2.5 %v/v of TFA; *iii)* after subsequent neutralization to pH 7 using $(NH_4)_2CO_3$ for 5 min.

A second experiment was run on both peptides **P1-Daa1-P2** and **P1-Daa1-P3** to unambiguously assess the thioester nature of the new peaks formed in acidic conditions. Two separate solutions of peptides **P1-Daa1-P2** and **P1-Daa1-P3** (320 µM) at pH 6 were thus acidified by adding 2.5 %v/v of TFA. The resulting mixtures, containing the amide and thioester isomers of each peptide, together with the thiazoline derivatives, were then treated with an equivalent volume of a 2 M solution of hydroxylamine (buffered at pH 5). As a result (Figure 20), for each of the peptides **P1-Daa1-P2** and **P1-Daa1-P3**, UPLC-ESI analysis showed the disappearance of the peaks corresponding to the two possible thioester isomers and

the formation of the hydroxamic acid of the N-terminal fragment together with the release of the C-terminal fragment (respectively **Daa1-P2** and **Daa1-P3**).

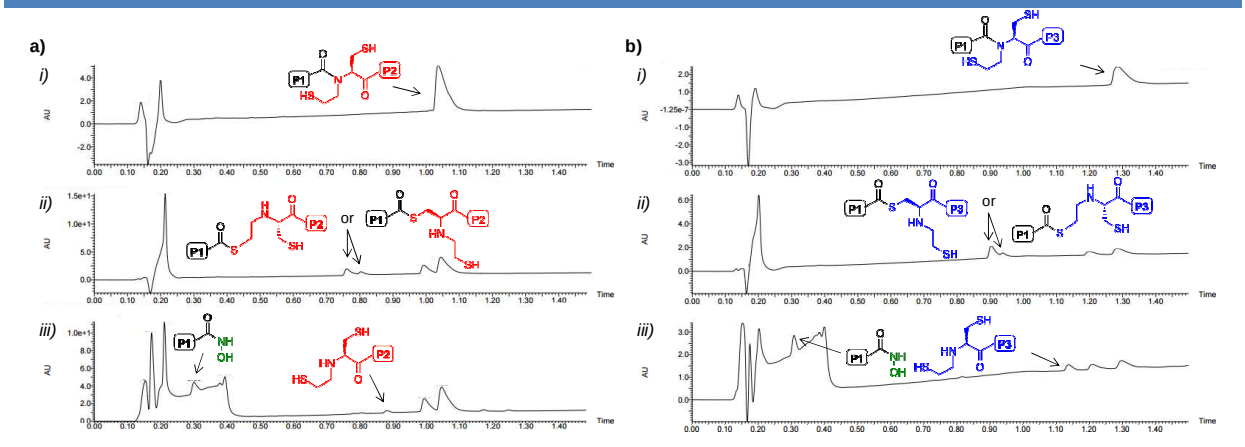


Figure 20 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of **a) P1-Daa1-P2** (320 µM), *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 20 min in the same solution + 2.5 %v/v of TFA; *iii)* after a 10 min treatment with an equivalent volume of a 2 M hydroxylamine solution buffered at pH 5; **b) P1-Daa1-P3** (320 µM), *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 120 min in the same solution + 2.5 %v/v of TFA; *iii)* after a 5 min treatment with an equivalent volume of a 2 M hydroxylamine solution buffered at pH 5. The poor quality of the baseline before 0.45 min is due to the presence of hydroxylamine at high concentrations.

iii. Characterization and quantification of the hydrolysis product

After long time of equilibration, and in particular at pH 8 and 9, the peak corresponding to the hydrolysis product of the **P1-DTT** thioester appeared in the reverse phase chromatograms. In order to show the stability of the exchange product **P1-Daa1-P3**, we have plotted its absolute concentration against time for each of the four pHs investigated. **P1-Daa1-P3** proved to be stable even long time after the thermodynamic equilibrium was reached, allowing us to conclude that the hydrolysis byproducts can be neglected in the timescale of the reaction (Figure 21).

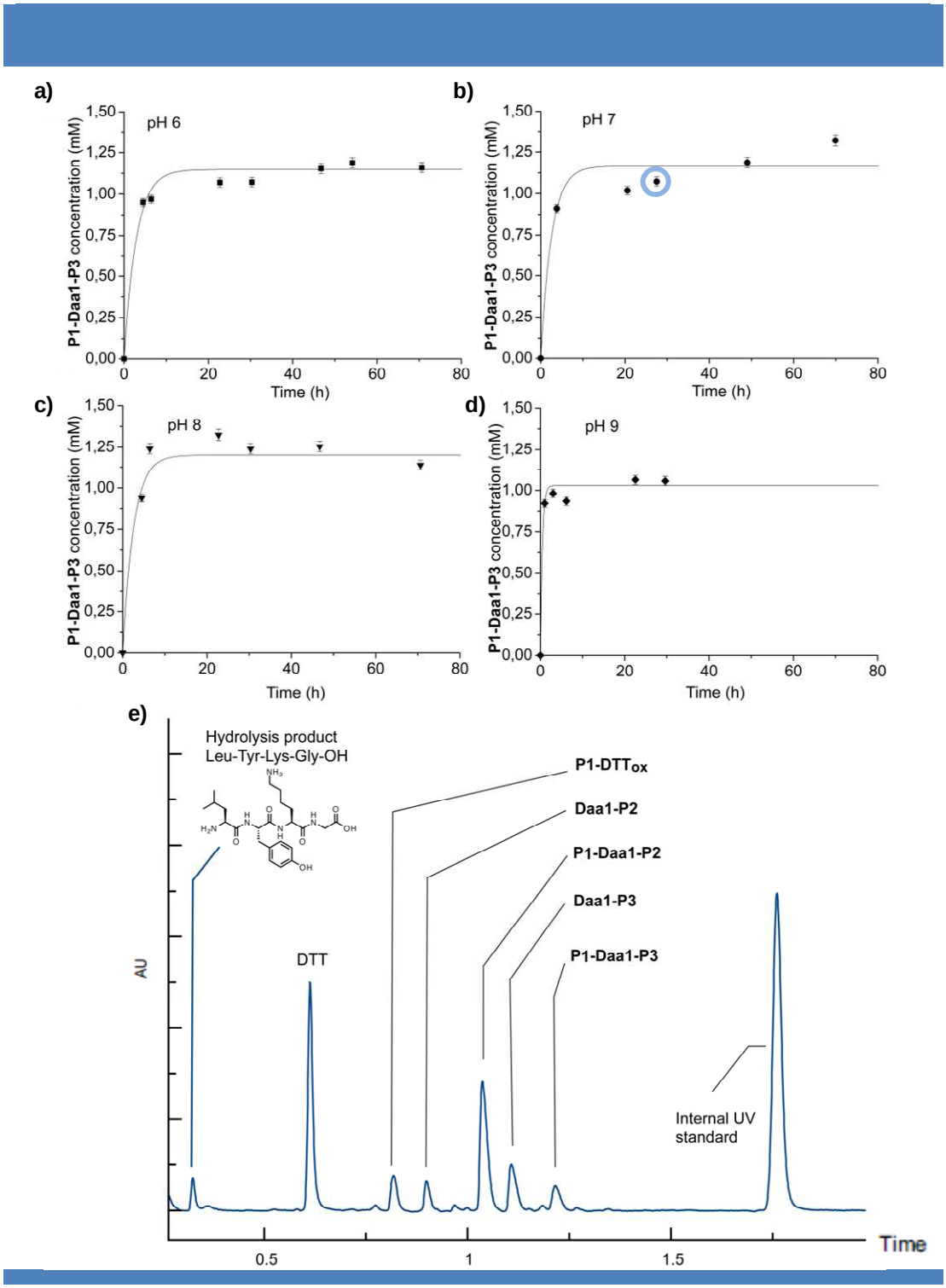


Figure 21 | Absolute concentration of exchange peptide **P1-Daa1-P3** plotted against the time of equilibration at pH 6 **a)**, 7 **b)**, 8 **c)** and 9 **d)**. **e)** Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection at pH 7 and after 28 h of equilibration, corresponding to the timepoint circled in **b)**.

e. Exchange reaction with 1,4-butanedithiol replacing DTT

The experiment in the presence of 1,4-butanedithiol was run to ultimately confirm the exchange reaction mechanism (Figure 22). This molecule is an analogue of DTT, but the lack of the two hydroxyl groups makes it a less ambiguous reducing agent to assess the reaction mechanism. The exchange reaction run at pH 7 showed a similar rate compared to the same reaction run in the presence of DTT; this shows that if both esters and thioesters can be formed from DTT during the course of the reaction, only the thioesters are mandatory to promote the exchange reaction.

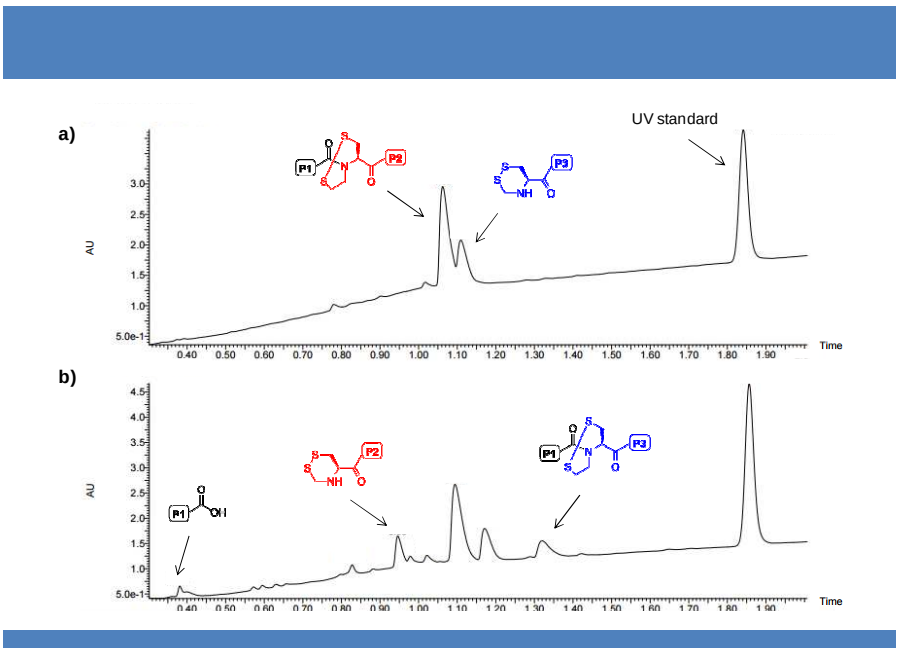


Figure 22 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the exchange reaction between **P1-Daa1-P2** and **Daa1-P3** in the presence of 1,4-butanedithiol at pH 7 at **a)** t = 0 and **b)** 90 h.

4. Exchange reaction between P1-Daa1-P3 and Daa1-P2

In a dynamic combinatorial library the distribution of its constituents is determined by their thermodynamic stability within this system. Provided that all the dynamic blocks are present in solution, the distribution of the constituents at equilibrium is not affected by the nature of the starting species. In our system this means that we expected the peptide distribution at equilibrium to be the same when starting the reaction from **P1-Daa1-P2** and **Daa1-P3**, or from **P1-Daa1-P3** and **Daa1-P2**.

We thus investigated the exchange reaction symmetric to the one hitherto described,

consisting in the dynamic exchange between **P1-Daa1-P3** and **Daa1-P2**. The setup and the analysis of this exchange reaction were performed as previously reported. At pH 7 and after three days of equilibration, the peptide distribution was very similar (Figure 23a *iii)*) to the one observed for the "forward" exchange reaction (Figure 23a *ii)*), starting from **P1-Daa1-P2** and **Daa1-P3**, proving the thermodynamic control of the system.

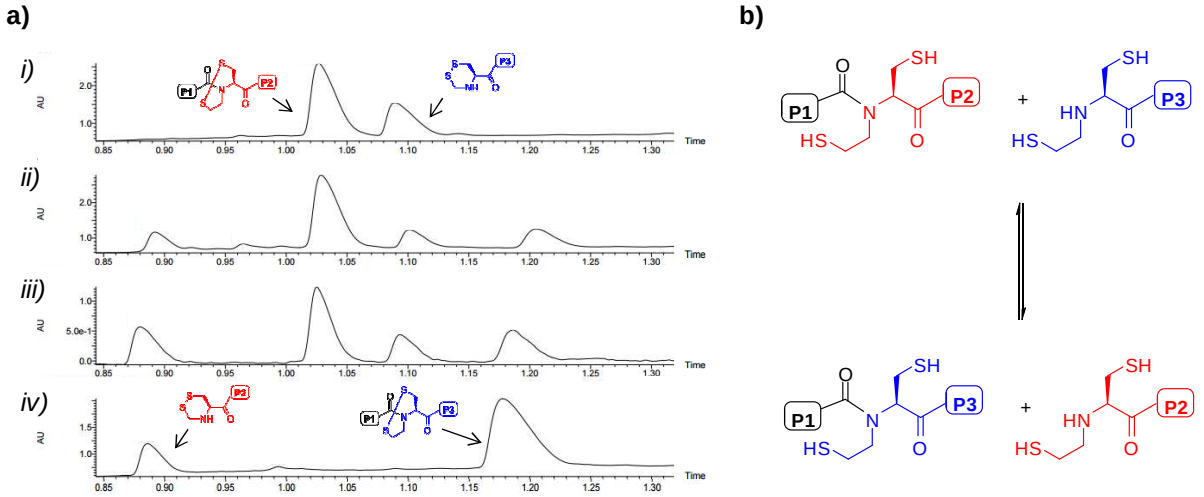


Figure 23 | a) Reverse phase chromatograms (UPLC-ESI) at t = 0 h (*i)* and *iv)*) and t = equilibrium (*ii)* and *iii)*) for the forward (*i)* and *ii)*) and reverse (*iii)* and *iv)*) exchange reactions at pH 7 described in **b)**.

5. Control reaction between P1-Cys-P2 and Cys-P3

In order to provide a control for the exchange reaction, two analogues of the dynamic peptides were synthesized (Figure 24). **Cys-P3** (**Cys**-Ala-Phe-Lys-Phe-$CONH_2$) and **P1-Cys-P2** (H_2N-Leu-Tyr-Lys-Gly-**Cys**-Ala-Lys-Leu-Leu-$CONH_2$) lack the N-(2-thioethyl) group present in the dynamic peptides **Daa1-P3** and **P1-Daa1-P2**.

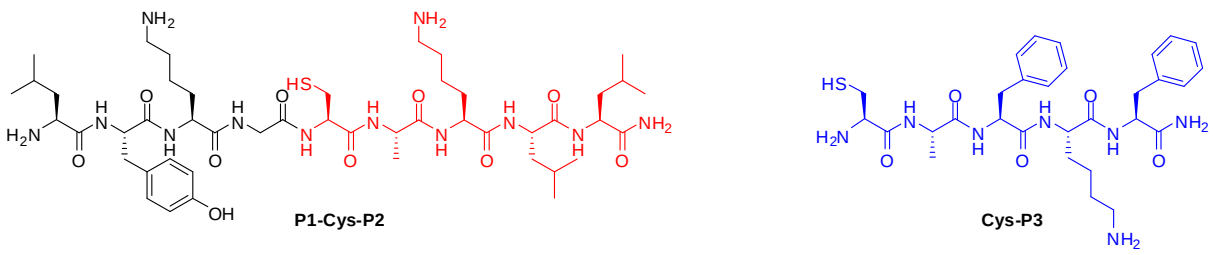


Figure 24 | Structures of the two control peptides **P1-Cys-P2** and **Cys-P3**.

The synthesis of these peptides was entirely performed with standard procedures via microwave-assisted solid phase peptide synthesis, using the standard Fmoc/*tert*-butyl strategy and HBTU/DIEA coupling system previously described (see page 55).

A reaction at pH 7 between peptides **P1-Cys-P2** and **Cys-P3** (Figure 25a) was investigated. Solutions of the two peptides were joined in a HPLC vial, frozen and lyophilized. This lyophilized powder containing the peptides (final concentration 2.5 mM) was then diluted under argon with 500 µL of pH 7 phosphate buffer containing DTT (25 mM). The reaction was then left stirring under positive argon pressure at room temperature in the HPLC vial inserted in a Schlenk tube. At regular time points, aliquots (25 µL) of the exchange reaction mixture were diluted with a solution (175 µL) containing TCEP and the internal UV standard 3,5-dimethoxybenzoic acid; 2 µL of the resulting solution were then submitted to UPLC-ESI analysis. Due to the higher stability of these peptides compared to their **Daa1**-containing counterparts, the conditions used in the setup and in the analysis of the reaction are slightly different. The impossibility of forming intramolecular disulfide bonds confers higher resistance to oxidation to these peptides. This explains the removal of TCEP in the reaction setup and the substitution of the full oxidation for full reduction prior to UPLC-ESI analysis.

As expected, this control reaction did not yield any detectable trace of exchange products after four weeks of equilibration (Figure 25b).

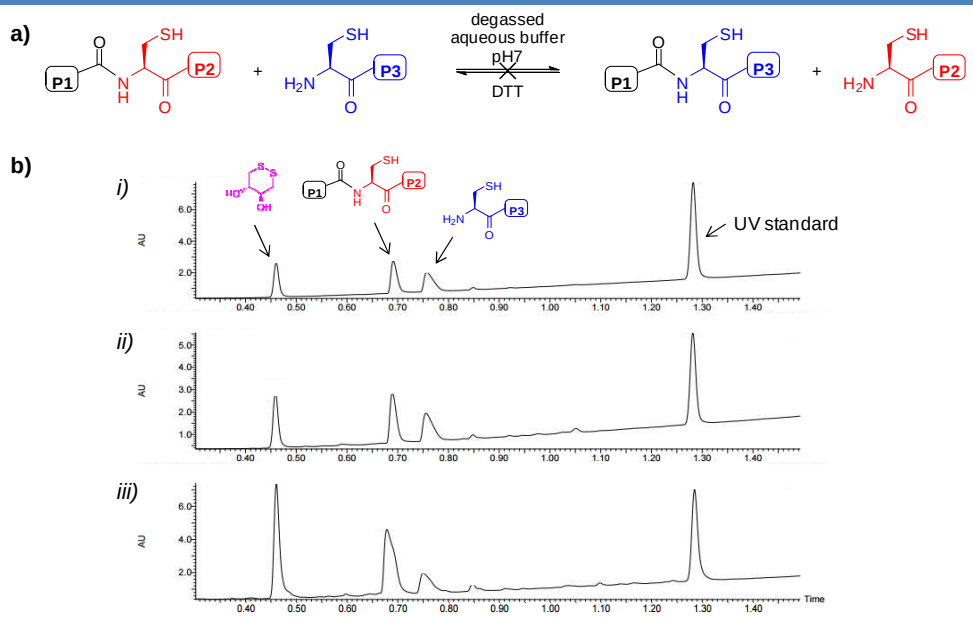


Figure 25 | a) Schematic representation of the control reaction between **P1-Cys-P2** and **Cys-P3**. **b)** Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the reaction between **P1-Cys-P2** and **Cys-P3** at pH 7 at *i)* t = 5 min, *ii)* t = 7 days and *iii)* t = 4 weeks.

Chapter 3. Peptides Incorporating N-Methyl-Cysteine

1. Introduction: N-methylation of peptides and proteins

a. N-methylation of peptides and proteins in nature

N-methylation has been observed on the side chains of amino acids such as lysine and arginine as well as on the peptide bond. However, the two biosynthetic pathways are extremely different.

N-methylation of lysine and arginine residues has been observed[97] in nuclear proteins in eukaryotic cells. The addition of one, two or three methyl groups on each of these amino acids is a common posttranslational modification, finely tuned by specific methyltransferases and demethylases. Among the proteins that can undergo N-methylation, histones are by far the most important ones. DNA wraps around these alkaline proteins forming nucleosomes and this process is strictly correlated to the patterns and extent of histone N-methylation. The abundance of alkaline amino acids on these peptides, and the possibility of each of these residues to be methylated up to three times, creates an enormous set of possible combinations. Since other posttranslational modifications can also occur, Jenuwein and Allis[98] proposed the existence of an “histone code” that is translated into biological functions when “read” by enzymes and other proteins. It has indeed been observed that changes in the methylation pattern are associated with the development of human cancer[99].

Conversely, N-methylation of the peptide bond is not a post translational modification and to date has not been observed in mammalian proteins[97]. It occurs in bacteria and fungi in large multi-enzyme complexes called non-ribosomal peptide synthetases (NRPSs), also responsible for the insertion of other moieties on the peptide backbone. Frequent modifications include the insertion of heterocyclic rings, fatty acid chains, glycosylated amino acids, D-amino acids and other moieties[100].

97. Chatterjee, J.; Rechenmacher, F.; Kessler, H. N-Methylation of Peptides and Proteins: An Important Element for Modulating Biological Functions. *Angew. Chem. Int. Ed.* **2013**, *52*, 254–269.
98. Jenuwein, T.; Allis, C. D. Translating the Histone Code. *Science* **2001**, *293*, 1074–1080.
99. Chi, P.; Allis, C. D.; Wang, G. G. Covalent Histone Modifications--Miswritten, Misinterpreted and Mis-Erased in Human Cancers. *Nat. Rev. Cancer* **2010**, *10*, 457–469.
100. Sieber, S. A.; Marahiel, M. A. Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics. *Chem. Rev.* **2005**, *105*, 715–738.

b. N-methyl peptides as drugs

With an estimated increase to 25.4 billion US dollars in 2018, the global peptide drug market is flourishing[101]. The therapeutic potential of peptides lies in their high specificity towards a broad range of targets, that results in high potency of action and relatively low side-effects. As their degradation products are amino acids, they show fewer side effects than small molecule drugs, whose metabolites are potentially toxic. If compared to protein-based biopharmaceuticals, they generally penetrate better into tissues, are less immunogenic and more stable.

However, naturally occurring peptides are often not suitable for use as therapeutics. Rapid degradation by proteolytic enzymes and high clearance, together with poor membrane permeability, are the main disadvantages of peptide drug candidates[102]. Another major drawback in the development of peptide drugs is their negligible oral availability. In addition to what described above, this is also due to low uptake via tight junctions and active export into the gut[103].

In order to overcome these issues, researchers have developed multiple strategies[104], mainly addressed to improving proteolytic stability and increasing hydrophobicity. Cyclization of the peptide sequence, replacement of natural amino acids with non natural residues and modifications of the amide bond or of the C- or N-terminus are the most common modifications. Moreover, peptide flexibility often results in low receptor subtype selectivity, therefore the insertion of suitable conformational constraints into the molecule has been shown to increase this selectivity.

Among the strategies listed above, N-methylation of the peptide backbone is widely used. Remarkably, the insertion of this small modification of only 14 Daltons has been shown to increase proteolytic stability[105], promote membrane permeability[103], inhibit peptide aggregation[106] and enhance selectivity[107] in biomolecular recognition processes. Recently

101. Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2014**, *20*, 122–128.
102. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. The Future of Peptide-Based Drugs. *Chem. Biol. Drug Des.* **2013**, *81*, 136–147.
103. Kessler, H.; Chatterjee, J.; Doedens, L.; Opperer, F.; Biron, E.; Hoyer, D.; Schmid, H.; Gilon, C.; Hruby, V. J.; Mierke, D. F. New Perspective in Peptide Chemistry by N-Alkylation. In; Valle, S. Del; Escher, E.; Lubell, W. D., Eds.; Advances in Experimental Medicine and Biology; Springer New York: New York, NY, 2009; Vol. 611, pp. 229–231.
104. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M. Synthetic Therapeutic Peptides: Science and Market. *Drug Discov. Today* **2010**, *15*, 40–56.
105. Fiacco, S. V; Roberts, R. W. N-Methyl Scanning Mutagenesis Generates Protease-Resistant G Protein Ligands with Improved Affinity and Selectivity. *ChemBioChem* **2008**, *9*, 2200–2203.
106. a) Gordon, D. J.; Sciarretta, K. L.; Meredith, S. C. Inhibition of B-Amyloid(40) Fibrillogenesis and Disassembly of B-Amyloid(40) Fibrils by Short B-Amyloid Congeners Containing N -Methyl Amino Acids at Alternate Residues. *Biochemistry* **2001**, *40*, 8237–8245. b) Lanning, J. D.; Hawk, A. J.; Derryberry, J.; Meredith, S. C. Chaperone-like N-Methyl Peptide Inhibitors of Polyglutamine Aggregation. *Biochemistry* **2010**, *49*, 7108–7118.
107. Chatterjee, J.; Ovadia, O.; Zahn, G.; Marinelli, L.; Hoffman, A.; Gilon, C.; Kessler, H. Multiple N-Methylation by a Designed Approach

N-methylation has also been correlated to enhanced intestinal permeability[108] and oral bioavailability[109]. All these effects can be explained with the increased hydrophobicity and conformational constraints of N-methylated peptides. It is known that steric hindrance on the amide bond lowers the energy difference between its *cis* and *trans* configurations[110]; in addition, peptide conformation is also affected by steric interactions between the N-methyl group and the amino acid side chains. In the particular case of long peptides, the removal of the hydrogen-bonding capability prevents the formation of secondary structural elements (i.e. peptide aggregation into β–sheets)[97].

The potential of N-methylation in increasing the affinity of a lead peptide for a given target is so striking that a N-methyl scan[111] method has been developed. In this approach, a library of peptides with the N-methyl group placed at different positions is synthesized and tested. The peptides with the most active conformations will thus be selected.

Several naturally occurring cyclic and linear N-methylated peptides have been described and their potential bioactivity investigated. Due to their different structures, they can interact with various subcellular components in mammalian cells, leading to a number of potential medical applications. Many of these natural peptides and their synthetic derivatives are currently on the market or in clinical trials for the treatment of different pathologies. Among them, cyclosporine A[112] is remarkable: this highly N-methylated natural peptide is administered orally despite it violates all the Lipinski's rules for oral availability. It is currently used as an immunosuppressive drug in organ transplantation and its structure is shown in Figure 26.

Enhances Receptor Selectivity. *J. Med. Chem.* **2007**, *50*, 5878–5881.

108. Ovadia, O.; Greenberg, S.; Chatterjee, J.; Laufer, B.; Opperer, F.; Kessler, H.; Gilon, C.; Hoffman, A. The Effect of Multiple N-Methylation on Intestinal Permeability of Cyclic Hexapeptides. *Mol. Pharm.* **2011**, *8*, 479–487.
109. Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H. Improving Oral Bioavailability of Peptides by Multiple N-Methylation: Somatostatin Analogues. *Angew. Chem. Int. Ed.* **2008**, *47*, 2595–2599.
110. Vitoux, B.; Aubry, A.; Cung, M. T.; Marraud, M. N-Methyl Peptides. *Int. J. Pept. Protein Res.* **2009**, *27*, 617–632.
111. Rajeswaran, W. G.; Hocart, S. J.; Murphy, W. A.; Taylor, J. E.; Coy, D. H. N- Methyl Scan of Somatostatin Octapeptide Agonists Produces Interesting Effects on Receptor Subtype Specificity. *J. Med. Chem.* **2001**, *44*, 1416–1421.
112. Laupacis, A.; Keown, P. A.; Ulan, R. A.; McKenzie, N.; Stiller, C. R. Cyclosporin A: A Powerful Immunosuppressant. *Can. Med. Assoc. J.* **1982**, *126*, 1041–1046.

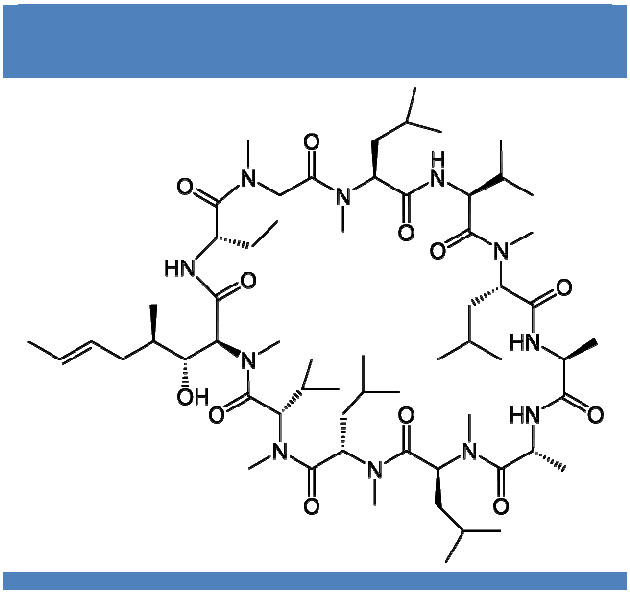


Figure 26 | Structure of cyclosporine A.

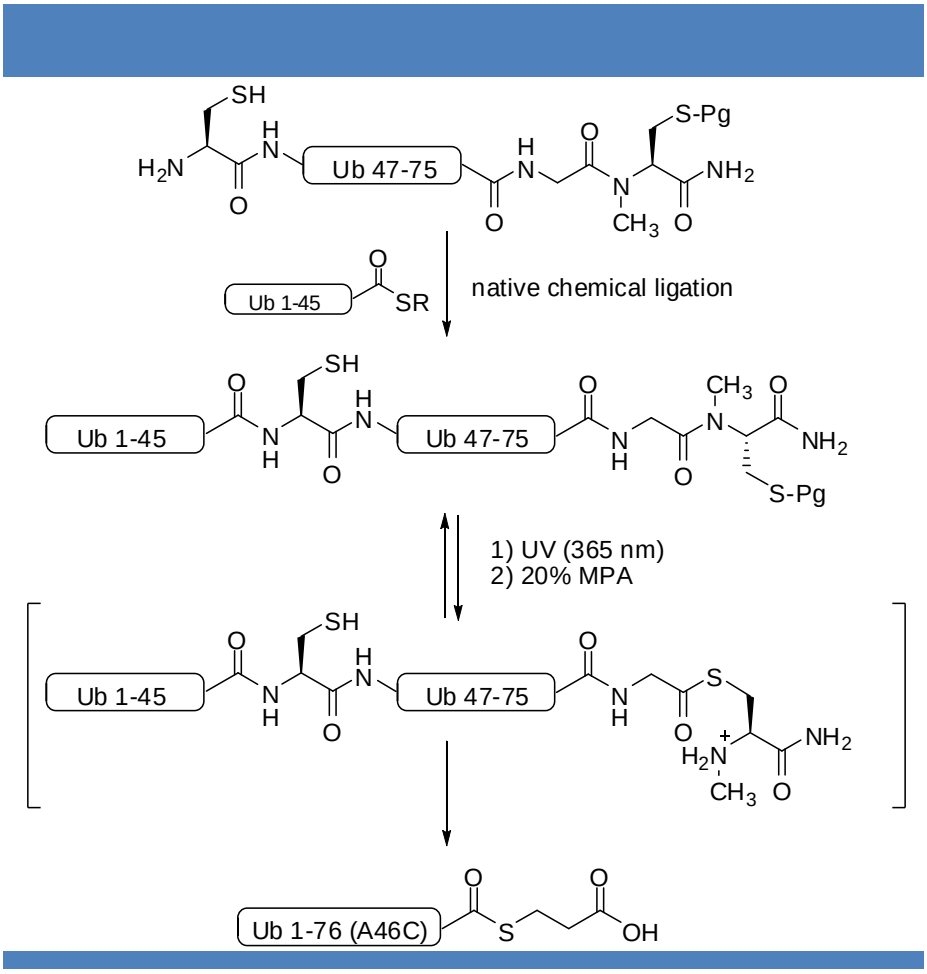
c. Design and objectives

As described in Chapter 2, the insertion of N-(2-thioethyl)-cysteine (**Daa1**) in peptides led to the first example of a dynamic NCL mechanism. The strength of this methodology rests in the formation of a reversible peptide bond in mild aqueous conditions and without relying on the use of enzymes. However, the dynamic unit N-(2-thioethyl)-cysteine is not a natural amino acid. Therefore, the potential biocompatibility and bioactivity of the compounds incorporating this unit remain to be investigated.

With the aim of developing a fully biocompatible reversible chemistry for the exchange of peptide fragments, we decided to further investigate the mechanism of the exchange reaction. As shown in Chapter 1, in acidic conditions we have observed the formation of both the possible thioester isomers of the peptides. This means that both the thiol groups on the dynamic unit can perform the nucleophilic attack leading to the N→S acyl shift. In turn, this is only possible if the peptide bond assumes both the *cis* and the *trans* conformation. Indeed, N^{α}-alkylation is known to increase the population of the *cis* conformation of that specific peptide bond, thereby distorting the peptide backbone. In the case of N-alkyl-cysteine, the distortion of the peptides into their *cis* conformation favors the intramolecular N→S acyl shift, thus the formation of their thioester isomers. For this reason N-alkyl cysteine (N-ethyl cysteine

in this case) has already been inserted[113] in peptides to be used as a thioester surrogate in native chemical ligation. Based on these premises, we focused our investigation on a new dynamic unit, N-methyl-cysteine (or dynamic amino acid 2, **Daa2**), a non ribosomal amino acid that can be found in natural products[114].

Similarly to N-ethyl-cysteine, this residue has also been used as a latent thioester functionality. In particular, N-methyl-cysteine at the C-terminus of ubiquitine has been used by Brik and coworkers[115] to generate the ubiquitine thioester *in situ*. In this strategy the N-methyl-cysteine was inserted as C-terminal residue and its thiol was initially protected with a photolabile group. This allowed the full synthesis of ubiquitine via native chemical ligation, followed by selective deprotection of the photolabile group and thiol-thioester exchange, leading to the synthesis of the final ubiquitine thioester (Scheme 15).



Scheme 15 | Synthesis of an ubiquitine thioester by using N-methyl cysteine as a latent thioester. Pg = 2-nitrobenzyl, R = $-CH_2-CH_2-COOH$, MPA = mercaptopropionic acid.

113 Asahina, Y.; Nabeshima, K.; Hojo, H. Peptidyl N-Alkylcysteine as a Peptide Thioester Surrogate in the Native Chemical Ligation. *Tetrahedron Lett.* **2015**, *56*, 1370–1373.

114 Pérez Baz, J.; Cañedo, L. M.; Fernández Puentes, J. L.; Silva Elipe, M. V. Thiocoraline, a Novel Depsipeptide with Antitumor Activity Produced by a Marine Micromonospora. II. Physico-Chemical Properties and Structure Determination. *J. Antibiot.* **1997**, *50*, 738–741.

115 Erlich, L. A.; Kumar, K. S. A.; Haj-Yahya, M.; Dawson, P. E.; Brik, A. N-Methylcysteine-Mediated Total Chemical Synthesis of Ubiquitin Thioester. *Org. Biomol. Chem.* **2010**, *8*, 2392–2396.

The development of a reversible native chemical ligation based on this new natural dynamic unit **Daa2** allowed the scrambling of peptide fragments to occur in fully biocompatible conditions. We can only speculate about the importance of these findings that could lead to the discovery of new bioactive compounds as well as new biocompatible materials.

2. Peptide synthesis

For the initial experiments, we incorporated N-methyl-cysteine in our model peptides **P1-Daa2-P2** (H_2N-Leu-Tyr-Lys-Gly-**Daa2**-Ala-Lys-Leu-Leu-$CONH_2$) and **Daa2-P3** (**Daa2**-Ala-Phe-Lys-Phe-$CONH_2$). Later, in order to further investigate the potential of the exchange reaction between peptides incorporating the new dynamic unit **Daa2**, we synthesized peptides **P1-Daa2-P3** (H_2N-Leu-Tyr-Lys-Gly-**Daa2**-Ala-Phe-Lys-Phe-$CONH_2$), **P1'-Daa2-P2** (H_2N-Leu-Tyr-Lys-Val-**Daa2**-Ala-Lys-Leu-Leu-$CONH_2$) and **P1''-Daa2-P2** (H_2N-Leu-Tyr-Lys-Lys-**Daa2**-Ala-Lys-Leu-Leu-$CONH_2$). All these peptides are shown in Figure 27.

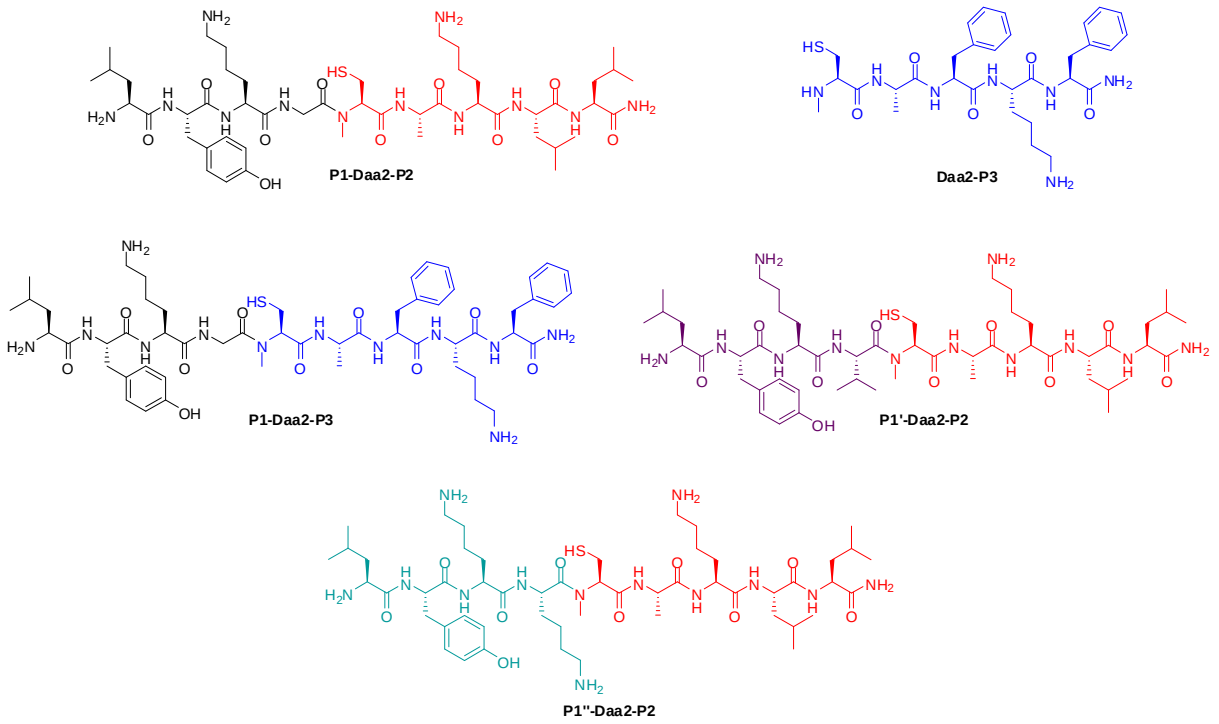


Figure 27 | Structures of the five peptides **P1-Daa2-P2**, **Daa2-P3**, **P1-Daa2-P3**, **P1'-Daa2-P2** and **P1''-Daa2-P2**.

a. General procedure

The peptides incorporating N-methyl-cysteine were synthesized and cleaved from the resin according to the general procedure reported in Chapter 2. The only differences concern the insertion of the alkyl chain on the cysteine amino group and the coupling of the following amino acid. Moreover, these peptides were less sensitive to oxidation, thereby allowing their synthesis, purification and handling in their reduced forms.

b. Insertion of the methyl group

Two main strategies have been described in the literature for the synthesis of peptides containing N-methyl residues.

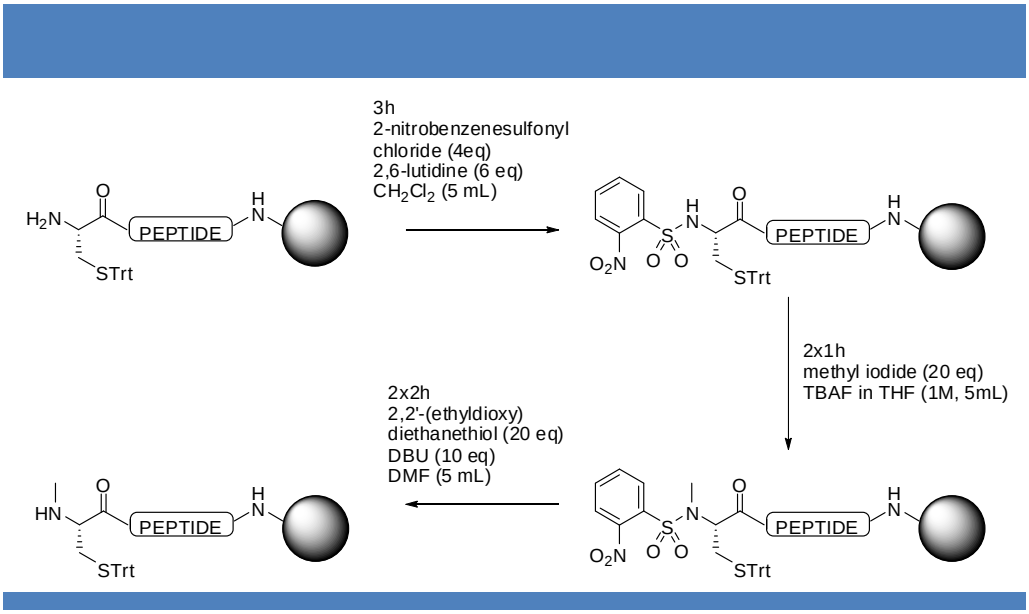
In the first strategy, an already N-methylated amino acid is introduced as a building block. The majority of Fmoc- and Boc- protected N-methyl amino acids are commercially available, although often expensive. In alternative, conveniently protected N-methyl amino can be synthesized in solution[116] and then used as building blocks in solid phase synthesis.

In the second strategy, the N-terminal residue is selectively methylated in a three step procedure. In the first step the terminal amino group is converted into a sulfonamide with 2-nitrobenzenesulfonylchloride and 2,6-lutidine or collidine, that is then removed in the third step with DBU and a thiol. N-methylation, the second step of this procedure, can be performed using different synthetic strategies. Initially, Miller and Scanlan [117] performed the deprotonation of the sulfonamide with MTBD, followed by methylation with methyl p-nitrobenzenesulfonate. This procedure was then optimized by Kessler and coworkers[118], who used DBU as base and dimethyl sulfate as alkylating agent. In alternative, a Mitsunobu reaction with methanol, DIAD and triphenylphosphine can be performed (see page 56). A different approach using methyl iodide and TBAF was reported by Erlich and coworkers[115].

For the solid phase methylation of the cysteine residue we used this last strategy. As shown in Scheme 16, the terminal amino group was first activated with 2-nitrobenzenesulfonylchloride and 2,6-lutidine in DCM under argon. Methylation with methyl iodide and TBAF (1M solution in THF) was then performed twice under argon atmosphere. Finally, the methylated amino group was deprotected by treating the resin with a

116. Sagan, S.; Karoyan, P.; Lequin, O.; Chassaing, G.; Lavielle, S. N- and Cα-Methylation in Biologically Active Peptides: Synthesis, Structural and Functional Aspects. *Curr. Med. Chem.* **2004**, *11*, 2799–2822.
117. Miller, S. C.; Scanlan, T. S. Site-Selective N-Methylation of Peptides on Solid Support. *J. Am. Chem. Soc.* **1997**, *119*, 2301–2302.
118. Biron, E.; Chatterjee, J.; Kessler, H. Optimized Selective N-Methylation of Peptides on Solid Support. *J. Pept. Sci.* **2006**, *12*, 213–219.

2,2'-(ethyldioxy)diethanethiol and DBU solution in DMF.



Scheme 16 | Schematic representation of the three-step procedure used for the N-methylation of the cysteine residue.

c. Glycine, lysine or valine coupling: acylation of N-methyl cysteine

As already extensively described in Chapter 1, the acylation of N-alkylated amino acids is challenging and generally requires harsh coupling conditions.

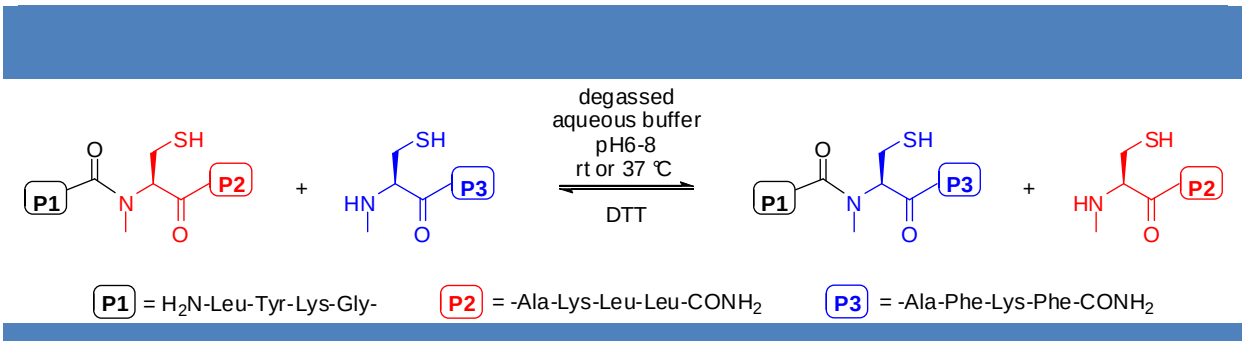
For the synthesis of peptides **P1-Daa2-P2** and **P1-Daa2-P3**, we performed the glycine coupling reaction in the microwave synthesizer using Fmoc-Gly-OH (4 eq.), HATU (4 eq.) and DIEA (5.4 eq.) in DMF (8 mL). This solution was prepared manually and added to the microwave reaction vessel; the coupling reaction was performed at 75 °C for 20 min at 30 W and repeated a second time.

For the synthesis of peptides **P1'-Daa2-P2** and **P1"-Daa2-P2**, the amino acid coupling (respectively Fmoc-Val-OH and Fmoc-Lys(Boc)-OH) was performed at higher concentrations. A solution of Fmoc-aa-OH (8 eq.), HATU (8 eq.) and DIEA (11 eq.) was added to the microwave reaction vessel and the coupling performed as reported above (i.e. 75 °C at 30 W for 20 min, repeated twice).

3. Exchange reaction between P1-Daa2-P2 and Daa2-P3

a. Exchange reaction setup

The exchange reaction described in Scheme 17 between **P1-Daa2-P2** and **Daa2-P3** was investigated at room temperature at different pHs (6, 7, 8) and at 37 °C at pH 7.



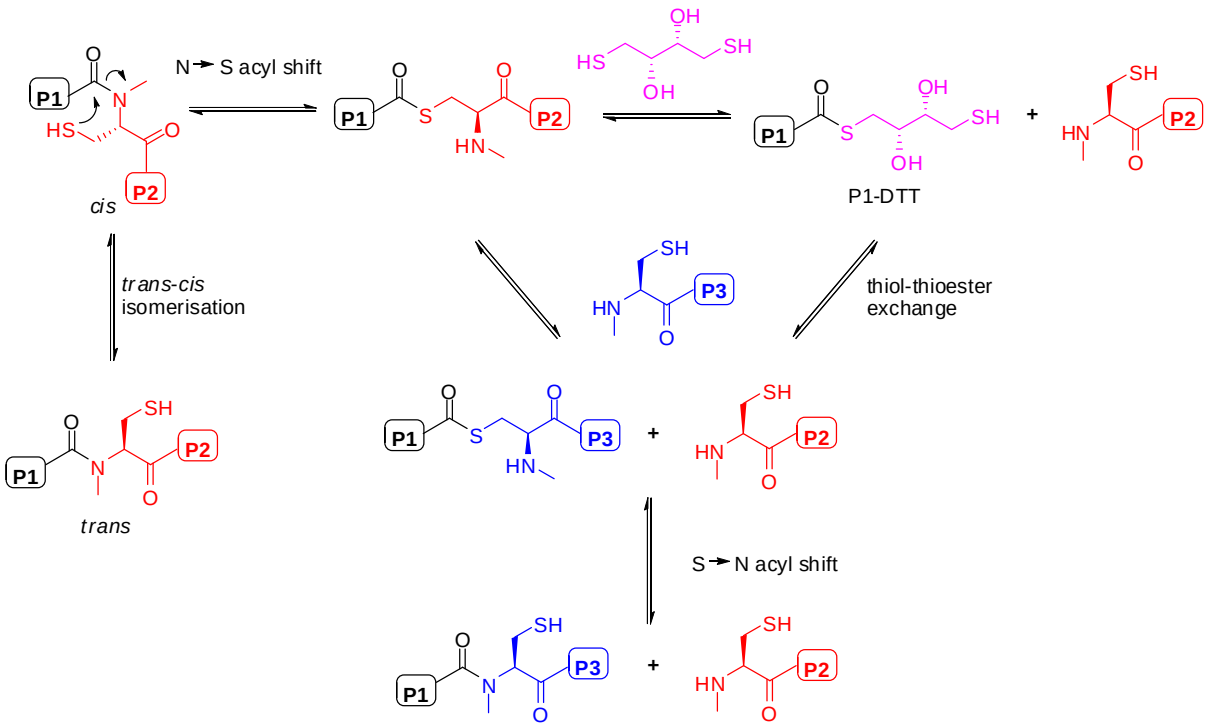
Scheme 17 | Schematic representation of a reversible NCL based on the use of **Daa2** N-methyl-cysteine.

0.2 M phosphate buffers were prepared at pH 6, 7 and 8. In order to prevent oxidation, these buffers were degassed using freeze/thaw cycles and placed under argon prior to use. For each pH, two solutions of **Daa2-P3** (formate salt) and **P1-Daa2-P2** (diformate salt) were mixed in an HPLC vial and immediately frozen and lyophilized. This lyophilized powder containing the peptides (final concentration 2.5 mM) was then diluted under argon with 500 µL of phosphate buffer (pH 6, 7 and 8) containing DTT (25 mM). The exchange reactions were then left stirring under positive argon pressure at room temperature in an HPLC vial inserted in a Schlenk tube. In the case of the exchange reaction run at 37 °C at pH 7, the Schlenk tube was put in an oil bath at constant temperature.

b. Exchange reaction mechanism

Peptide **P1-Daa1-P2** needs to be distorted in its *cis* conformation for the intramolecular nucleophilic attack of the thiol group to occur. The resulting **P1-Daa1-P2** thioester can then undergo transthioesterification to afford the **P1-Daa1-P3** thioester and **Daa1-P2**. This reaction can occur directly with **Daa1-P3** or through the formation of the **P1-DTT** intermediate that can in turn undergo thiol-thioester exchange with **Daa1-P3**. Finally, the **P1-Daa1-P3** thioester can rearrange, via S→N acyl shift, to afford peptide **P1-Daa1-P3**. All the reactions shown in

Scheme 18 are fully reversible.



Scheme 18 | Schematic representation of the *trans-cis* isomerisation, N→S acyl shift, thiol-thioester exchange and S→N acyl shift equilibriums involved in a reversible NCL based on the use of **Daa2**.

c. Exchange reaction analysis

Quantitative analysis of the exchange reaction kinetics was determined by UPLC-ESI analysis. At regular time points, aliquots (25 µL) of the exchange reaction mixture were diluted with a solution (175 µL) containing TCEP and the internal UV standard 3,5-dimethoxybenzoic acid; 2 µL of the resulting solution were then submitted to UPLC-ESI analysis and a second 2 µL injection was performed. TCEP was added to avoid the partial oxidation of the exchange reaction mixture, thus facilitating the analysis.

Typical chromatograms to monitor the exchange reactions are shown in Figure 27. At t = 0 h only the two starting peptides **P1-Daa2-P2** and **Daa2-P3** are present in solution (together with DTT and the UV standard); although we could clearly observe the formation of the exchange products **P1-Daa2-P3** and **Daa2-P2**, together with **P1-DTT**, after several hours, Figure 28 shows the chromatograms after 4 days of equilibration.

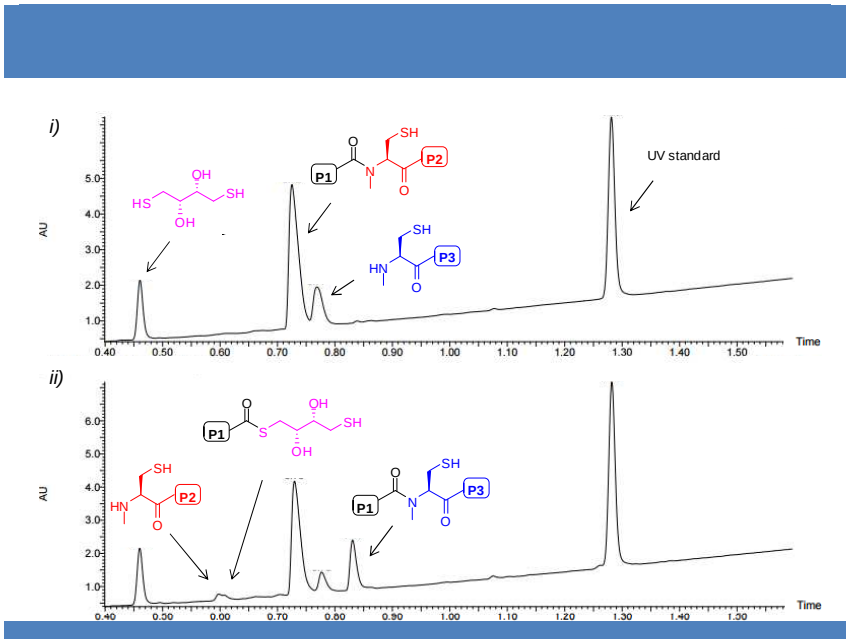


Figure 28 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the exchange reaction between **P1-Daa2-P2** and **Daa2-P3** at pH 7 at *i)* t = 0 h and *ii)* t = 96 h.

Calibration curves (Figure 29) were established for peptides **P1-Daa2-P2** and **P1-Daa2-P3** by injecting 2 µL of the peptides at known concentrations (125.00, 156.25, 312.50, 500 and 625.00 µM) under the same conditions used for the dilution of the aliquots of the exchange reactions. Each standard point corresponds to an average of 3 replicates. After a manual baseline correction with the peak analyzer tool of OriginPro 9.0, the area of the peak corresponding to the peptide was corrected according to the area of the internal UV standard and plotted against the concentration. A linear regression analysis gave 2 linear calibration curves with R^2 of 0.999 and 0.991 that where then used to calculate the concentration of the peptides in the exchange mixtures.

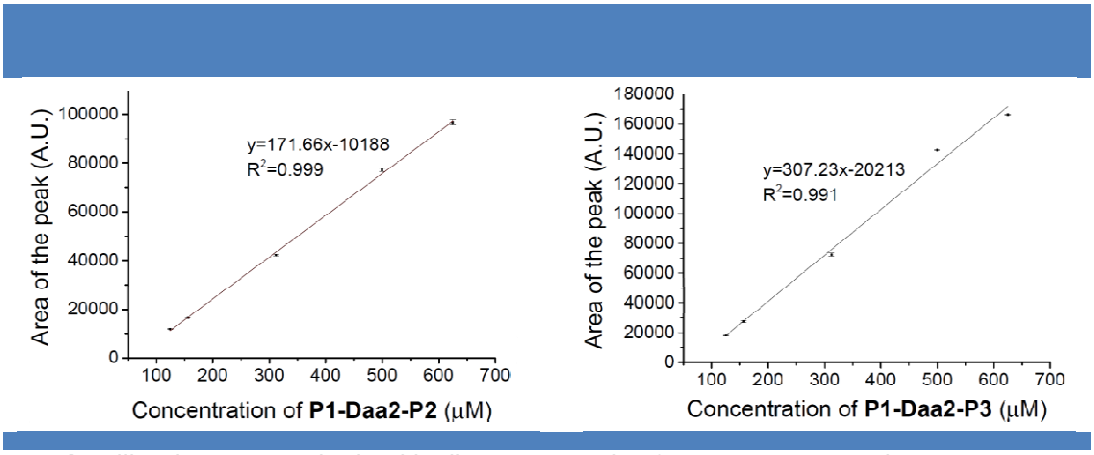


Figure 29 | Calibration curves obtained by linear regression for **P1-Daa2-P2** and **P1-Daa2-P3**.

For each exchange reaction, each timepoint was determined as a mean of two separate injections in the UPLC-ESI. In order to calculate the kinetics parameters of the exchange reaction at each of the four conditions investigated, the UPLC-ESI chromatograms recorded with UV detection were processed as follows. After a manual baseline correction with the peak analyzer tool of OriginPro 9.0, the area values of the peaks corresponding to **P1-Daa2-P2** and **P1-Daa2-P3** were extracted and corrected according to the area value of the UV standard for each injection. For each timepoint, the mean between the two area values obtained was then expressed as concentration [**P1-Daa2-P2**] and [**P1-Daa2-P3**] of the two peptides, according to the calibration curves reported above. Finally, we expressed the % of the exchange product [**P1-Daa2-P3**] as [**P1-Daa2-P3**]/([**P1-Daa2-P2**]+[**P1-Daa2-P3**]) (Figure 30) and calculated fitting curves for each set of values.

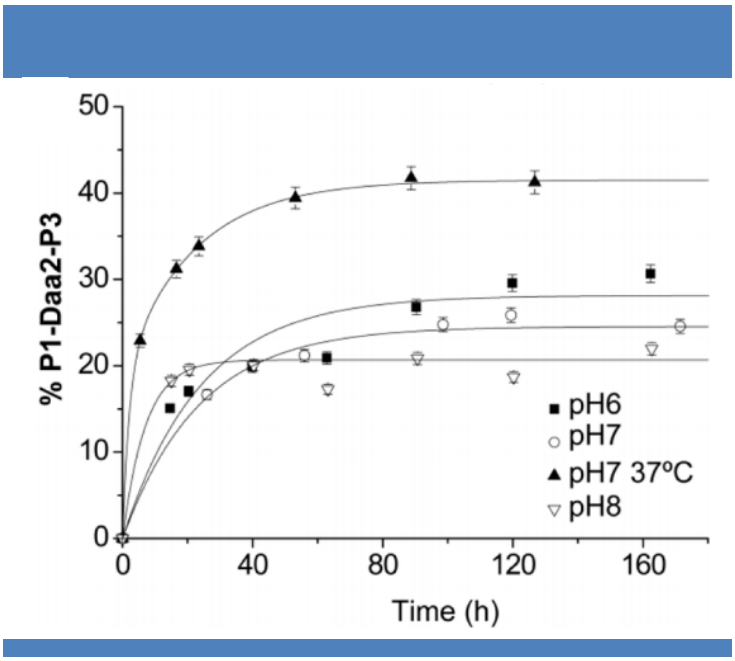


Figure 30 | % of exchange product [**P1-Daa2-P3**] ([**P1-Daa2-P3**]/([**P1-Daa2-P2**]+[**P1-Daa2-P3**]) as a function of time starting from 2.5 mM solutions of **P1-Daa2-P2** and **P1-Daa2-P3** at room temperature at pHs 6, 7 and 8 and at 37 ℃ at pH 7 in a phosphate buffer (200mM, DTT 25mM).

We could extrapolate the half-times of equilibration ($t_{1/2}$) 25, 20 and 6 h and initial rates (V_0) 1.11×10^{-8}, 1.38×10^{-8} and 4.16×10^{-8} $M \cdot s^{-1}$respectively at pH 6, 7 and 8 at room temperature. The V_0 values were calculated using the "tangent" tool of the Origin 9.0 software at t = 0 h. If compared to the exchange reaction between peptides incorporating **Daa1**, the average rates of the reactions are slower. The presence of only one thiol group in this dynamic unit, in contrast to the two thiol groups present in **Daa1**, halves the possibility of the initial intramolecular shift occurring.

Finally, we investigated the exchange reaction at pH 7 and at 37 °C. The set of kinetic

parameters obtained ($t_{1/2}$ = 10 h and V_0 = 5.00 × 10^{-8} M·s^{-1}) ranges in a reasonable time scale to extend dynamic covalent chemistry to peptides in fully biocompatible conditions. Importantly, we also observed an increase of the equilibrium constant (K), that is known to be affected by the reaction medium and in particular by temperature[55], other than by pH.

Interestingly, in the case of peptides incorporating N-methyl-cysteine, we observed only traces of thioester **P1-DTT** even when performing the exchange reaction at pH 6.

d. Exchange reaction in detail

i. Characterization of the thioester intermediates

Previous work [119] on peptides incorporating N-alkyl-cysteine, including N-methyl-cysteine, shows that low pH favors their thioester isomer, whereas the amide configuration is predominant at neutral pH. We were thus confident that all the observed peptides were in their amide configuration at all the pHs investigated. However, we confirmed our hypothesis by lowering the pH of the solutions, thus promoting the formation of the corresponding thioester isomer.

A first experiment was run on peptide **P1-Daa2-P2**; a solution of the peptide in a pH 6 phosphate buffer containing TCEP (28 mM) was acidified with 2.5 %v/v of TFA and let stir for 20 min. As a result we could observe a new peak corresponding to the peptide thioester. Neutralization of the solution with $(NH_4)_2CO_3$ led to its reconversion into the more stable amide peptide **P1-Daa2-P2** (Figure 31).

119. Hojo, H.; Onuma, Y.; Akimoto, Y.; Nakahara, Y.; Nakahara, Y. N-Alkyl Cysteine-Assisted Thioesterification of Peptides. *Tetrahedron Lett.* **2007**, *48*, 25–28.

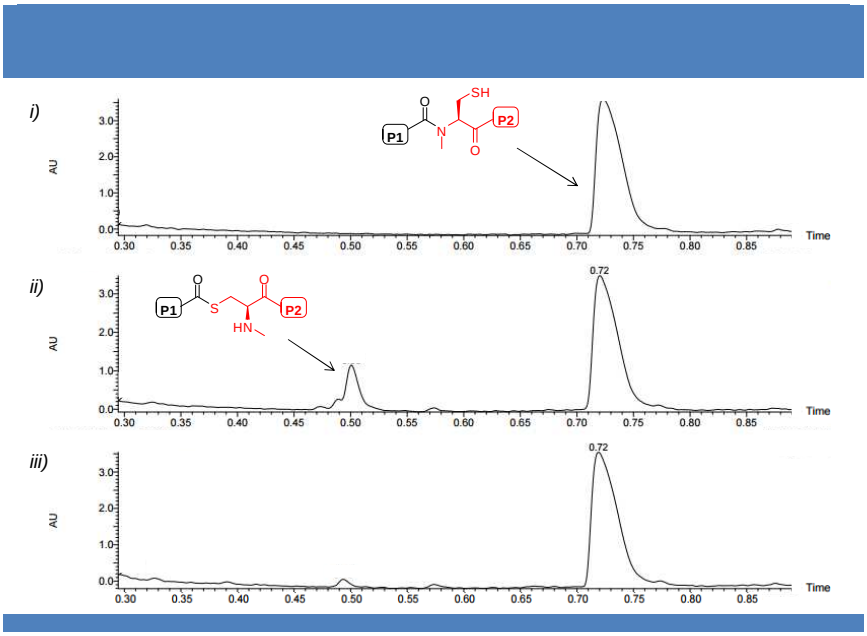


Figure 31 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of **P1-Daa2-P2** (320 µM), *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 20 min in the same solution + 2.5 %v/v of TFA; *iii)* after subsequent neutralization to pH 7 using $(NH_4)_2CO_3$ for 5 min.

A second experiment was run on both peptides **P1-Daa2-P2** and **P1-Daa2-P3** to unambiguously assess the thioester nature of the new peak formed in acidic conditions. Two separate solutions of peptides **P1-Daa1-P2** and **P1-Daa1-P3** (320 µM) at pH 6 were thus acidified by adding 2.5 %v/v TFA. The resulting mixtures, containing the amide and thioester isomers of each peptide, were then treated with an equivalent volume of a 2 M solution of hydroxylamine (buffered at pH 7). As a result (Figure 32), for each of the peptides **P1-Daa2-P2** and **P1-Daa2-P3**, UPLC-ESI analysis showed the disappearance of the peaks corresponding to the thioester isomer and the formation of the hydroxamic acid of the N-terminal fragment together with the release of the C-terminal fragment (respectively **Daa2-P2** and **Daa2-P3**).

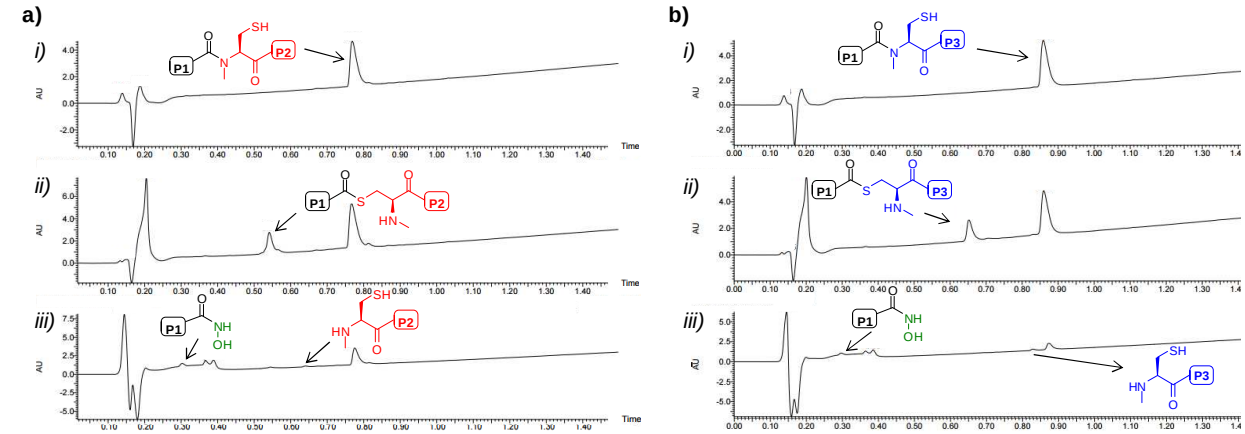


Figure 32 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of **a) P1-Daa2-P2** (320 μM), *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 20 min in the same solution + 2.5 %v/v of TFA; *iii)* after a 5 min treatment with an equivalent volume of a 2 M hydroxylamine solution buffered at pH 5; **b) P1-Daa2-P3** (320 μM), *i)* at pH 6 under reductive conditions in a 28 mM solution of TCEP; *ii)* after 180 min in the same solution + 2.5 %v/v of TFA; *iii)* after a 1 min treatment with an equivalent volume of a 2 M hydroxylamine solution buffered at pH 5. The poor quality of the baseline before 0.45 min is due to the presence of hydroxylamine at high concentrations.

ii. Characterization and quantification of the hydrolysis product

After long time of equilibration, and in particular at pH 8, the peak corresponding to the hydrolysis product of the **P1-DTT** thioester appeared in the reverse phase chromatograms. In order to show the stability of the exchange product **P1-Daa2-P3**, we have plotted its absolute concentration against the time for each of the four conditions investigated. **P1-Daa2-P3** proved to be stable even long time after the thermodynamic equilibrium was reached, allowing us to conclude that the hydrolysis byproduct can be neglected in the timescale of the reaction (Figure 33).

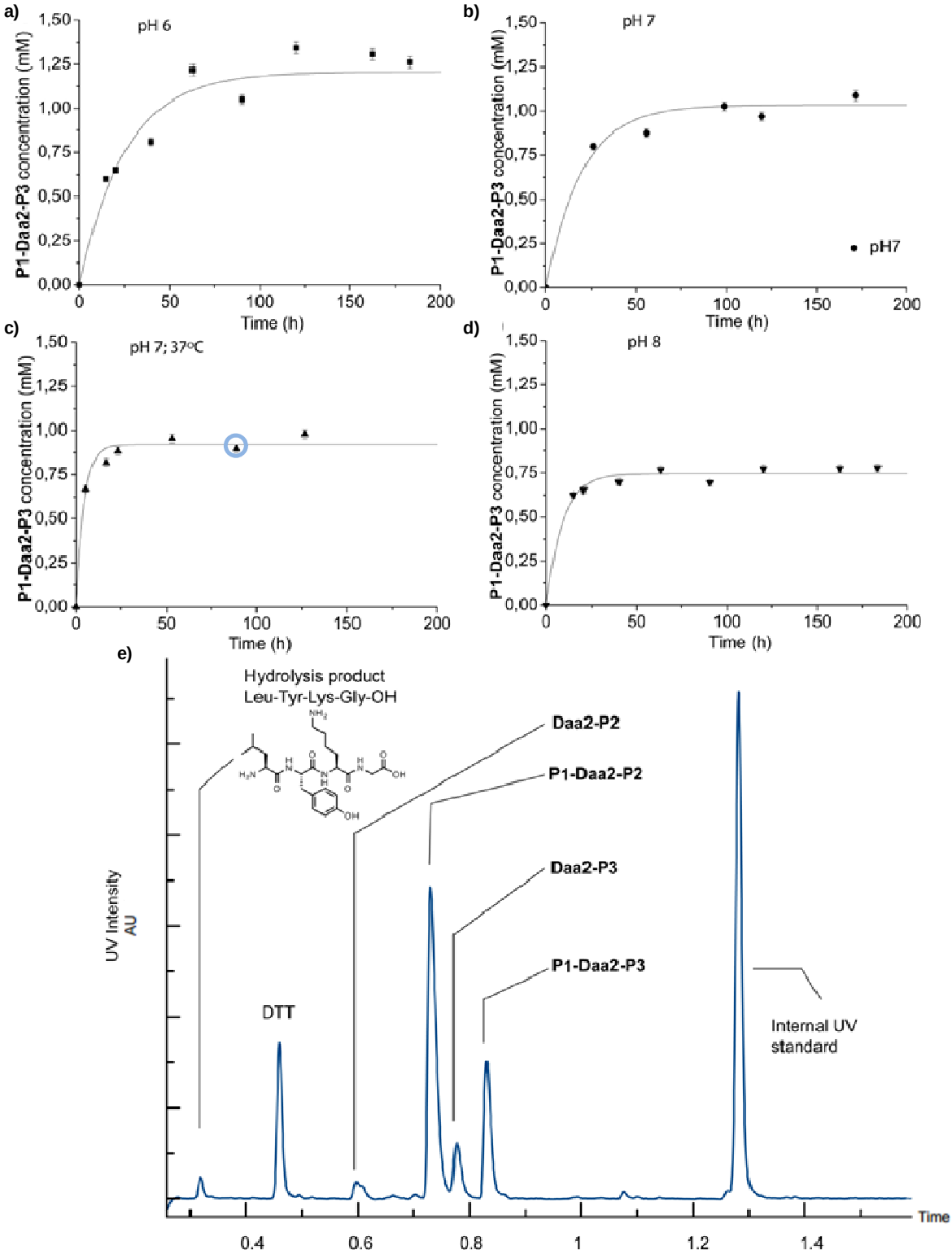


Figure 33 | Absolute concentration of exchange peptide **P1-Daa2-P3** plotted against the time of equilibration at **a)** pH 6 and rt, **b)** pH 7 and rt, **c)** pH 7 and 37 ℃ and **d)** pH 8 and rt. **e)** Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection at pH 7, 37 ℃ and after 98 h of equilibration, corresponding to the timepoint circled in **c)**.

iii. Exchange reaction in the presence of TCEP

TCEP is a stronger reducing agent than DTT and less sensitive to air oxidation. Moreover, trialkylphosphines have been shown[120] to accelerate thiol-thioester exchanges by activating the thioesters as phosphonium salts. On the other hand, as already mentioned in Chapter 2, TCEP is known to promote desulfurization[95] of the cysteine moiety into an alanine residue through a radical mechanism.

We performed the exchange reaction at pH 7 and room temperature in the same conditions reported above (i.e. same procedure, peptide concentration 2.5 mM) but in the presence of TCEP (25 mM) instead of DTT. In these conditions, after only 19.5 h we could observe a strong proportion of desulfurized starting material and the absence of exchange products. We then repeated the same reaction in the presence of both DTT and TCEP (25 mM each), expecting DTT to act as a radical scavenger and prevent desulfurization. Unfortunately, after 29 h we obtained a mixture of the exchange products together with the desulfurized starting material (Figure 34).

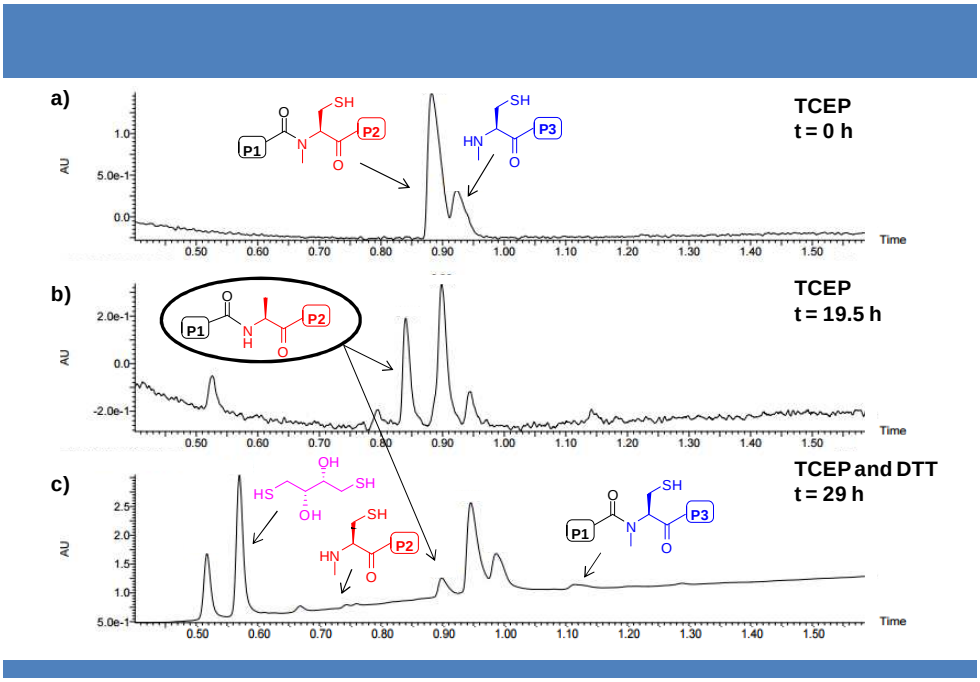


Figure 34 | Reversed phase chromatograms (UPLC-ESI) recorded with UV and MS detection (**a)** t = 0 h; **b)** t = 19.5 h) of the exchange reaction between **P1-Daa2-P2** and **Daa2-P3** at pH 7 and room temperature in the presence of TCEP (25 mM) instead of DTT. **c)** Reversed phase chromatograms (UPLC-ESI) recorded with UV and MS detection of an exchange reaction run in the same conditions and in a 1:1 mixture of DTT (25 mM) and TCEP (25 mM) showing the formation of both exchange products and desulfurized starting material after 29 h.

120. Tam, J. P.; Lu, Y. A.; Liu, C. F.; Shao, J. Peptide Synthesis Using Unprotected Peptides through Orthogonal Coupling Methods. *Proc. Natl. Acad. Sci.* **1995**, *92*, 12485–12489.

4. Exchange reaction between P1'-Daa2-P2 and Daa2-P3

The rate and efficacy of native chemical ligation have been shown[22] to be strictly correlated to the nature of the C-terminal residue thioester. In the case of sterically hindered amino acids such as proline, isoleucine and valine, the reaction proceeds prohibitively slowly even in the presence of thiophenol catalysts.

In order to probe the limit of our reversible native chemical ligation approach, we have synthesized **P1'-Daa2-P2**, bearing a valine residue (instead of a glycine) in α-position to **Daa2** and investigated its exchange reaction with **Daa2-P3**. The conditions were optimized to promote the formation of the exchange products: peptide and DTT concentrations were increased (respectively to 7.5 mM and 50 mM) and the reaction was run at pH 8 and 37 °C. Within three days of equilibration we could indeed observe the formation of the exchange products **P1'-Daa2-P3** and **Daa2-P2** (Figure 35), confirming the claim that our methodology is strictly related to the original non-reversible NCL. Interestingly, we did not observe any trace of the **P1'-DTT** thioester, nor of the hydrolysis byproducts.

The precise measurement of the kinetic parameters encompassed the scope of this experiment. Due to the long time of equilibration, the exchange reaction was stopped after the detection of the exchange products, before it could reach its thermodynamic equilibrium. Moreover, the precise analysis of the reaction kinetics was precluded by precipitation, presumably of the exchange product **P1'-Daa2-P3**, during the course of the reaction.

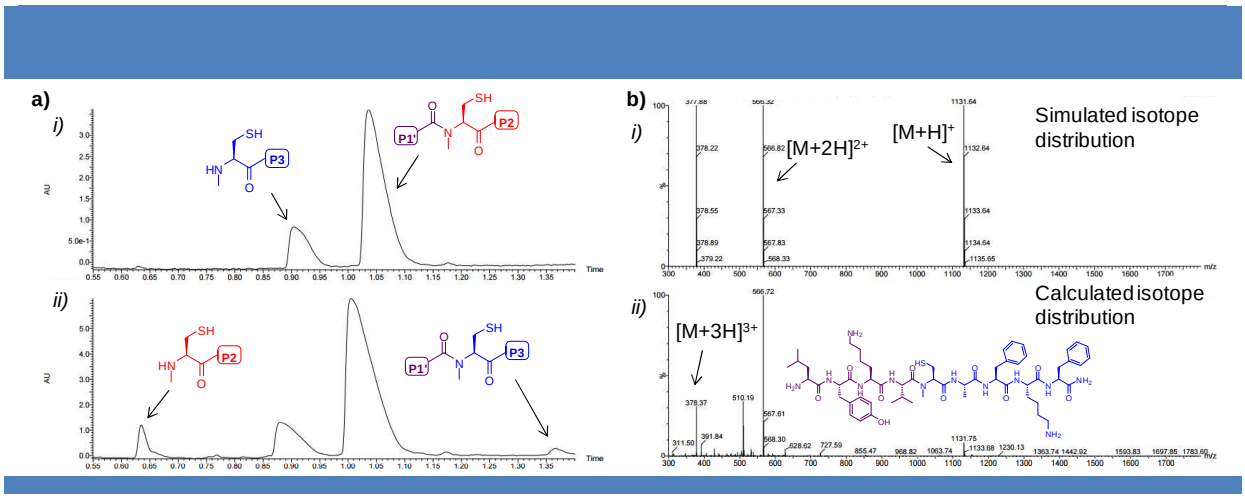


Figure 35 | a) Reversed phase chromatograms (UPLC-ESI) recorded with UV and MS detection at *i)* t = 0 h and *ii)* t = 19.5 h) of the exchange reaction between **P1'-Daa2-P2** and **Daa2-P3** (7.5 mM) at Ph 8 and 37 ℃ in the presence of DTT (50 mM). **b)** *i)* Calculated and *ii)* experimentally obtained mass spectra for the exchange product **P1'-Daa2-P3**.

5. Exchange reaction between P1”-Daa2-P2 and P1-Daa2-P3

The reactions reported so far occurred between two peptides bearing the dynamic unit (**Daa1** or **Daa2**) internally in one case and at the C-terminus in the other case (peptide transamidation). In order to broaden the scope of our methodology, thus extending it to peptide metathesis, we investigated the exchange reaction occurring between two peptides containing an internal **Daa2**.

Model peptides **P1”-Daa2-P2** and **P1-Daa2-P3** were thus synthesized according to the procedure previously reported. The exchange reaction between them (Figure 36a) was then initially performed according to the standard protocol reported above (peptide concentration 2.5 mM, DTT concentration 25 mM) at pH 7. Although we could detect the exchange products, we optimized the conditions to promote their formation and repeated the reaction. In order to clearly characterize the species involved in the equilibrium, we increased the concentration of starting peptides (5 mM) and DTT (125 mM) and we brought the temperature to 37 °C. As a result, we could indeed observe the formation of the two exchange products **P1”-Daa2-P3** and **P1-Daa2-P2**, together with the other species involved in the equilibrium, i.e. **Daa2-P2**, **Daa2-P3**, **P1”-DTT**, **P1-DTT** and the starting peptides **P1”-Daa2-P2** and **P1-Daa2-P3** (Figure 36b).

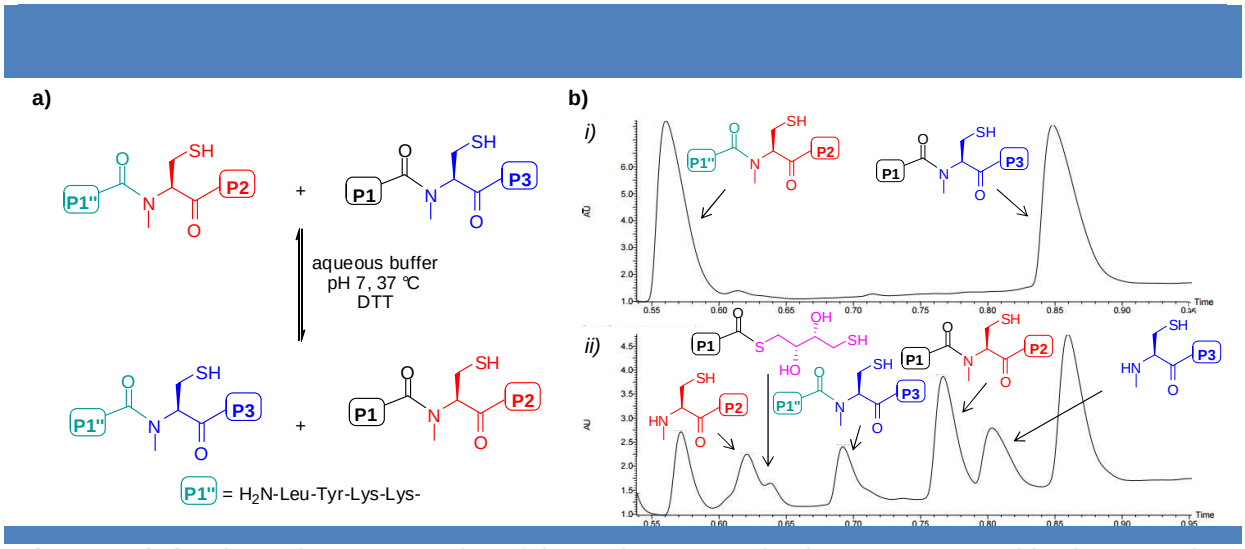


Figure 36 | a) Schematic representation of the exchange reaction between two peptides incorporating internal dynamic units **P1”-Daa2-P2** and **P1-Daa2-P3**. **b)** Corresponding reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the exchange reaction described in **a)** at *i)* t = 0 and *ii)* t = 48 h for an initial concentration of 5 mM of each peptide, at pH 7 and 37 °C in a phosphate buffer containing DTT (125 mM).

Chapter 4 - Affibody Molecules: Toward a First Biological Application of Reverse NCL

1. Introduction

a. Affinity proteins

Affinity proteins are nowadays invaluable tools in the fields of biotechnology and medicine. Their ability to bind a plethora of different molecules with high affinity and specificity has been extensively explored, leading to a variety of applications.

Immunoglobulins (Ig, also called antibodies, Ab) are naturally designed affinity proteins, used by the immune system to identify a specific antigen on pathogens, such as bacteria or viruses. Due to their high affinity to a specific molecular target, they have been routinely employed in research and diagnostics. Moreover, a number of antibody-based drugs have been approved for therapeutic use and others are currently in clinical development[121]. Despite being the most widely used affinity proteins for a plethora of different applications, immunoglobulins have some disadvantages. These large (150 kDa) proteins are composed of different polypeptide chains (Figure 37) and have complex glycosylation patterns, making them difficult to manufacture. Moreover, for therapeutic applications they show limitations in terms of impaired interactions with the immune system, pharmacokinetics and tissue accessibility[122].

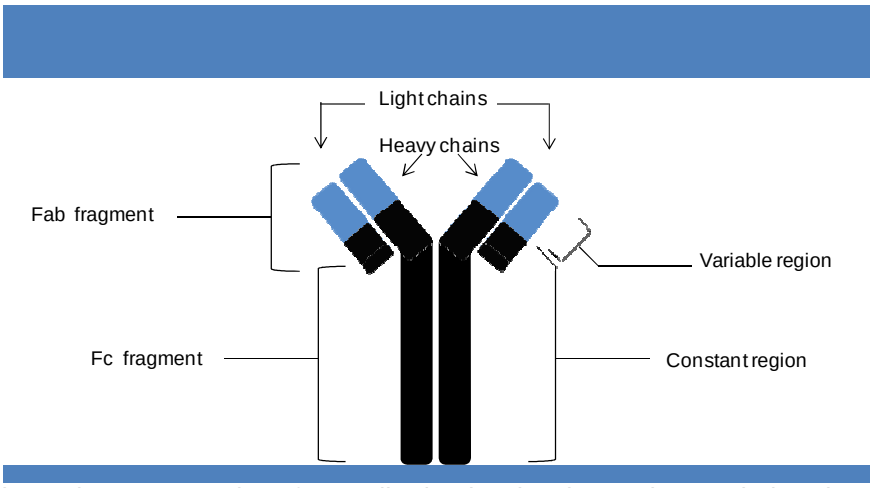


Figure 37 | Schematic representation of an antibody, showing the two heavy chains, the two light chains, the variable and constant regions and the Fab and Fc fragments.

121. Leavy, O. Therapeutic Antibodies: Past, Present and Future. *Nat. Rev. Immunol.* **2010**, *10*, 297.
122. Chames, P.; Van Regenmortel, M.; Weiss, E.; Baty, D. Therapeutic Antibodies: Successes, Limitations and Hopes for the Future. *Br. J. Pharmacol.* **2009**, *157*, 220–233.

Advances in protein engineering, together with an increased understanding of the immunoglobulin chemistry and interactions, have led to the development of recombinant antibody fragments[123]. The smaller size of these engineered proteins allows one to overcome the majority of limitations associated to the clinical use of antibodies.

Although very important biomolecules, immunoglobulins are far from being the only naturally occurring binding proteins. Other affinity proteins are present in nature and have been investigated as a starting point for the development of new affinity ligands[124]. Around 50 different protein scaffolds have hitherto been reported and, although very different in terms of structure and function, they show similar general features. They are relatively small proteins, often composed by a single polypeptide chain, and are based on stable architectures. They can thus be obtained in high yields by bacterial expression and, in some cases, by solid phase synthesis, allowing the insertion of non-biological groups. The general absence of cysteine residues improves stability and, in addition, opens the possibility to the insertion of one of these moieties in a specific position for successive conjugation. Finally, affinity protein structures have to be tolerant to a certain degree of randomization. This is of fundamental importance for the creation of libraries and the consequent selection of their members specific for a given target.

b. Affibody molecules

Affibody molecules are among the most successfully used and documented non-Ig affinity molecules, originally derived from the Ig-binding domain B in the staphylococcal protein A. The high solubility and folding properties of these relatively short peptides, suggested further investigation; they were thus engineered in order to increase chemical stability. The engineered variant of this 58 amino acid peptide, denoted Z domain[125], retained its affinity toward the Fc fragment of the antibody and showed interesting properties. It consists of three α-helices (Figure 38a) forming a very stable three-bundle structure, characterized by rapid folding kinetics. Due to small size (6.5 kDa), affibody molecules can be either produced in a prokaryotic host, with low manufacturing costs, or synthesized by SPPS, enabling the insertion of different chemical modifications. This domain has thus been investigated as a scaffold for the selection of novel affinity peptides.

123. Holliger, P.; Hudson, P. J. Engineered Antibody Fragments and the Rise of Single Domains. *Nat. Biotechnol.* **2005**, *23*, 1126–1136.
124. Grönwall, C.; Ståhl, S. Engineered Affinity Proteins--Generation and Applications. *J. Biotechnol.* **2009**, *140*, 254–269.
125. Nilsson, B.; Moks, T.; Jansson, B.; Abrahmsén, L.; Elmblad, A.; Holmgren, E.; Henrichson, C.; Jones, T. A.; Uhlén, M. A Synthetic IgG-Binding Domain Based on Staphylococcal Protein A. *Protein Eng. Des. Sel.* **1987**, *1*, 107–113.

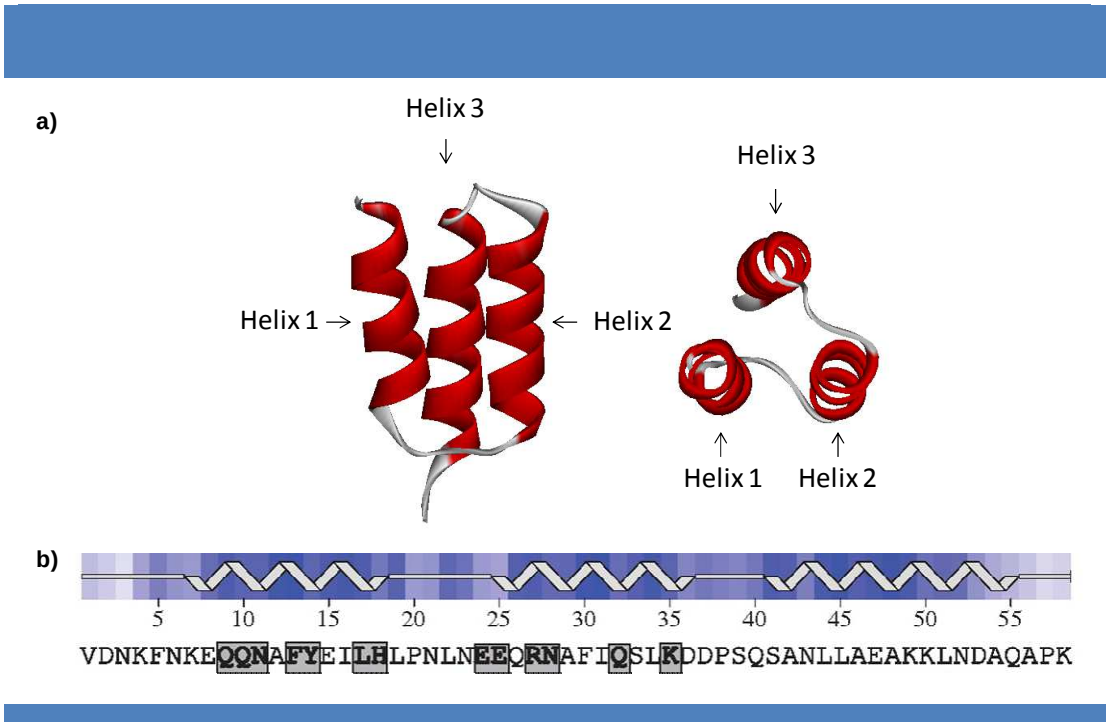


Figure 38 | a) Frontal and top view of the affibody molecule (Z domain). **b)** Sequence of the Z domain; the randomized positions in affibody libraries are shown in grey (adapted from ref. 126).

Combinatorial libraries were created by randomization of 13 surface-exposed residues on helices 1 and 2 (Figure 38b), including many amino acids originally involved in the Fc-binding region of the Z domain and their members denoted affibodies. The classical strategy involved the creation of libraries by phage display, followed by biopanning against different targets[127]. Different affibody molecules were thus selected, demonstrating high affinity (K_D) in the pM to μM range toward specific targets[128].

Initially investigated for biotechnological applications, affibody molecules have recently proved to be invaluable tools in imaging as well as in therapy[129]. An overview of the applications of affibody molecules is shown in Figure 39.

126. Eigenbrot, C.; Ultsch, M.; Dubnovitsky, A.; Abrahmsén, L.; Härd, T. Structural Basis for High-Affinity HER2 Receptor Binding by an Engineered Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 15039–15044.

127. Nord, K.; Gunneriusson, E.; Ringdahl, J.; Ståhl, S.; Uhlén, M.; Nygren, P. A. Binding Proteins Selected from Combinatorial Libraries of an Alpha-Helical Bacterial Receptor Domain. *Nat. Biotechnol.* **1997**, *15*, 772–777.

128. Nygren, P.-A. Alternative Binding Proteins: Affibody Binding Proteins Developed from a Small Three-Helix Bundle Scaffold. *FEBS J.* **2008**, *275*, 2668–2676.

129. Löfblom, J.; Feldwisch, J.; Tolmachev, V.; Carlsson, J.; Ståhl, S.; Frejd, F. Y. Affibody Molecules: Engineered Proteins for Therapeutic, Diagnostic and Biotechnological Applications. *FEBS Lett.* **2010**, *584*, 2670–2680.

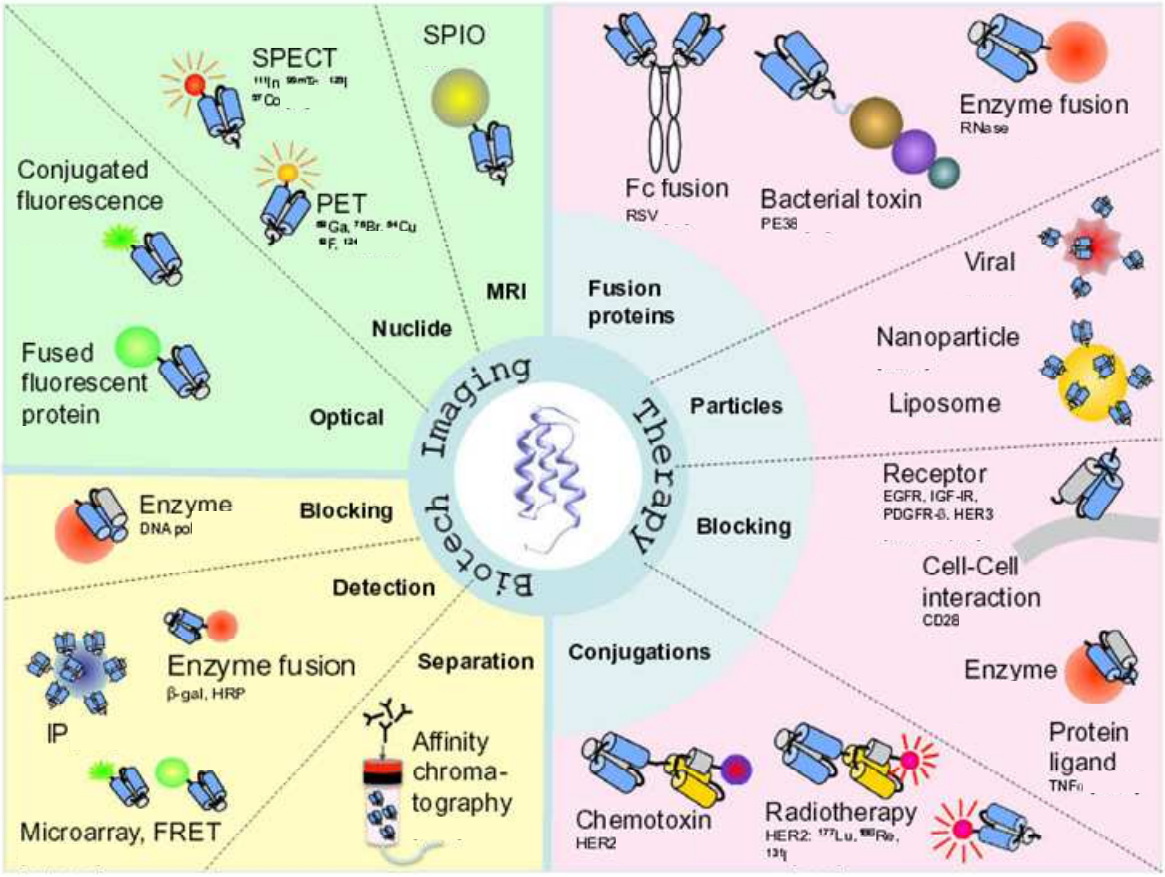


Figure 39 | Overview of selected applications of affibody molecules (adapted from ref. 129).

c. Design and objectives

The description of the first example of abiotic reversible native chemical ligation was indeed an important achievement. However, an important milestone in the project was to use reversible NCL for the preparation or discovery of bioactive peptides.

With an affinity (K_D) of 22 pM[130] for its important biological target, Z_{HER2} is probably the most extensively studied affibody. HER2 is a member of the human epidermal growth factor receptor family and, being overexpressed in an aggressive form of breast cancer, has become an important biomarker and target of therapy. The expression status of HER2 predicts the response of HER2-target therapy; it is therefore important to mention that a ^{111}In-Z_{HER2} has recently been used for in-human imaging of the HER2 expression in breast cancer metastases[131]. As shown in

130. Orlova, A. Tumor Imaging Using a Picomolar Affinity HER2 Binding Affibody Molecule. *Cancer Res.* **2006**, *66*, 4339–4348.
131. Sörensen, J.; Sandberg, D.; Sandström, M.; Wennborg, A.; Feldwisch, J.; Tolmachev, V.; Åström, G.; Lubberink, M.; Garske-Román, U.; Carlsson, J.; Lindman, H. First-in-Human Molecular Imaging of HER2 Expression in Breast Cancer Metastases Using the 111In-ABY-025 Affibody Molecule. *J. Nucl. Med.* **2014**, *55*, 730–735.

Figure 40, this affibody specifically binds HER2 at the junction of domains III and IV in its extracellular domain (ecd), in a position distant from those targeted by the traditional therapeutic antibodies trastuzumab and pertuzumab.

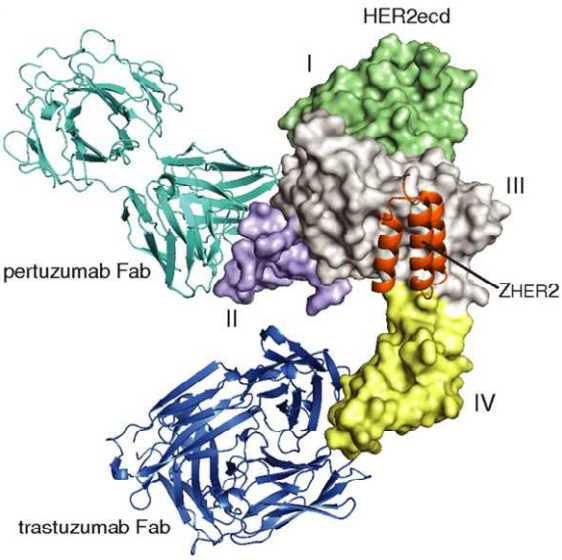


Figure 40 | Crystal structure of the affibody molecule Z_{HER2} (red) in a complex with HER2ecd. For reference, Fab fragments of the therapeutic antibodies pertuzumab (cyan) and trastuzumab (blue) are shown in the positions where they bind HER2ecd (from ref. 126).

Our goal was the preparation of affibodies using a dynamic chemistry approach via reversible NCL. Interestingly, NCL has already been used to synthesize affibody molecules[132] specific for IgG (Z_{IgG}), HER2 (Z_{HER2}) and insulin ($Z_{Insulin}$). The substitution of Gln-Ser with Gly-Cys in positions 40-41, in the loop between helices 2 and 3, determined only minor changes in the binding affinities of these molecules for their targets. The authors also substituted Asp with Glu in position 2, not involved in the binding, to prevent possible aspartimide formation during peptide synthesis. We modified this Z_{HER2} sequence (Figure 41a) by introducing **Daa2** (N-methyl-cysteine) in the two loops between the three helices of Z_{HER2}. We synthesized the three peptides corresponding to the three α-helices of Z_{HER2} and incorporating **Daa2** moieties (helices 1, 2 and 3, respectively in Figures 41b, 41c and 41d). We then planned to investigate the exchange reaction among these three fragments in the presence or absence of the target molecule HER2. In this latter case, we expected to observe an amplification of the exchange peptide analogue to Z_{HER2}, the H1-H2-H3 trimer (Figure 41e).

132. Lindgren, J.; Ekblad, C.; Abrahmsén, L.; Eriksson Karlström, A. A Native Chemical Ligation Approach for Combinatorial Assembly of Affibody Molecules. *ChemBioChem* **2012**, *13*, 1024–1031.

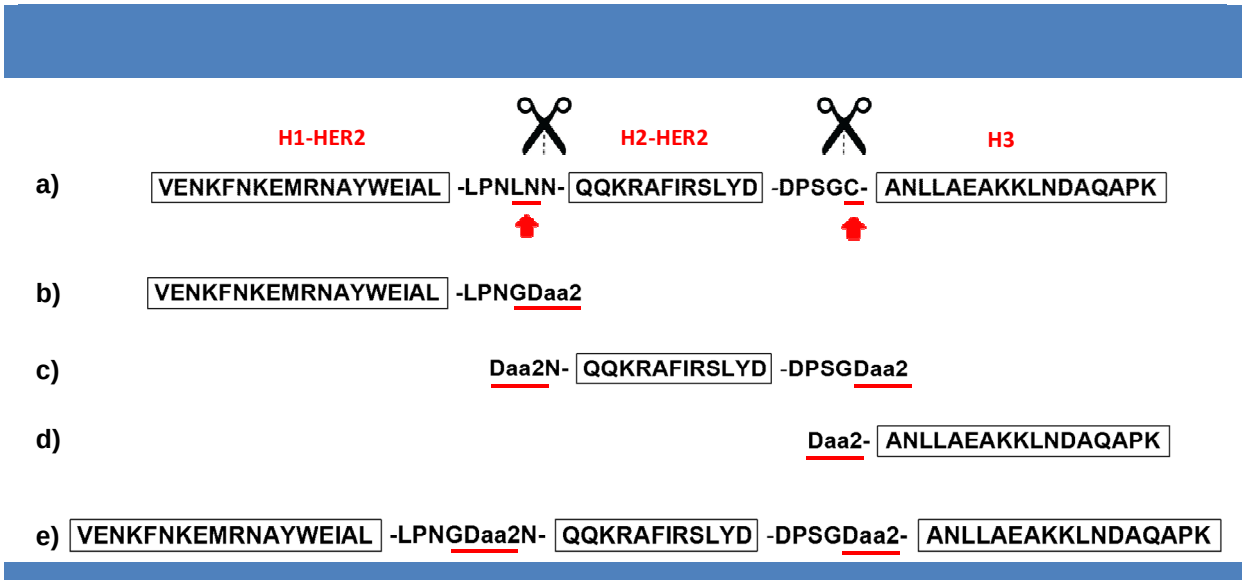


Figure 41 | a) Z_{HER2} sequence synthesized by NCL in ref. 132; design of affibody helices **b) H1-HER2**, **c) H2-HER2** and **d) H3**; **e)** H1-H2-H3 exchange product trimer. Helical sequences are shown inside the rectangles, whereas the residues in between correspond to the loops connecting the helices; the amino acids underlined in red correspond to the residues that have been modified from the original Z_{HER2} shown in **a)**.

Despite the high affinity of Z_{HER2} for its target molecule, the high cost (1 mg ~ 1000 €) of HER2 was a serious limitation for our investigations. We thus started a collaboration with Dr. Marc Ruff in the IGMBC (Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France) to express the extracellular domain of HER2 in cells. After a first failed attempt to express this domain in *Escherichia coli*, following a published procedure[133], he is currently trying to express it in CHO (Chinese Hamster Ovary) cells[134].

In parallel, we investigated Z_{IgG}, that specifically binds a much cheaper target protein (50 mg ~ 380 €), although with a lower affinity (K_D = 22 nM). We thus synthesized a second set of **Daa2**-containing affibody helices, derived from Z_{IgG} and investigated the exchange reaction among them. In the presence of the protein (IgG) we expected to observe an amplification of the H1-H2-H3 trimer, analogue to Z_{IgG} and shown in Figure 42b.

Whereas helix 3 is the same for both the HER2-based and IgG-based libraries, helices 1 and 2 differ for 13 residues that, as explained earlier in this Chapter, have been randomized in order to select affibodies that strongly bind their specific target.

133. Liu, X.; He, Z.; Zhou, M.; Yang, F.; Lv, H.; Yu, Y.; Chen, Z. Purification and Characterization of Recombinant Extracellular Domain of Human HER2 from Escherichia Coli. *Protein Expr. Purif.* **2007**, *53*, 247–254.
134. Franklin, M. C.; Carey, K. D.; Vajdos, F. F.; Leahy, D. J.; de Vos, A. M.; Sliwkowski, M. X. Insights into ErbB Signaling from the Structure of the ErbB2-Pertuzumab Complex. *Cancer Cell* **2004**, *5*, 317–328.

a) VENKFNKEMRNAYWEIAL -LPNGDaa2N- QQKRAFIRSLYD -DPSGDaa2- ANLLAEAKKLNDAQAPK

b) VENKFNKEQQNAFYEILH -LPNGDaa2E- EQRNAFIQSLKD -DPSGDaa2- ANLLAEAKKLNDAQAPK

Figure 42 | H1-H2-H3 exchange product trimers specific for **a)** Z_{HER2} and **b)** Z_{IgG}. The sequences corresponding to the three helices are shown in the rectangles; the randomized residues are shown in grey; the amino acids underlined correspond to the residues that have been modified from the original **a)** Z_{HER2} and **b)** Z_{IgG} sequences described in ref. 132.

2. Peptide synthesis

In order to investigate the exchange reaction among affibody helices, we synthesized the helices specific for Her2 and IgG in their StBu protected form: helix 1 (Her2) **H1(StBu)-HER2**, helix 2 (Her2) **H2(StBu)$_2$-HER2**, helix 3 (Her2 and IgG) **H3(StBu)**, helix 1 (IgG) **H1(StBu)-IgG** and helix 2 (IgG) **H2(StBu)$_2$-IgG**. Moreover, helix 2 (Her2) incorporating just one dynamic end was synthesized **H2a(StBu)-HER2**. All these peptides are shown in Figure 43.

H1(StBu)-HER2:	NH_2-VENKFNKEMRNAYWEIALLPNG-Daa2(StBu)-$CONH_2$
H2(StBu)$_2$-HER2:	NH_2-Daa2(StBu)-NQQKRAFIRSLYDDPSG-Daa2(StBu)-$CONH_2$
H3(StBu):	NH_2-Daa2(StBu)-ANLLAEAKKLNDAQAPK-$CONH_2$
H1(StBu)-IgG:	NH_2-VENKFNKEQQNAFYEILHLPNG-Daa2(StBu)-$CONH_2$
H2(StBu)$_2$-IgG:	NH_2-Daa2(StBu)-EEQRNAFIQSLKDDPSG-Daa2(StBu)-$CONH_2$
H2a(StBu)-HER2	NH_2-Daa2(StBu)-NQQKRAFIRSLYDDPSG-$CONH_2$

Figure 43 | Structures of peptides **H1(StBu)-HER2**, **H2(StBu)$_2$-HER2**, **H3(StBu)**, **H1(StBu)-IgG**, **H2(StBu)$_2$-IgG** and **H2a(StBu)-Her2**.

a. General procedure

Syntheses were performed using microwave-assisted SPPS and the standard Fmoc/*tert*-butyl strategy on a Rink-amide-ChemMatrix resin (0.51 mmol/g loading, 0.1 mmol scale). ChemMatrix resins are chemically stable and hydrophilic, combining the advantages of

polystyrene-based and PEG-grafted resins. They can be used in different polar solvents due to their high polarity. In particular, their high swelling properties have to be taken into consideration when performing peptide synthesis as the volume occupied by the resin in the reaction vessel is much higher than for standard resins and extensive washings are needed after each step. ChemMatrix resins have been shown to greatly improve the purity of long peptides, e.g. allowing the efficient synthesis of HIV-1 protease[135] (99 amino acids) and polypeptide Aβ[136] (42 amino acids).

For the synthesis of the affibody helix peptides ranging from 18 to 23 amino acid residues, we added DMF and DCM washings after each coupling and deprotection step to remove the residual reagents trapped in the resin beads. Moreover, while performing the synthesis at 0.1 mmol scale, we set the microwave synthesizer parameters at 0.25 mmol. As a result, we increased the reaction volumes while keeping the different reagents at the same concentration (standard 0.1 mmol scale volumes were not sufficient due to the high swelling properties of this resin).

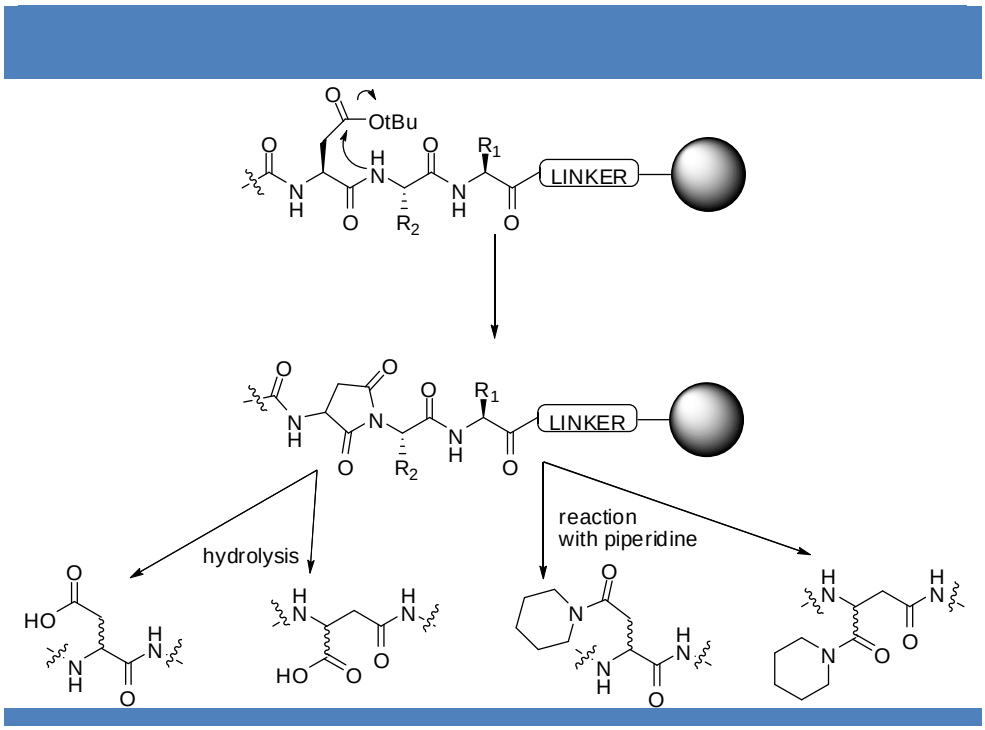
We generally performed the peptide couplings at 70 ºC and 35 W (microwave power) for 5 min using ten-fold molar excess of each Fmoc-aa-OH (5 mL of a 0.2 M solution in DMF), HBTU (2 mL of a 0.5 M solution in DMF) and DIEA (1 mL of a 2 M solution in NMP). This standard coupling procedure was slightly different in the case of few amino acids. Double couplings were performed for Fmoc-Arg(Pbf)-OH due to the steric hinderance of the Pbf protecting group (2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl). In the case of residues sensitive to racemization, milder conditions were used. Fmoc-His(Trt)-OH was coupled for 2 min without microwave heating and then at 50 ºC with 25 W for 4 min, the insertion of the N-methyl-cysteine moiety is extensively described below.

b. Fmoc group deprotection

Fmoc groups were deprotected with 2 successive treatments with a solution of 5 %v/v piperazine in DMF (10 mL) containing 0.1 M HOBt at 70 ºC and 55 W for 3 min. This deprotection cocktail was preferred to the more standard 20 %v/v piperidine in DMF (used in the syntheses described in chapters 2 and 3) as it has been shown[67] to reduce aspartimide formation, together with racemization. During Fmoc deprotection of peptides containing Asp,

135. Frutos, S.; Tulla-Puche, J.; Albericio, F.; Giralt, E. Chemical Synthesis of 19F-Labeled HIV-1 Protease Using Fmoc-Chemistry and ChemMatrix Resin. *Int. J. Pept. Res. Ther.* **2007**, *13*, 221–227.
136. Bacsa, B.; Bősze, S.; Kappe, C. O. Direct Solid-Phase Synthesis of the B-Amyloid (1−42) Peptide Using Controlled Microwave Heating. *J. Org. Chem.* **2010**, *75*, 2103–2106.

aspartimide formation can occur, in particular in the presence of the "Asp-aa" sequence where aa = Gly, Asn, Ser or Ala. The nitrogen atom attached to the α-carboxy group attacks the aspartic acid side chain ester forming a five member ring that can easily react with a nucleophile, i.e. piperidine or be hydrolyzed (Scheme 19).



Scheme 19 | Mechanism of aspartimide formation and successive hydrolysis or reaction with piperidine.

The addition of 0.1 M HOBt to the standard 20 %v/v piperidine solution in DMF has been shown to reduce aspartimide formation, but better results were achieved after substitution of piperidine with piperazine. In addition, this deprotection mixture has been shown to reduce racemization. This base-induced side reaction has also been shown to occur during the coupling step and it's particularly common in peptides containing a C-terminal cysteine. Due to the presence of **Daa2** (often in C-terminal position), His and Asp in the affibody helix peptides, the solution of piperazine in DMF (5 %v/v containing 0.1 M HOBt) was used to achieve Fmoc deprotection for this set of peptides.

c. Insertion of N-methyl-cysteine

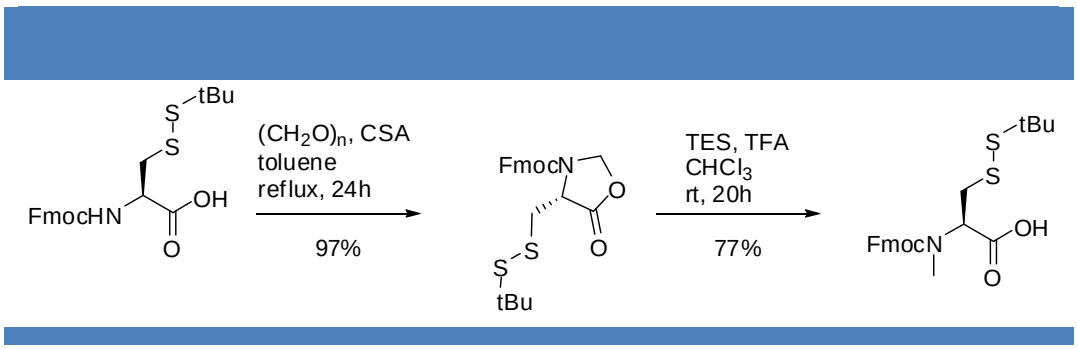
i. First approach: In situ N-methylation

We initially performed the synthesis of **H3** using the same protocol described in Chapter 3 for the N-methylation of peptides in solid phase. This three step protocol consists in activating

the N-terminal cysteine group as NBS, performing N-methylation with methyl iodide and TBAF and finally removing the NBS group by treatment with a thiol and DBU. However, the purity of the final peptides was not satisfactory. UPLC analysis of a small portion of the resin (mini-cleavage) at regular intervals during the synthesis showed an increased amount of side products after the N-methylation steps. Considering that all the target peptides incorporate one or two (in the case of **H2-Her2** and **H2-IgG**) N-methyl-cysteine residues, we investigated a different synthetic strategy.

ii. Second approach: Coupling of N-methyl-cysteine

In alternative to the in-situ methylation, a second method for the synthesis of N-methylated peptides, as mentioned in Chapter 3, is the insertion during peptide synthesis of an already N-methylated amino acid as a building block. Many of these Fmoc-protected N-methyl amino acids are commercially available or can generally be synthesized in few steps from cheaper standard precursors. In the case of certain amino acids, including cysteine, the synthesis of N-methylated analogues can be more challenging and not many strategies are described in the literature. For this reason we selected a two step procedure reported by Ruggles *et al.*[137] for the N-methylation of Fmoc-Cys(StBu)-OH to afford Fmoc-NMeCys(StBu)-OH in good yields (Scheme 20). The authors tested the oxazolidinone method for the synthesis of Fmoc-N-methyl-cysteine bearing different protecting groups on the side chains. They found that the StBu group allowed for the synthesis of Fmoc-NMeCys(StBu)-OH with high yields and minimal formation of side products, i.e. the thioazolidine derivative.



Scheme 20 | Two-step synthesis of Fmoc-Cys(StBu)-OH, as reported in ref. 137.

Fmoc-NMeCys(StBu)-OH was easily synthesized; however, the presence of the StBu protecting group had to be taken into consideration. This protecting group is not cleaved in acidic conditions, unlike the more common Trt group. We synthesized our peptides in their

137. Ruggles, E. L.; Flemer, S.; Hondal, R. J. A Viable Synthesis of N-Methyl Cysteine. *Biopolymers* **2008**, *90*, 61–68.

StBu-protected form as we considered that the reducing conditions (DTT) used in our exchange reactions were sufficient to obtain in-situ removal of this protecting group.

In contrast with our first approach, our newly designed synthetic scheme allowed for the synthesis of the peptides in higher purity. The two chromatograms shown in Figure 44 correspond to the crude peptides obtained with the two different approaches reported above.

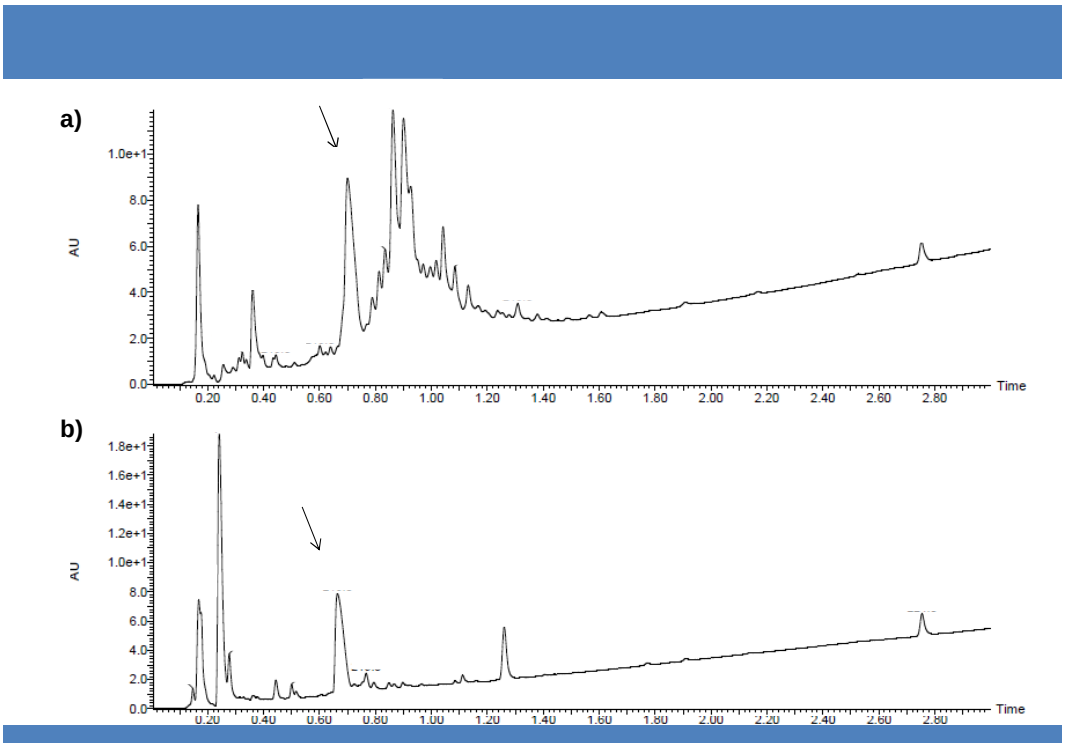


Figure 44 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of the crude product **H3(StBu)** synthesized with **a)** approach *i* (*in situ* N-methylation) or **b)** approach *ii* (coupling of N-methyl cysteine). The peak corresponding to **H3(StBu)** is indicated by the arrow.

d. Glycine coupling: Acylation of N-methyl cysteine

As already described in Chapter 2 and 3, the acylation of N-alkylated amino acids is challenging and generally requires harsh conditions. For the synthesis of these peptides a concentrated solution of Fmoc-Gly-OH (10 eq.), HATU (10 eq.) and DIEA (11 eq.) in DMF (4 mL) was prepared and added to the microwave reaction vessel; the coupling reaction was then performed at 75 °C for 20 min at 30 W and repeated a second time with a fresh solution.

e. Simultaneous cleavage from the resin and side-chain deprotection

In the case of the affibody helix peptides, the resin was suspended in a freshly prepared 5 mL solution of TFA/phenol/H_2O/TIS: 88/5/5/2 vol%. After 2 h, the resin was filtered and

washed with TFA (5mL). The combined filtrates were concentrated in vacuo and then added to cold diethylether (200 mL). The resulting precipitate was then separated by centrifugation, washed with cold diethylether (2 x 40 mL) and dried under argon to afford crude peptides.

During the synthesis, a mini-cleavage was performed at regular intervals to verify that the reactions had gone to completion. A minimum amount of the resin was cleaved with a 1 mL solution of TFA/phenol/H_2O/TIS: 88/5/5/2 vol%. After 30 min the solvent was then reduced under vacuo and the crude peptide precipitated in diethyl ether. The precipitate was then isolated by centrifugation, dissolved in acidic water (0.1 %v/v formic acid) and/or methanol and analyzed via UPLC-ESI analysis.

3. Exchange reactions among H1-HER2, H2-HER2 and H3 or H1-IgG, H2-IgG and H3

We investigated the exchange reaction among the three affibody helices specific for HER2 or for IgG, represented in Figure 45. We aimed to observe the *in situ* formation of the **H1-H2-H3** trimer and its amplification in the presence of the target protein; however other exchange products are involved in the exchange reaction and they will be further discussed later in this chapter.

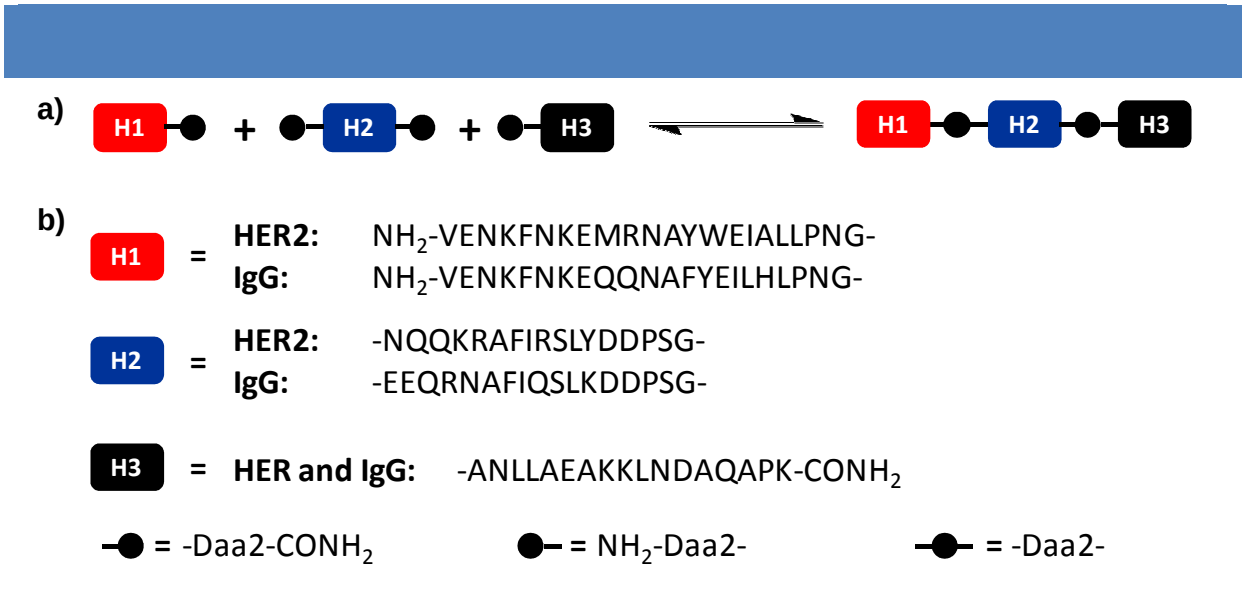


Figure 45 | a) Exchange reaction among the three affibody helices, showing only the desired H1-H2-H3 trimer for simplicity; **b)** explanation of the symbols used in **a)**.

a. Exchange reaction setup

The exchange reactions were run in 0.2 M phosphate or ammonium bicarbonate buffers prepared at pH 7.4. In order to prevent oxidation, these buffers were degassed using freeze/thaw cycles and placed under argon prior to use. Solutions of each of the three helices were mixed together in a vial, frozen and lyophilized. This lyophilized powder containing the peptides (final concentration ranging from 50 µM to 1 mM) was then diluted under argon with the degassed buffer containing DTT (1.3 to 625 mM). The exchange reactions were then left stirring in a vial placed into a Schlenk tube equipped with a Young PTFE valve and containing a 250 mM DTT solution in degassed buffer at the bottom. This solution acts as scavenger to prevent traces of oxygen to oxidize the species present in the reaction mixture (Figure 46d).

Initial attempts at performing the exchange reaction using a standard Schlenk tube under a positive pressure of argon (Figure 46b) were unsuccessful due to rapid oxidation of the mixture (Figure 46a). In order for the exchange reaction to occur, DTT needs to be in its reduced form to remove the StBu protecting groups and to keep the cysteine thiols reduced. As shown in Figure 46c, with this second setup only a small amount of DTT was oxidized after 4 days of equilibration.

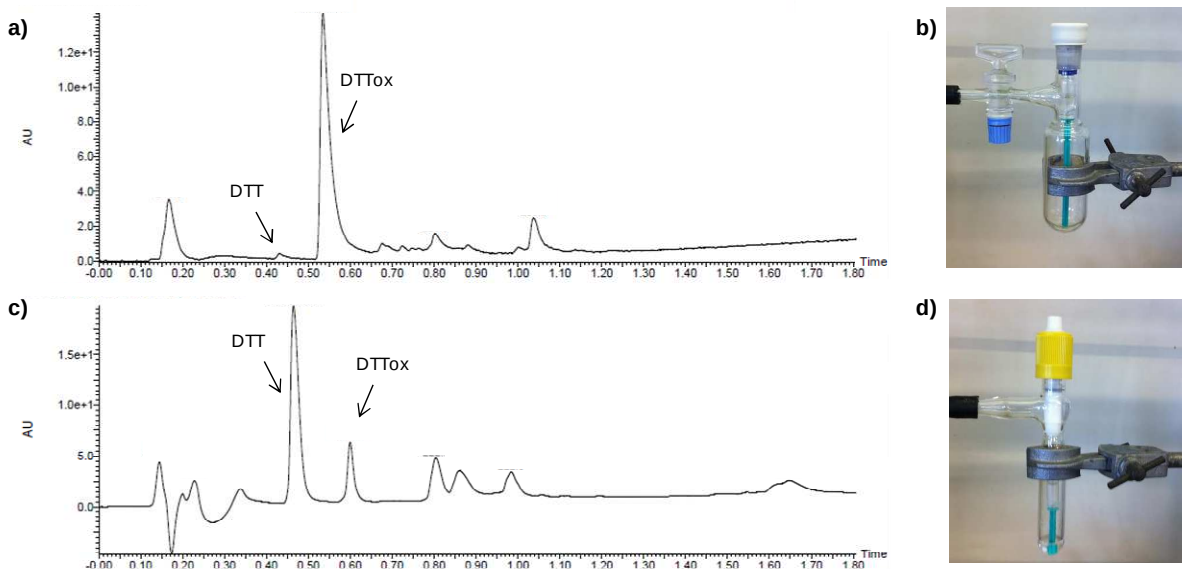


Figure 46 | a) Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of an exchange reaction among the affibody helices based on Z_{HER2} after 12 h of equilibration in **b)** the first setup. **c)** Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of an exchange reaction among the affibody helices based on Z_{HER2} after 4 days of equilibration in **d)** the second setup.

b. Exchange reaction analysis

i. UPLC-ESI analysis

Studying the exchange reaction among affibody helices in the presence of the specific proteic target and demonstrating the amplification of its best binder were the objectives of this project. We envisaged to study the system with a 1:1 affibody helices/protein ratio, which led to high dilutions of the exchange mixtures in order to avoid protein aggregation.

The UPLC-ESI that we have in our laboratories was crucial for the investigation of the exchange reaction between model peptides incorporating either **Daa1** or **Daa2** and described in the Chapters 2 and 3. However, in the case of the exchange reaction among affibody helices, the UV and ESI-MS detectors were not sensitive enough to detect low amounts of the exchange products. In addition to the **H1-H2-H3** trimer, possible exchange products include the four dimers **H1-H2**, **H2-H3**, **H1-H3** and **H2-H2**, **H2** cyclized, **H2-H2** cyclized, different trimers and tetramers. Moreover, each peptide incorporating a C-terminal **Daa2** can undergo transthioesterification with DTT, forming the corresponding thioester that can in turn be hydrolysed (Figure 47e).

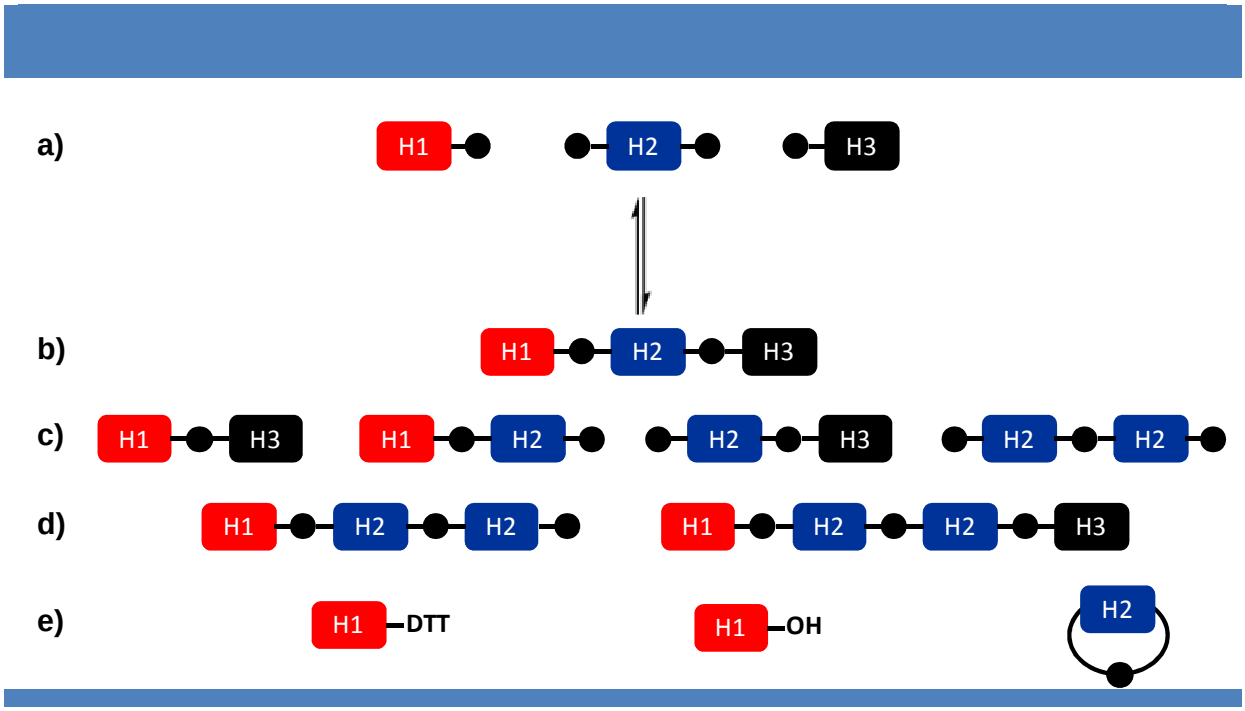


Figure 47 | Schematic representation of the exchange reaction among **a)** affibody helices 1, 2 and 3 can lead to the formation of **b)** the **H1-H2-H3** trimer but also to several **c)** dimers, **d)** different trimers or tetramers, **e)** DTT thioesters, hydrolysis and cyclization products. The aim of this figure is to give a general overview of the complexity of the exchange reaction mixture, other exchange products, DTT thioesters and hydrolysis products can be formed but have not been included.

At peptide concentrations of 50 µM, the only detectable species were DTT, the starting peptides (initially StBu protected) and the cyclization product of **H2**. **H2-HER2** and **H2-IgG** are the only starting peptides incorporating two dynamic units and they can therefore undergo intramolecular cyclization resulting in the loss of one N-methyl-cysteine moiety. We had to increase to peptide concentration to 1 mM in order to observe other exchange products, including the **H1-H2-H3** trimer (Figure 48).

In addition, the UPLC-ESI mass detector has a molecular weight cut-off of 2 kDa, that complicates the detection of the exchange peptides. With molecular weights around 4 and 6.5 kDa, dimers and trimers were observed as $[M+2H]^{2+}$ to $[M+6H]^{6+}$ ions.

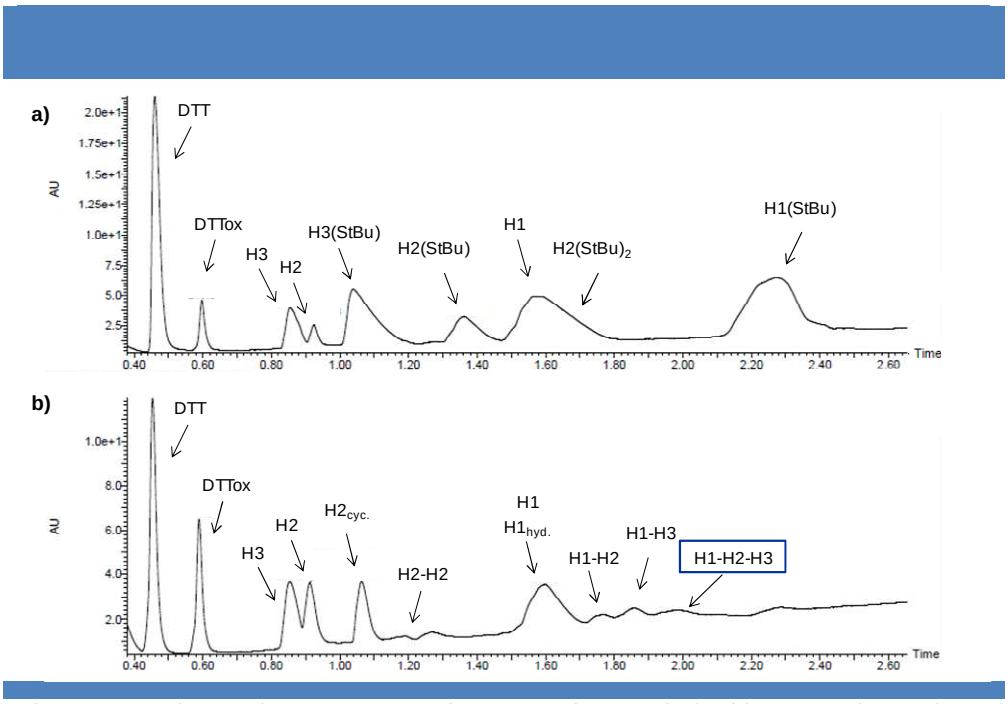


Figure 48 | Reverse phase chromatograms (UPLC-ESI) recorded with UV and MS detection of an exchange reaction among the affibody helices based on Z_{IgG} in a 0.2 M ammonium bicarbonate buffer at pH 7.4 (peptide concentration 1 mM, DTT concentration 25 mM) after **a)** 1 h and **b)** 7 days of equilibration. cyc. = cyclized product; hyd. = hydrolised product

ii. MALDI-TOF analysis

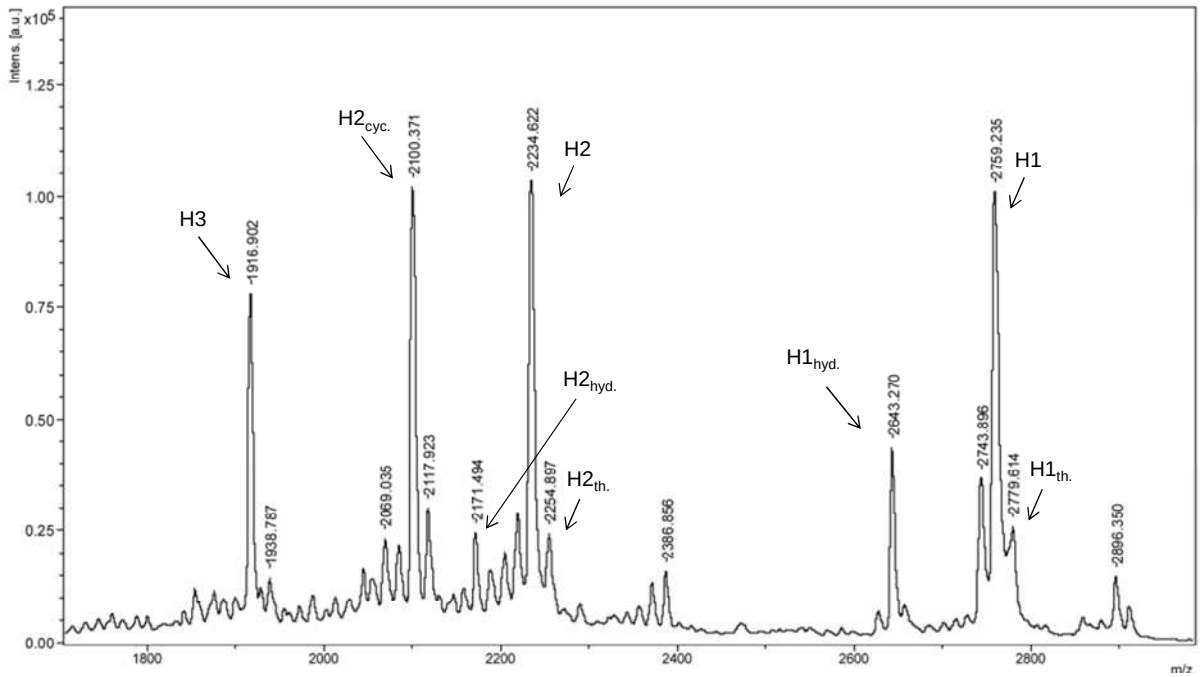
The need for an improvement in the analysis of the exchange mixture led us to initiate a collaboration with Dr. Jean-Marc Strub from the LSMBO (Laboratoire de Spectrométrie de Masse BioOrganique) in Strasbourg, France.

We initially analyzed our exchange mixtures via matrix-assisted laser desorption/ionization time-of-flight analysis (MALDI-TOF). Species with high molecular weights can be detected by this relatively soft ionization technique. In addition, its high sensitivity allows one to overcome the necessity for concentrated samples (i.e., as for

UPLC-ESI).

In our case, a 30 μL aliquot of the exchange reaction mixture was diluted with 20 μL of a formic acid solution (0.5 %v/v in water) in order to prevent oxidation. 1 μL of this mixture was then diluted 1000 fold with a 0.1 %v/v TFA solution. Only 1 μL of this final solution was then analyzed by MALDI. It was co-deposed on a metal plate with a solution of the matrix, generally α-cyano-4-hydroxycinnamic acid (CHCA) or 2,5-dihydroxybenzoic acid (DHB). A laser pulse was then applied under high vacuum conditions and the resulting ionized species detected according to their m/z ratio. As salts can interfere with the matrix recrystallization, MALDI analysis is very sensitive to their concentration in the sample; in particular, it is not compatible with phosphate buffers[138], that we routinely used for our exchange reactions. We therefore substituted it with an ammonium bicarbonate buffer.

We initially analyzed a library based on Z_{HER2} (peptide concentration 333 μM, DTT concentration 125 mM) in a 0.2 M ammonium bicarbonate buffer and we could characterize different species involved in the exchange reaction (Figure 49). However, after 3 days of equilibration, we did not detect any trace of the helix1-helix2-helix3 trimer.



138. Kallweit, U.; Börnsen, K. O.; Kresbach, G. M.; Widmer, H. M. Matrix Compatible Buffers for Analysis of Proteins with Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 845–850.

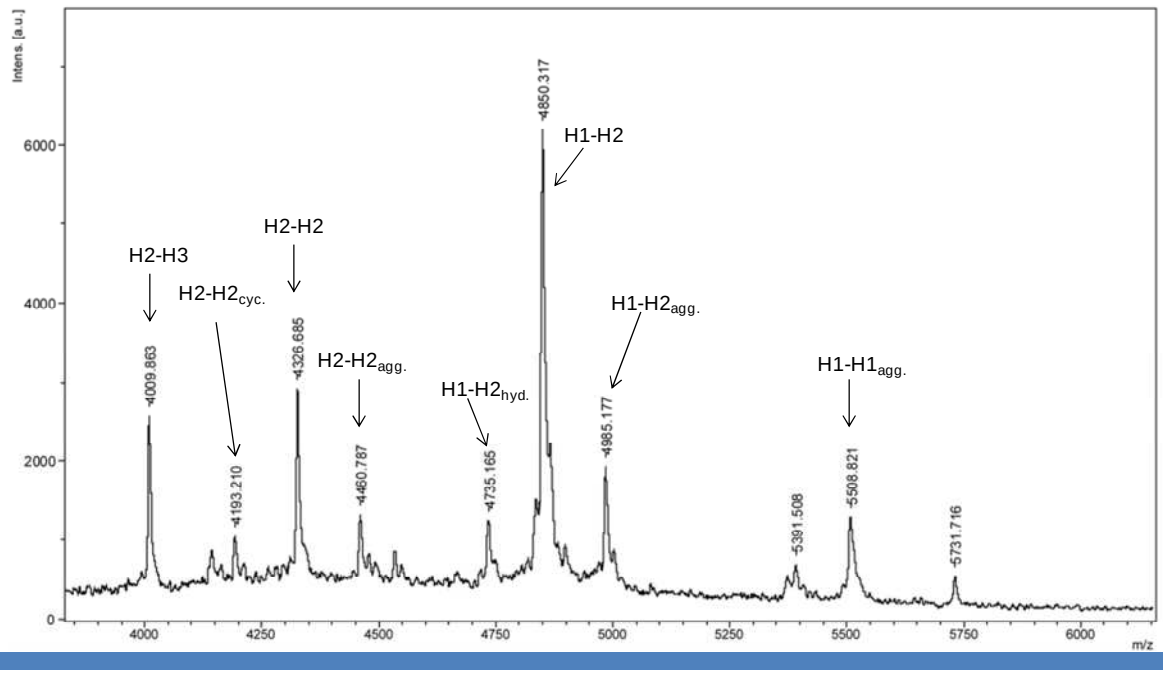


Figure 49 | MALDI-TOF spectra of the analyzed exchange reaction using a CHCA matrix. The molecular weights correspond to the masses of the peptides with a +7 increment, probably due to a problem during the instrument calibration. th. = thioester product; agg. = peptide aggregate.

Although promising, the intensity of the peaks corresponding to the exchange peptide dimers was one order of magnitude lower than the one of the starting peptides. In addition, we observed problems of reproducibility. We performed MALDI-TOF analysis on libraries based on either Z_{HER2} or Z_{IgG} at different peptide and DTT concentrations and we observed either no formation of dimers and trimers, or trace amounts of them.

In order to improve the detection of the exchange products present in lower amounts than the starting peptides, we decided to fractionate the exchange mixture prior to MALDI-TOF analysis. To do so, we injected an aliquot (30 µL) of the reaction mixture in the UPLC-ESI, manually collected up to 9 different fractions in the 5 minute chromatography and individually analyzed them by MALDI-TOF. In addition, this procedure also allowed us to monitor the exchange reactions run in a phosphate buffer.

Unfortunately, MALDI-TOF analysis only showed the original starting peptides together with thioester and hydrolysis products and cyclized helix 2. It is possible that dimer and trimer exchange products were lost during the filtration procedure necessary before injection in the UPLC-ESI. This eventuality was further investigated with another analytical technique and will be discussed later in this chapter.

iii. UPLC-ESI-TOF

Although an improvement if compared to UPLC-ESI, MALDI-TOF analysis didn't meet

our expectations. Thanks to the collaboration with Dr. Jean-Marc Strub, we had access to a second instrument: a *nanoAcquity* UPLC coupled to an ESI-TOF detector. Its 25 cm column provides for a better separation of the peptide mixture and its *SYNAPT* detector for a more accurate analysis of the species present in the reaction mixture. A 1 µL aliquot of the reaction mixture was diluted with 100 µL of a 0.1 %v/v TFA solution in water and just 2 µL of the resulting solution were injected in the UPLC-ESI-TOF.

We initially analyzed a library based on Z_{IgG} (peptide concentration 333 µM, DTT concentration 625 mM) in a 0.2 M ammonium bicarbonate buffer. After 10 days of equilibration we could characterize the four possible dimers, cyclization products (of **H2** and **H2-H2** dimer), thioester and hydrolysis products and the **H1-H2-H3** trimer. Importantly, the state of charge of each peak was well defined, allowing us to inequivocably attribute the peaks to the different species, even in the case of peptides with high molecular weights and complex ionization patterns. Figure 50 shows only the peaks corresponding to the **H1-H2-H3** trimer; however, all the peptides present in the exchange mixture were characterized in the same fashion.

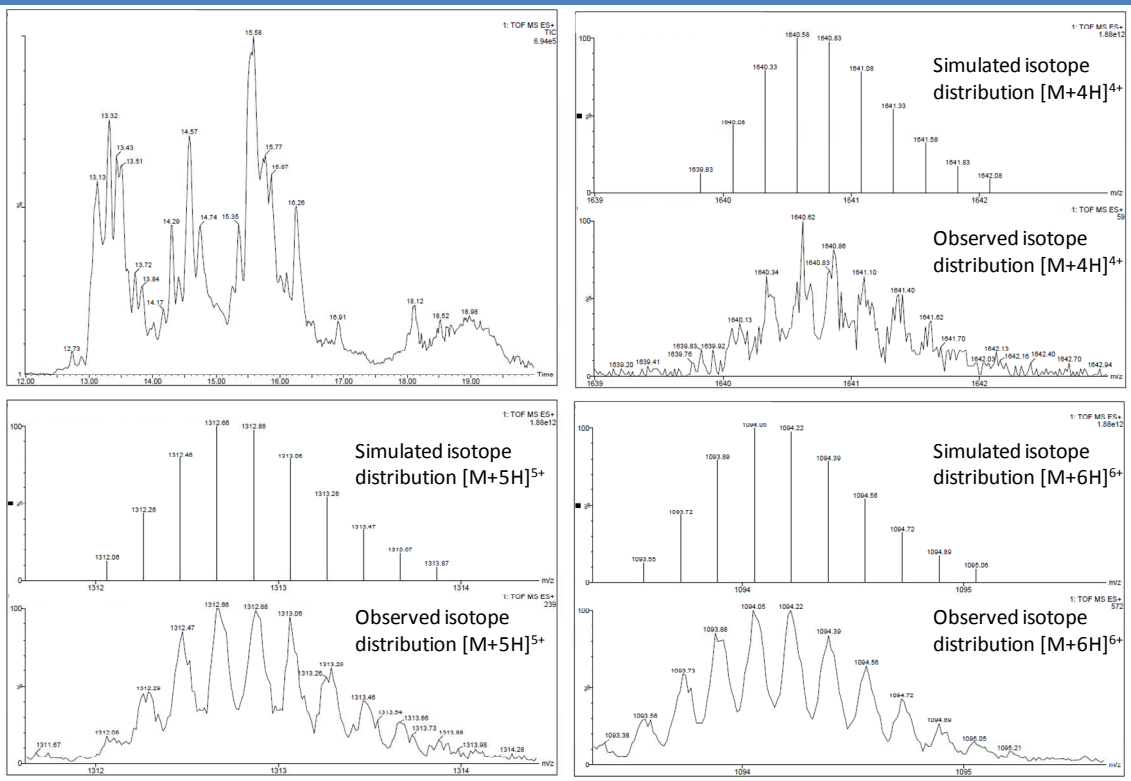


Figure 50 | Reverse phase TIC chromatogram (UPLC-ESI-TOF) of the exchange reaction among the affibody helices based on Z_{IgG} in a 0.2 M ammonium bicarbonate buffer at pH 7.4 (peptide concentration 333 µM, DTT concentration 625 mM) after 10 days of equilibration and isotope distribution corresponding to the **H1-H2-H3** trimer.

We then analyzed exchange reactions run in different conditions with the same instrument. Interestingly, in the case of a library based on Z_{IgG} with a peptide concentration of 1 mM (DTT concentration 25 mM, in a 0.2 M ammonium bicarbonate buffer) we could also characterize the **H1-H2-H2** trimer and the **H1-H2-H2-H3** tetramer.

c. Influence of the filtration on the library

As already mentioned earlier in this Chapter, filtering was mandatory prior to injection of exchange reaction aliquots in the UPLC-ESI. In the case of UPLC-ESI-TOF, high dilutions and the presence of a pre-column made this step superfluous. We could thus investigate the influence of the filtration step on the exchange mixture composition detected.

We analyzed a library based on Z_{IgG} (peptide concentration 333 μM, DTT concentration 125 mM) in a 0.2 M ammonium bicarbonate buffer. After dilution of a 1 μL aliquot with 100 μL of a 0.1 %v/v TFA solution, 2 μL of the resulting mixture were either injected as such or after filtration through the Corning pipette tips. Indeed, the resulting chromatograms were not superimposable (Figure 51a) and the m/z distribution in the area corresponding to the elution of dimers and trimers (t_r = 32-38 min) revealed the presence of only traces of different exchange products (Figure 51b) in the case of the filtered sample.

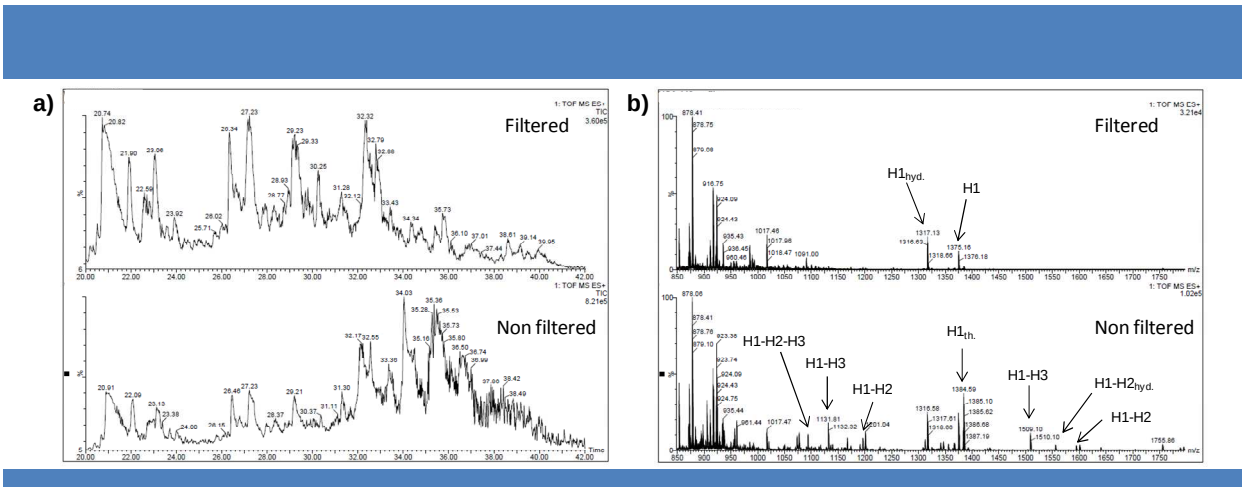


Figure 51 | a) Reverse phase TIC chromatograms (UPLC-ESI-TOF) of the exchange reaction among the affibody helices based on Z_{IgG} in a 0.2 M ammonium bicarbonate buffer at Ph 7.4 (peptide concentration 333 μM, DTT concentration 125 mM) after 5 days of equilibration; the analysis was run either on a filtered or a non-filtered sample. **b)** MS of the area corresponding to t_r = 32-38 min, showing the loss of different species after filtration.

This experiment might explain the reason why we could not observe dimers and trimers when analyzing fractions of the exchange mixture via MALDI-TOF. The sample aliquot was

filtered prior to injection in the UPLC-ESI and the filter retained different exchange peptides. It is also probable that part of the starting peptides were retained but not completely, due to their relatively high concentration.

d. Influence of the exchange reaction medium on the library

We then investigated whether the buffer composition influenced the rate or the extent of the exchange reaction. To do so, we performed four exchange reactions (peptide concentration 50 μM, DTT concentration 25 mM) in parallel among affibody helices based on Z_{IgG}. Two reactions were run in a phosphate buffer and the other two in an ammonium bicarbonate buffer. Moreover, for each of these reaction couples, in one case 150 mM NaCl and 0.02 %w/v NaN_3 were added to the buffer.

At this stage, we had a large amount of data concerning exchange reactions run in either the phosphate or the bicarbonate buffer (introduced to facilitate MALDI-TOF analysis). We wanted to verify if these data were superimposable. Moreover, with the objective of studying the system in the presence of the protein target, we investigated the influence of the addition of NaCl (150 mM) and NaN_3 (0.02 %w/v) on the rate and extent of the exchange reaction. While NaCl is often recommended in protein reconstitution protocols, NaN_3 is a commonly used anti-microbial agent to prevent protein (and peptide) degradation.

Interestingly, we did not observe any significant difference among the four different conditions used. We decided to continue our investigations in a phosphate buffer containing NaCl (150 mM) and NaN_3 (0.02 %w/v).

e. StBu group deprotection and exchange reaction kinetics

Another step in the understanding of our system was the investigation of the kinetics involved. In order for the exchange reaction to occur, the StBu protecting groups on the initial peptides need to be removed.

Two different reactions take place: the initial StBu deprotection and the exchange reaction among the deprotected species. Careful monitoring of different libraries based on Z_{IgG}, gave us insights into their kinetics.

We initially analyzed a library based on Z_{IgG} (peptide concentration 50 μM, DTT concentration 25 mM) in a 0.2 M phosphate buffer containing 150 mM NaCl and 0.02 %w/v

NaN_3. Interestingly, and although complete removal of the StBu group was observed after 12-24 h for **H1** and **H3** and 33 h for **H2**, after only 45 min we could already detect all the four possible dimers and cyclized **H2** and **H2-H2**. However, for the peak corresponding to the **H1-H2-H3** trimer we had to wait 49 h (Figure 52).

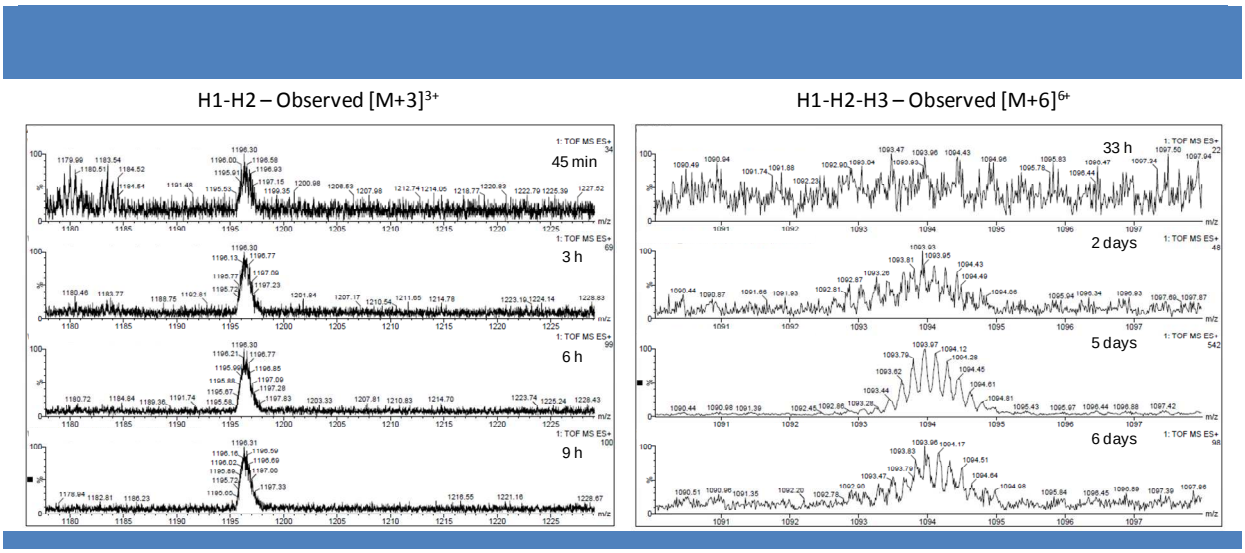


Figure 52 | MS peaks corresponding to the exchange products **H1-H2** and **H1-H2-H3** over time of an exchange reaction among the affibody helices based on Z_{IgG} in a 0.2 M phosphate buffer containing NaCl (150 mM) and NaN_3 (0.02 %w/v) at pH 7.4 (peptide concentration 52 μM, DTT concentration 25 mM).

This experiment demonstrated that equilibration of the exchange mixture starts as soon as a little amount of deprotected starting peptides is present in solution. Unfortunately, it also shows the slow rate of the StBu group deprotection.

4. Exchange reaction between H1-Her2 and H2a-Her2

The initial investigation of the exchange reaction among the three helices of Z_{HER2} showed the rapid propensity of helix 2 to cyclize. Although theoretically able to undergo exchange reactions with other species in equilibrium, this cyclic peptide is likely to act as a spectator in the library. As a result, the amount of helix 2 actually participating in the exchange reaction is diminished. To address this issue, we investigated the exchange between **H1-HER2** and **H2a-HER2**, an analogue of **H2-HER2** containing only one N-methyl-cysteine moiety, in N-terminal position.

While helix3 provides to the overall stability of the affibody architecture[139], helices 1 and

139. Honarvar, H.; Jokilaakso, N.; Andersson, K.; Malmberg, J.; Rosik, D.; Orlova, A.; Karlström, A. E.; Tolmachev, V.; Järver, P. Evaluation of Backbone-Cyclized HER2-Binding 2-Helix Affibody Molecule for in Vivo Molecular Imaging. *Nucl. Med. Biol.* **2013**, *40*, 378–386.

2 are directly involved in the recognition process. One can thus hypothesize that the helix1-helix2 dimer would probably retain a certain degree of binding affinity to HER2.

The exchange reaction between **H1-HER2** and **H2a-HER2** was performed in a bicarbonate buffer (200 mM, peptide concentration 333 μM, DTT concentration 125 mM) and monitored by UPLC-ESI-TOF that showed the formation of the helix1-helix2 dimer (Figure 53).

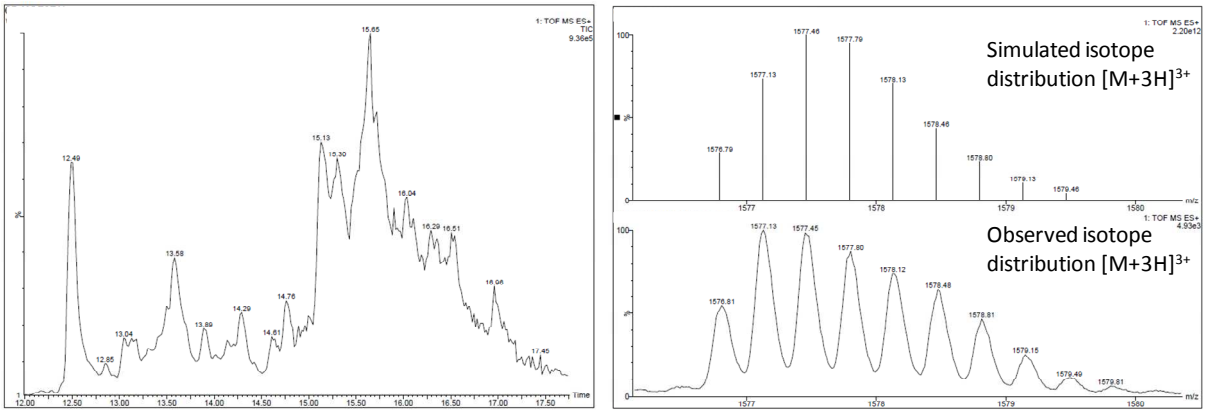


Figure 53 | Reverse phase TIC chromatogram (UPLC-ESI-TOF) of the exchange reaction between **H1-HER2** and **H2a-HER2** in a 0.2 M ammonium bicarbonate buffer at pH 7.4 (peptide concentration 333 μM, DTT concentration 125 mM) after 9 days of equilibration and isotope distribution corresponding to the exchange product **H1-H2a**.

5. Exchange reaction in the presence of the protein: preliminary studies

We realized that both the target proteins of Z_{HER2} and Z_{IgG} are sensitive to the reduction conditions (DTT) used to promote the exchange reaction. The extracellular domain of the HER2 protein contains multiple disulfide bonds[140] and the whole structure of antibodies, including IgG[141], is based on disulfide bonds. In order to overcome this issue, we considered different options.

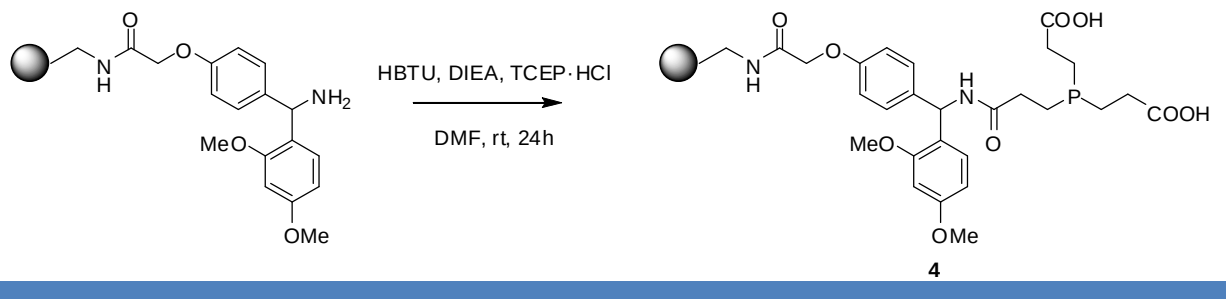
Initially we hypothesized that a minimum amount of DTT would have been sufficient to promote the exchange reaction without being detrimental for the protein. However, Singh and Whitesides reported[142] that a DTT concentration of 4.8 mM could completely reduce IgG in 1

140. Rawale, S. V.; Kaumaya, P. T. P. Design, Synthesis and Analysis of Constrained B-Cell Epitope from HER-2 Protein. In *Understanding Biology Using Peptides*; Blondelle, S. E., Ed.; Springer New York: New York, NY, 2006; pp. 124–125.
141. Liu, H.; May, K. Disulfide Bond Structures of IgG Molecules: Structural Variations, Chemical Modifications and Possible Impacts to Stability and Biological Function. *MAbs 4*, 17–23.
142. Singh, R.; Whitesides, G. M. Reagents for Rapid Reduction of Native Disulfide Bonds in Proteins. *Bioorg. Chem.* **1994**, *22*, 109–115.

h. Moreover, when we analyzed libraries based on Z_{IgG} (peptide concentration 50 μM in a 0.2 M phosphate buffer containing 150 mM NaCl and 0.02 %w/v NaN_3) in the presence of reduced concentrations of DTT such as 5 mM or 1.3 mM, we could not detect any **H1-H2-H3** trimer after 5 days of equilibration. Therefore, this idea was quickly abandoned.

We then considered performing the two reactions separately: the StBu removal first and, in a second moment, the exchange reaction. The experiment design consisted in removing the StBu protecting group with a resin-bound reducing agent and then transferring the solution of the deprotected peptides on a second resin, capable of promoting the exchange reaction.

To do so, we coupled TCEP on a ChemMatrix resin, following a procedure reported by Subra and coworkers[143] and shown in Scheme 21.



Scheme 21 | Synthesis of resin-bound TCEP.

ATR-IR analysis of the resulting resin showed vibration bands corresponding to the newly formed amide linkage (1655 cm^{-1}) and the carboxylic acids (1731 cm^{-1}). TCEP loading can indeed be determined by ICP-AES; however, we have not performed this analysis at this preliminary stage of investigation. We tested the reducing power of this TCEP-bound resin on a library based on Z_{IgG} (peptide concentration 200 μM) in a 0.2 M phosphate buffer containing 150 mM NaCl and 0.02 %w/v NaN_3. Since metal cations are known to reduce the TCEP activity, we improved the initial results by adding EDTA (20 mM) to the buffer. Using 35 mg of TCEP-bound resin in these conditions, after 2 h we observed complete removal of the StBu group from the starting peptides and only traces of the corresponding desulfurized products. We then transferred this solution to a second vial containing a mercaptoalkyl-PEG200-Hypogel resin, as we expected the terminal thiol group on this resin to promote the exchange reaction. However, after 5 days of equilibration we could not observe any trace of exchange products.

We could indeed overcome this issue by using a different resin-supported thiol. For

143. Miralles, G.; Verdié, P.; Puget, K.; Maurras, A.; Martinez, J.; Subra, G. Microwave-Mediated Reduction of Disulfide Bridges with Supported (tris(2-Carboxyethyl)phosphine) as Resin-Bound Reducing Agent. *ACS Comb. Sci.* **2013**, *15*, 169–173.

instance, we could envisage to immobilize MPAA, the typical NCL catalyst, on a solid support. However, the in-situ removal of StBu groups leads to the formation of *tert*-butyl thiol, that could be detrimental for the protein stability in the medium.

The best approach to investigate the system in the presence of a protein target probably consists in running the exchange reaction among peptides with free N-methyl-cysteine thiol groups. The slow rate of StBu deprotection and the possible incompatibility of the resulting *tert*-butyl thiol with the protein targets, are indeed major hurdles. Different thiols could be covalently attached on a resin to promote the exchange reaction and the protein either kept in a separate compartment by a semipermeable membrane or, in the case of IgG, immobilized on a second resin. However, these are only speculations as there was no time for me to perform the synthesis of the deprotected affibody helices.

Chapter 5. Dynamic Glycopeptides: Toward Biological Applications of Reverse NCL

1. Introduction: Protein glycosylation

Carbohydrates are important biomolecules that, unlike proteins and nucleic acids, are able of forming many different structures. The multiple functional groups (mostly hydroxyl groups) present on monosaccharides allow for the formation of a plethora of different carbohydrate structures, including branched ones, from a relatively small number of sugar units.

Besides their fundamental role as energy sources and structural materials, carbohydrates are involved in a variety of biological functions. In particular, when conjugated to lipids or proteins and exposed on the cell surface, they play a crucial role in various recognition processes, intercellular communication and transduction events (Figure 54)[144].

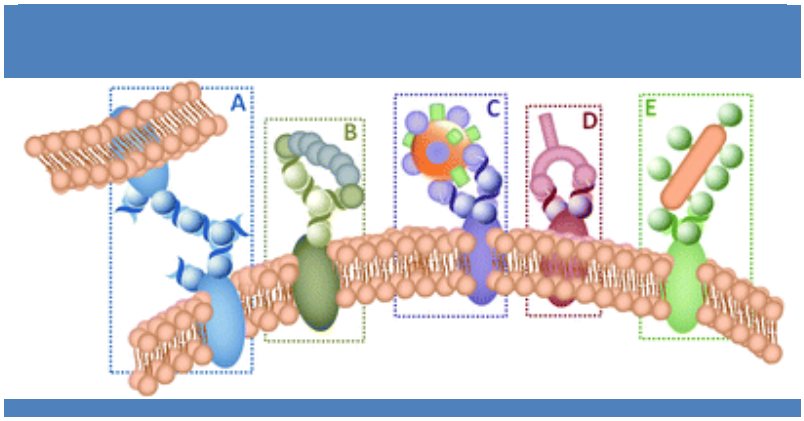


Figure 54 | Participation of cell surface carbohydrates in recognition events with **a)** another cell, **b)** toxins, **c)** viruses, **d)** antibodies and **e)** bacteria (from ref. 144).

Glycoprotein biosynthesis by the cellular enzymatic machinery is a complex phenomena and many elaborate glycosylation routes have been identified. Indeed, the attachment of sugar residues is considered one of the most complicated post-translational modifications that a protein can undergo[145]. Thirteen different monosaccharides and eight amino acids are involved in the carbohydrate-peptide linkage, giving rise to a variety of possible combinations. Despite other bonds exist, naturally occurring glycopeptides most commonly incorporate N- and O-glycosidic linkages. Interestingly, S-linked glycoproteins have been extensively investigated as analogues of O- or N-linked glycoproteins. Their increased enzymatic resistance and chemical

144. Pashkuleva, I.; Reis, R. L. Sugars: Burden or Biomaterials of the Future? *J. Mater. Chem.* **2010**, *20*, 8803.
145. Spiro, R. G. Protein Glycosylation: Nature, Distribution, Enzymatic Formation, and Disease Implications of Glycopeptide Bonds. *Glycobiology* **2002**, *12*, 43R – 56R.

stability[146], together with their high tolerability, makes them attractive synthetic targets for the development of therapeutic agents.

As already mentioned above, carbohydrate-protein interaction has a critical role in many biological events; however, most saccharides bind their protein receptors only weakly, with association constants seldom beyond 10^6 M^{-1}. Nature has circumvented this low affinity through multivalency. Lectins are natural carbohydrate-binding proteins, specific for different mono- or polysaccharides and are aggregated into ordered oligomeric structures that simultaneously bind several saccharide units. These multivalent interactions are referred to as the "cluster glycoside effect"[147] and can determine an enhancement of the binding affinity of several orders of magnitude.

In glycoproteins, only a few carbohydrate residues of their complex oligosaccharide architectures are directly involved in the recognition process. The role of the other residues is limited to that of a structural matrix. Based on this observation, glycodendrimers[148] were first developed, followed by a number of multivalent scaffolds based on very different architectures[149]. The highest contribution to the binding enhancement is generally attributed to the so-called chelate effect, strictly correlated to the distance among the binding sites. In this case, the nature of the spacer between the multivalent scaffold and the carbohydrate moieties plays an important role. Another mechanism for binding enhancement is statistical rebinding, the quick replacement of a bound ligand by a nearby one due to proximity and high local concentration[150].

Considering the variety of biological events involving carbohydrate-protein interactions, the development of specific multivalent inhibitors has a great therapeutical potential. For instance, a number of multivalent glycoconjugate-based vaccines for various bacterial infections are already commercially available and others under clinical trials[151].

146. Capon, B. Mechanism in Carbohydrate Chemistry. *Chem. Rev.* **1969**, *69*, 407–498.
147. Lundquist, J. J.; Toone, E. J. The Cluster Glycoside Effect. *Chem. Rev.* **2002**, *102*, 555–578.
148. Roy, R.; Zanini, D.; Meunier, S. J.; Romanowska, A. Solid-Phase Synthesis of Dendritic Sialoside Inhibitors of Influenza A Virus Haemagglutinin. *J. Chem. Soc. Chem. Commun.* **1993**, 1869.
149. Renaudet, O.; Roy, R. Multivalent Scaffolds in Glycoscience: An Overview. *Chem. Soc. Rev.* **2013**, *42*, 4515–4517.
150. Pieters, R. J. Maximising Multivalency Effects in Protein-Carbohydrate Interactions. *Org. Biomol. Chem.* **2009**, *7*, 2013–2025.
151. Bhatia, S.; Dimde, M.; Haag, R. Multivalent Glycoconjugates as Vaccines and Potential Drug Candidates. *Med. Chem. Commun.* **2014**, *5*, 862–878.

2. Dynamic glycopeptides with high affinity for Wheat Germ Agglutinine (WGA)

a. Design and objectives

The formation of active biomolecules by means of reversible native chemical ligation was one of our research goals, as already reported in Chapter 4. In this case, we aimed at extending this approach to glycopeptides, using a specific lectin to drive the formation of its best binder. Interestingly, Fiore *et al.* [152] have recently described a tetravalent glycocyclopeptide with high affinity (IC_{50} = 1.5 nM) for wheat germ agglutinine, a lectine from *Triticum vulgaris* that specifically binds N-acetyl-glucosamine (GlcNAc) and N-acetylneuraminic acid. It is a stable homodimer protein and each polypeptide chain is organized into four domains, containing binding sites at a distance of 14-15 Å [153]. The tetravalent glycopeptides described belongs to the RAFT (regioselectively addressable functionalized template) family. Initially described by Mutter and coworkers[154], RAFTs are cyclic decapeptides whose well-defined geometry is conferred by a proline-glycine and a proline-x sequence (where x is mostly a glycine, Figure 55). Importantly, these molecules have an upper and a lower face, that can respectively contain up to 4 or 2 side chains of the remaining residues. These side chains (often lysines) can thus be conveniently orthogonally functionalized forming multivalent systems anchored to different substrates or conjugated to different molecules[155].

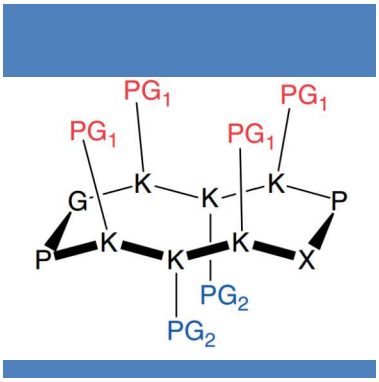


Figure 55 | Structure of a RAFT cyclopeptide with orthogonally protected lysine side chains on the two faces of the molecule. PG_1 = protecting group 1; PG_2 = protecting group 2.

152. Fiore, M.; Berthet, N.; Marra, A.; Gillon, E.; Dumy, P.; Dondoni, A.; Imberty, A.; Renaudet, O. Tetravalent Glycocyclopeptide with Nanomolar Affinity to Wheat Germ Agglutinin. *Org. Biomol. Chem.* **2013**, *11*, 7113–7122.
153. Schwefel, D.; Maierhofer, C.; Beck, J. G.; Seeberger, S.; Diederichs, K.; Möller, H. M.; Welte, W.; Wittmann, V. Structural Basis of Multivalent Binding to Wheat Germ Agglutinin. *J. Am. Chem. Soc.* **2010**, *132*, 8704–8719.
154. Dumy, P.; Eggleston, I. M.; Cervigni, S.; Sila, U.; Sun, X.; Mutter, M. A Convenient Synthesis of Cyclic Peptides as Regioselectively Addressable Functionalized Templates (RAFT). *Tetrahedron Lett.* **1995**, *36*, 1255–1258.
155. Boturyn, D.; Defrancq, E.; Dolphin, G. T.; Garcia, J.; Labbe, P.; Renaudet, O.; Dumy, P. RAFT Nano-Constructs: Surfing to Biological Applications. *J. Pept. Sci.* **2008**, *14*, 224–240.

In collaboration with the group of Prof. Stéphane Vincent at the University of Namur (Belgium), we have designed a RAFT similar to the one described by Fiore *et al.*[152] (Figure 56a), where N-methyl-cysteine (**Daa2**) replaces the alanine residues (Figure 56b). An exchange reaction among two linear glycopeptides shown in Figure 56c, in the presence of WGA, would lead to the formation and amplification of the corresponding cyclic dimer.

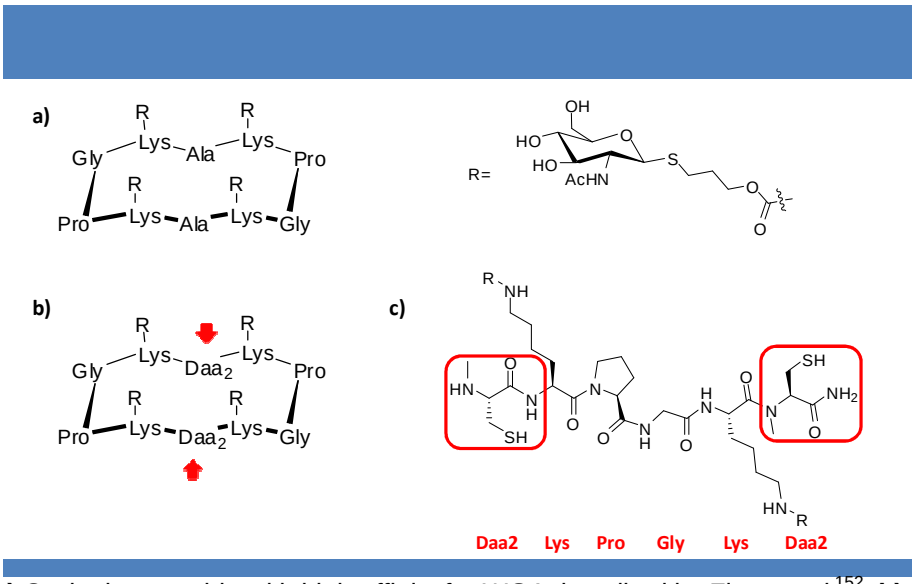
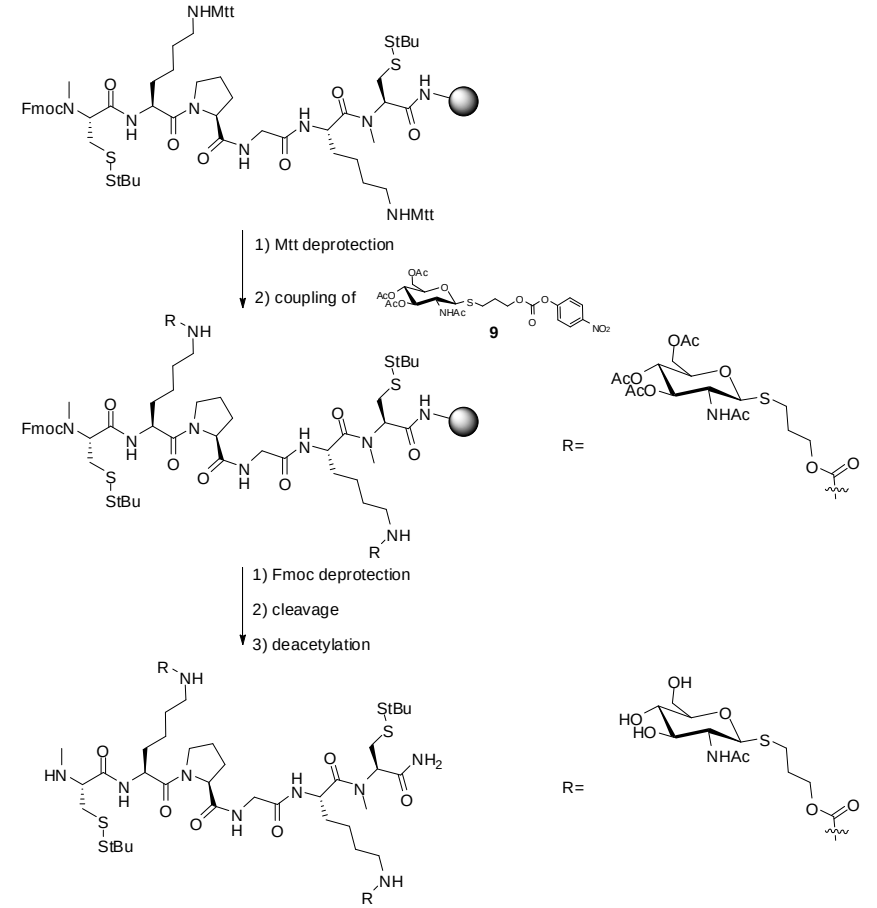


Figure 56 | a) Cycloglycopeptide with high-affinity for WGA described by Fiore *et al.*[152]; **b)** design of the dynamic cycloglycopeptide; **c)** design of the linear dynamic glycopeptides whose cyclization affords **b)**.

b. Glycopeptide synthesis

We planned to perform the synthesis of the linear peptide Fmoc-**Daa2**(StBu)-Lys(Mtt)-Pro-Gly-Lys(Mtt)-**Daa2**(StBu)-● in standard SPPS conditions as shown in Scheme 22; this short peptide incorporates two StBu-protected N-methyl-cysteine units (**Daa2**) and two lysines protected as 4-methyltrytil (Mtt). As already reported in Chapter 4, the StBu group on the N-methyl-cysteine moieties was expected to be easily removed in the reducing conditions (DTT) used for the exchange reaction. In the case of lysine, Mtt was preferred to the standard Boc protecting group as it is cleaved is mild acidic conditions[156], allowing for a selective deprotection of this moiety without causing simultaneous cleavage of the peptide from the resin. In alternative to Mtt, we also considered the possibility of using the allyloxycarbonyl group (Alloc)[157], that can be selectively cleaved by $Pd(PPh_3)_4$ in the presence of $PhSiH_3$ as a scavenger.

156. Li, D.; Elbert, D. L. The Kinetics of the Removal of the N-Methyltrityl (Mtt) Group during the Synthesis of Branched Peptides. *J. Pept. Res.* **2002**, *60*, 300–303.
157. Thieriet, N.; Alsina, J.; Giralt, E.; Guibé, F.; Albericio, F. Use of Alloc-Amino Acids in Solid-Phase Peptide Synthesis. Tandem Deprotection-Coupling Reactions Using Neutral Conditions. *Tetrahedron Lett.* **1997**, *38*, 7275–7278.

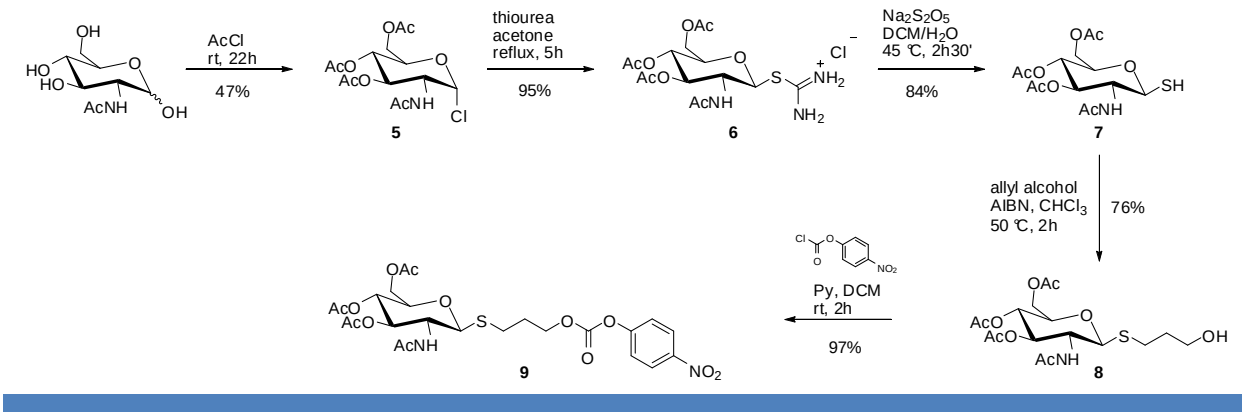


Scheme 22 | Planned synthesis of the linear glycopeptide to be used as starting material in the exchange reaction.

After selective deprotection of the side chain amino groups on the two lysine moieties, we planned to couple the carbohydrate unit **9** on solid phase with DIEA in NMP, according to a previously reported protocol[158]. After standard[159] Fmoc deprotection, cleavage from the resin and deacetylation in solution, we expected to obtain the final glycopeptide. Deacetylation is normally performed with NaOMe in MeOH; however, ammonia-saturated methanol[160] is a valid milder alternative for the synthesis of glycopeptides, whose backbone can undergo racemization in harsh basic conditions.

Compound **9** was synthesized according to Scheme 23.

158. Wittmann, V.; Seeberger, S. Spatial Screening of Cyclic Neoglycopeptides: Identification of Polyvalent Wheat-Germ Agglutinin Ligands. *Angew. Chem. Int. Ed.* **2004**, *43*, 900–903.
159. a) Sjölin, P.; Elofsson, M.; Kihlberg, J. Removal of Acyl Protective Groups from Glycopeptides: Base Does Not Epimerize Peptide Stereocenters, and B-Elimination Is Slow. *J. Org. Chem.* **1996**, *61*, 560–565. b) Hojo, H.; Nakahara, Y. Recent Progress in the Field of Glycopeptide Synthesis. *Biopolymers* **2007**, *88*, 308–324.
160. Kunz, H.; Waldmann, H. Construction of DisaccharideN-Glycopeptides - Synthesis of the Linkage Region of the Transmembrance-Neuraminidase of an Influenza Virus. *Angew. Chem. Int. Ed.* **1985**, *24*, 883–885.



Scheme 23 | Planned synthesis of the linear glycopeptide, to be used as a building block in the exchange reaction.

Compound **7** was synthesized according to a protocol reported in the literature[161] starting from commercial N-acetyl-D-glucosamine. Simultaneous acetylation of the hydroxyl groups in positions 3, 4 and 6 and chlorination of the anomeric carbon with acetyl chloride afforded compound **5** with a 47% yield. Successive reaction with thiourea in acetone provided the corresponding isothiouronium chloride **6** with a 95% yield and compound **7** was then obtained with sodium metabisulfite in phase-transfer conditions (84% yield). For the synthesis of compound **8**, we explored two different strategies in parallel. While alkylation of the thiol group with bromopropanol and triethylamine[162] in DCM allowed for the synthesis of compound **8** with a 54% yield, the same reaction performed with allyl alcohol and azobisisobutyronitrile (AIBN) in chloroform gave compound **8** with a 76% yield. This second strategy, adapted from a protocol reported by Knapp and Myers[163], was thus preferred to the first one. Finally, compound **9** was obtained with a 97% yield by reacting compound **8** with p-nitrophenylchloroformate and pyridine in DCM.

The synthesis of compound **9** was carried out in the laboratories of Prof. Stéphane Vincent in Namur (Belgium). Although planned in detail, the synthesis of the final glycopeptide was not performed. An exchange reaction between two or more molecules of this glycopeptide would lead to a library of linear and cyclic components. Being an analogue to the RAFT cycloglycopeptide, we expect the cyclic dimer to be amplified in the presence of its protein target, WGA. Unfortunately, the structure of WGA is stabilized by 32 disulfide bonds, 16 in

161. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **2009**, *48*, 7798–7802.
162. Hoang, K. L. M.; Bai, Y.; Ge, X.; Liu, X.-W. Exploring the Native Chemical Ligation Concept for Highly Stereospecific Glycosylation Reactions. *J. Org. Chem.* **2013**, *78*, 5196–5204.
163. Knapp, S.; Myers, D. S. Synthesis of α-GalNAc Thioconjugates from an α-GalNAc Mercaptan. *J. Org. Chem.* **2002**, *67*, 2995–2999.

each subunit[164]. This project was thus momentarily stopped in order to further investigate the exchange reaction on a simpler system, with the objective of developing a protocol that allows the exchange reaction to occur without being detrimental for the protein target.

3. Dynamic glycopeptides with high affinity for O-GlcNAcase (OGA)

a. Protein β-O-GlcNAcylation

Protein glycosylation is fundamental in a number of biological events. As discussed above, protein glycosylation generally results in branched polysaccharidic structures, often presented on the cell surface or in the extracellular matrix. Unlike the "classical" protein glycosylation, β-O-GlcNAcylation on the side chains of serine and threonine residues has been mostly observed within the cytoplasm or nucleoplasm. Moreover, the O-GlcNAc units are not further modified nor incorporated in complex polysaccharidic structures.

O-GlcNAcylation is involved in fundamental biological processes, such as regulation of transcription, protein trafficking and turnover, and response to cellular stress. Importantly, its dysregulation has been associated with the aethiology of different human diseases, including diabetes[165] and neurodegenerative disorders[166].

It is interesting to underline that the mechanism and timescale of the O-GlcNAc cycling are very similar to those of phosphorylation[167]. These two events are independently regulated and can occur on the same polypeptide or protein, providing cells with a rapid response mechanism to different stimuli. Fine tuning of protein interactions and functions strictly depends on the enormous molecular diversity that can be generated by these two modifications. Importantly, while in some proteins O-GlcNAc and O-phosphate compete for the same hydroxyl group of serine or threonine, in other cases they occupy adjacent sites.

Only two enzymes regulate O-GlcNAcylation: O-GlcNAc transferase (OGT), that catalyzes the installation of this monosaccharide on proteins from the substrate donor UDP-GlcNAc, and O-GlcNAcase (OGA), responsible for its cleavage (Figure 57).

164. Portillo-Téllez, M. D. C.; Bello, M.; Salcedo, G.; Gutiérrez, G.; Gómez-Vidales, V.; García-Hernández, E. Folding and Homodimerization of Wheat Germ Agglutinin. *Biophys. J.* **2011**, *101*, 1423–1431.
165. Wells, L.; Vosseller, K.; Hart, G. W. A Role for N-Acetylglucosamine as a Nutrient Sensor and Mediator of Insulin Resistance. *Cell. Mol. Life Sci.* **2003**, *60*, 222–228.
166. Leroy, J. G. Congenital Disorders of N-Glycosylation Including Diseases Associated with O- as Well as N-Glycosylation Defects. *Pediatr. Res.* **2006**, *60*, 643–656.
167. Hart, G. W.; Housley, M. P.; Slawson, C. Cycling of O-Linked B-N-Acetylglucosamine on Nucleocytoplasmic Proteins. *Nature* **2007**, *446*, 1017–1022.

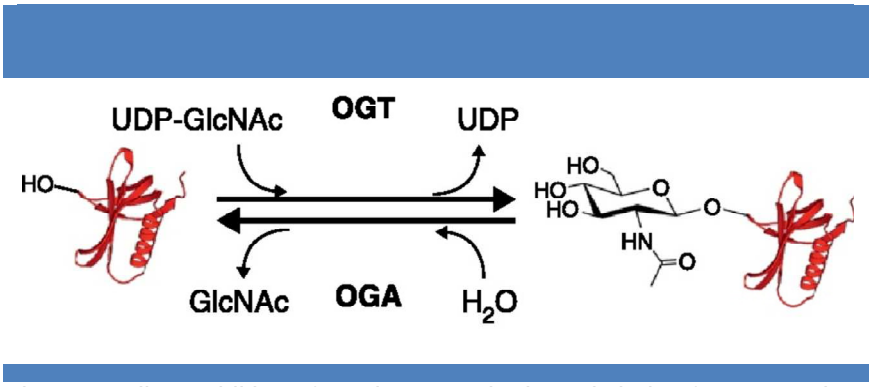


Figure 57 | O-GlcNAc cycling: addition of O-GlcNAc to the lateral chain of a Ser or Thr residue by OGT and its removal by OGA. The protein shown is not in scale with the sugar (adapted from ref. 168).

OGA is a 130 kDa enzyme, bearing the hydralase catalytic domain in its N-terminus. This enzyme is responsible for the removal of O-GlcNAc from more than 1000 peptides and proteins[169]; however, van Aalten and coworkers[170] have pointed out some common features in the binding of different O-GlcNAcylated peptides to OGA. The sugar always occupies the same position in the OGA binding pocket and assumes a $^{1,4}B$ boat conformation. Regarding the peptide backbone, it assumes a "V-shaped" conformation around the O-GlcNAc moiety and its side chains point away from the surface of the enzyme.

Among the glycopeptides investigated in this study, an 11-amino acid fragment of protein p53 was shown (Figure 58a) to have high affinity for both human OGA (hOGA, K_M = 21 μM) and *Clostridim perfrigens* OGA (*Cp*OGA, K_M = 3 μM). The authors noticed that the "V-shaped" conformation of this peptide in the OGA binding pocket was stabilized by an hydrogen bond between the Asp and Thr residues adiacent to the Ser(O-GlcNAc) (Figure 58b). Disruption of this hydrogen bond by substitution of either Asp or Thr with Val resulted in a drastic reduction in the binding affinity of the peptide to OGA (Figure 58c).

168. Shen, D. L.; Gloster, T. M.; Yuzwa, S. A.; Vocadlo, D. J. Insights into O-Linked N-Acetylglucosamine ([0-9]O-GlcNAc) Processing and Dynamics through Kinetic Analysis of O-GlcNAc Transferase and O-GlcNAcase Activity on Protein Substrates. *J. Biol. Chem.* **2012**, *287*, 15395–15408.
169. Ostrowski, A.; van Aalten, D. M. F. Chemical Tools to Probe Cellular O-GlcNAc Signalling. *Biochem. J.* **2013**, *456*, 1–12.
170. Schimpl, M.; Borodkin, V. S.; Gray, L. J.; van Aalten, D. M. F. Synergy of Peptide and Sugar in O-GlcNAcase Substrate Recognition. *Chem. Biol.* **2012**, *19*, 173–178.

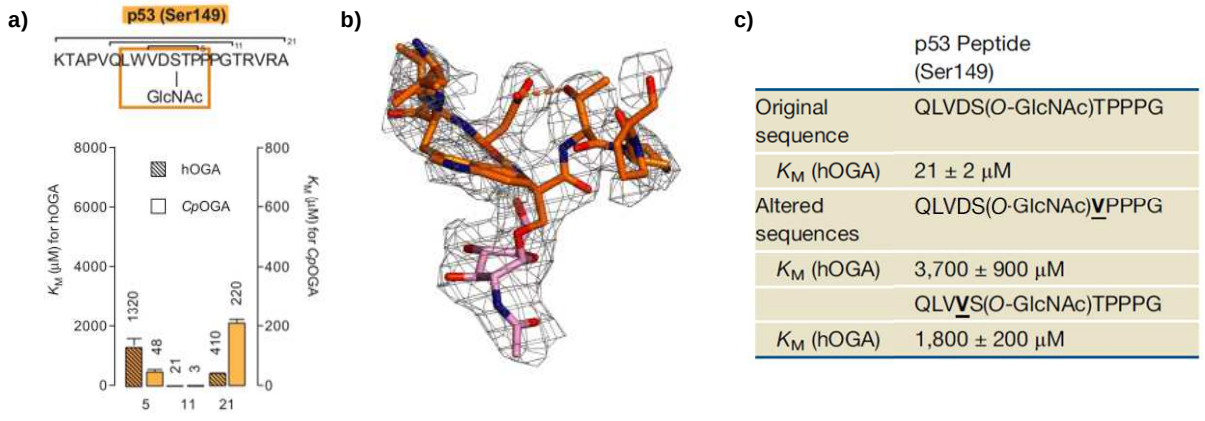


Figure 58 | a) Sequence and binding affinity of different peptides derived from protein p53. **b)** Conformation of the peptide bound in the active site of CpOGA as determined by X-ray crystallography and showing the intramolecular hydrogen bond between the lateral chains of Asp and Thr **c)** Table showing the reduced affinity of the glycopeptides for hOGA after disruption of the hydrogen bond by replacement of either Thr or Asp with Val (adapted from ref. 170).

b. Design and objectives

In this project, we aimed at investigating the formation of glycopeptides with high affinity for OGA by means of reversible native chemical ligation.

We designed a glycopeptide based on the fragment of protein p53 shown above, by substituting the Thr residue adiacent to Ser(O-GlcNAc) with a N-methyl-cysteine moiety (**Daa2**). Moreover, we replaced Ser(O-GlcNAc) with Cys(S-GlcNAc) that should in principle be bound but not cleaved by OGA (Figure 59a). This assumption is based on the observation that the pseudosubstrate Ala-Cys(S-GlcNAc)-Ala inhibits *Cp*OGA[171].

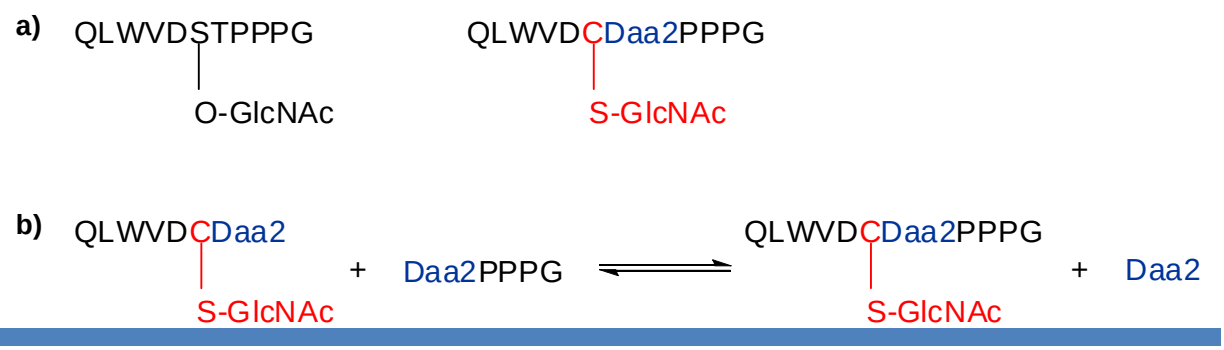


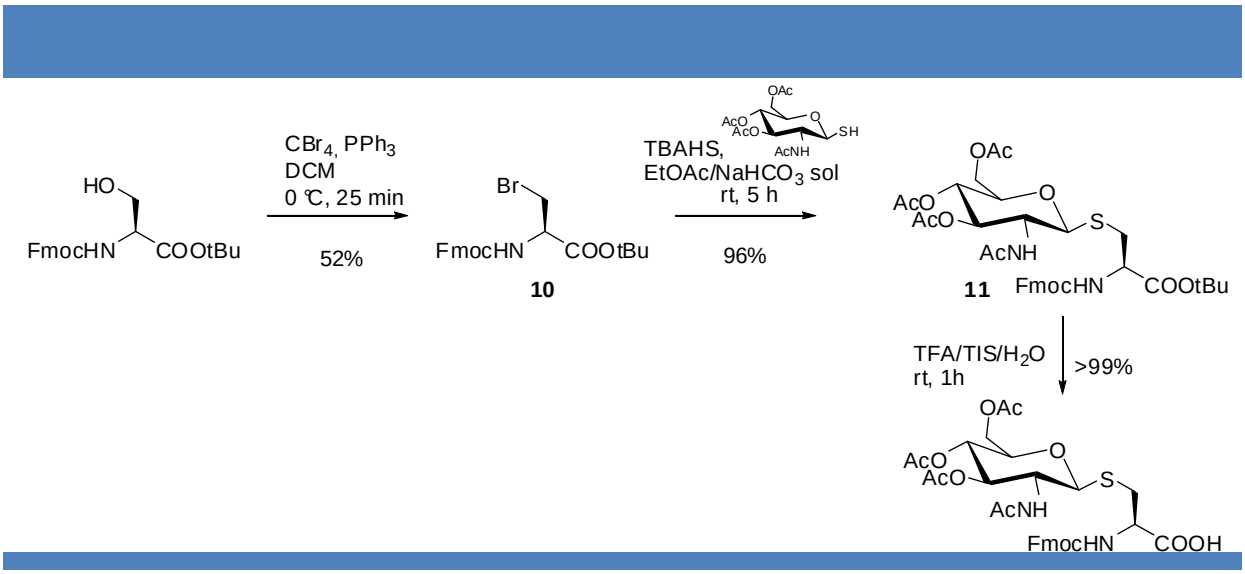
Figure 59 | a) Design of a dynamic glycopeptide, analogue to the one described in ref. 170. **b)** Exchange reaction between two fragments of this glycopeptides, leading to its formation.

171. Rao, F. V; Dorfmueller, H. C.; Villa, F.; Allwood, M.; Eggleston, I. M.; van Aalten, D. M. F. Structural Insights into the Mechanism and Inhibition of Eukaryotic O-GlcNAc Hydrolysis. *EMBO J.* **2006**, *25*, 1569–1578.

The exchange reaction shown in Figure 59b between two fragments of this glycopeptide should lead to its formation together with **Daa2**. Addition of randomized glycopeptides fragments (i.e. bearing a mannose unit or having a Val residue replacing the Asp one) and of OGA to the library should lead to the amplification of the full glycopeptide NH_2-QLVDC(S-GlcNAc)**Daa2**PPPG-$CONH_2$.

c. Glycopeptide synthesis

We planned to perform the synthesis of both glycopeptide fragments by solid phase peptide synthesis. The strategy consists in introducing a Fmoc-Cys(S-GlcNAc(Ac)$_3$)-OH unit in the sequence. The synthesis of this molecule was planned as shown in Scheme 24.



Scheme 24 | Planned synthesis of Fmoc-Cys(S-GlcNAc(Ac)$_3$)-OH.

Compound **11** was obtained according to a procedure reported by Zhu and Schmidt[172]. Commercial Fmoc-Ser-OtBu was initially brominated under standard conditions to afford compound 10 with a 62% yield. The resulting bromo alanine derivative was then coupled in phase-transfer conditions to compound **7**, synthesized according to the procedure previously reported in this Chapter. The final C-terminus deprotection in acidic conditions and in the presence of TIS and H_2O as scavengers was not performed; however, it is reported[173] to occur with quantitative yield.

The synthesis of compound **11** was performed in the laboratories of Prof. Stéphane

172. Zhu, X.; Schmidt, R. R. Efficient Synthesis of S-Linked Glycopeptides in Aqueous Solution by a Convergent Strategy. *Chemistry* **2004**, *10*, 875–887.
173. Hsieh, Y. S. Y.; Wilkinson, B. L.; O'Connell, M. R.; Mackay, J. P.; Matthews, J. M.; Payne, R. J. Synthesis of the Bacteriocin Glycopeptide Sublancin 168 and S-Glycosylated Variants. *Org. Lett.* **2012**, *14*, 1910–1913.

Vincent in Namur (Belgium) and we planned to start a collaboration with Prof. Tony Lefebvre from the University of Lille (France), who can provide us with OGA. However, before studying the system in the presence of this precious proteic target, we wanted to improve few parameters of the exchange reaction by testing it on a simpler system. Although planned in detail, this project was thus momentarily stopped.

CONCLUSIONS AND PERSPECTIVES

The main objective of my PhD was to develop a reversible chemistry to scramble peptide fragments in mild, bio-compatible conditions. To achieve this goal, we have investigated the potential reversibility of a well-known reaction in peptide chemistry: native chemical ligation (NCL). The beauty of this two-step reaction rests in the easy formation of a peptide bond in mild aqueous conditions. Importantly, a cysteine residue is essential in this process.

With the idea of further enhancing the reactivity of cysteine, we have attached a 2-thioethyl chain on its amino group. Peptides incorporating this new dynamic amino acids were shown to quickly exchange fragments. Hence, we have developed the first example of a reversible chemistry that can be applied to peptides at the amide bond level (Chapter 2). However, one of the limitations of this strategy is the necessity to incorporate a non-natural amino acid in the peptide sequence. Since the unknown bioactivity and biocompatibility of N-(2-thioethyl)-cysteine is a potential issue, we investigated which naturally occurring cysteine modification could lead to a reversible amide junction (Chapter 3). N-methyl cysteine, that can be found in many natural peptides, was thus introduced in our model peptides. Indeed, it was shown to be a new dynamic unit. We have thus described two small modifications of the peptide backbone at the cysteine level that can activate a specific bond for exchange reactions. Importantly, the peptide fragments are scrambled in the presence of DTT in physiological conditions: aqueous buffer, mildly acidic to mildly basic pH and room temperature (20-40 °C). With this project we have added reversible NCL to the list of covalent interactions that can be used in dynamic combinatorial chemistry. One can only speculate about the possible applications of this reaction, that can range from drug discovery to materials science.

A second important goal of this project was to provide a practical application of this strategy. Using reversible NCL, we wanted to create dynamic combinatorial libraries of peptides to test against a given biological target. To this end, we have designed a project based on two affibody molecules specific for either the HER2 protein or IgG. In Chapter 4 I have extensively described the different issues that emerged in this project. From a library of affibody helices, we have demonstrated the formation of different polypeptides, including the full affibody trimer. However, complications arised due to the incompatibility of the target molecules with the reducing conditions used for the exchange reactions.

In order to overcome these problems, we decided to further investigate the exchange reaction in a simpler system. First, we have synthesized short model peptides incorporating a

non-protected N-methyl-cysteine, obviating the need for in-situ StBu deprotection during the exchange reaction. Second, we have greatly simplified the setup, allowing the reaction to occur in an Eppendorf tube under ambient atmosphere. This was possible by substituting DTT with a mixture of TCEP and ascorbic acid. While TCEP is a reducing agent less sensitive to air oxidation than DTT, ascorbic acid was added to prevent TCEP-induced desulfurization of the peptides (i.e. cysteine into alanine). We have then investigated the exchange reaction between the model peptides in this new setup and explored the catalytic behavior of different thiol molecules. These first results are indeed encouraging and similar experiments will soon be performed in the presence of resin-bound TCEP. After optimization of all the parameters, it will be possible to perform the exchange reaction among the affibody helices in the presence of their biological target. Although I have actively participated in the design of these experiments, this work has been mainly carried out by Cristian Reţe and is thus not further discussed in this manuscript.

During this PhD, I have also tried to extend the "dynamic peptide" approach to glycopeptides (Chapter 5). To this end, I have designed two different projects in collaboration with Prof. Stéphane Vincent. In one case the target molecule is the lectine WGA, in the other case the enzyme OGA. However, these two projects are still at an initial stage. Due to the complications emerged in the project with the affibody molecules, we preferred to momentarily stop other experiments and focus on further investigating the exchange reaction in the simpler system discussed above.

So far my PhD work has resulted in one publication and one patent:

- Ruff, Y., Garavini, V., Giuseppone, N. Reversible native chemical ligation: A facile access to dynamic covalent peptides. *J. Am. Chem. Soc.* **2014**, *136*, 6333-6339
- Ruff, Y., Giuseppone, N., Garavini, V. Réactions d'échanges peptidiques par ligation chimique. *Eur. Appl.* **2013**, EP13306843.7

EXPERIMENTAL SECTION

1. Solvents and reagents

All reagents and solvents were purchased from Sigma-Aldrich, Carlo Erba, Iris-Biotech, Polypeptide, Alfa Aesar and Fisher Scientific at the highest commercial quality and used without further purification unless stated otherwise. Dry solvents were obtained using a double column *SolvTech* purification system. Water was deionized by using a Millipore *Q-POD Milli-Q* system.

2. Peptide synthesis

Solid phase peptide synthesis was performed on a microwave synthesizer (CEM *Liberty 1*) and on a parallel synthesizer (Heidolph *Synthesis 1*). Unless stated otherwise, all Fmoc amino acids were introduced with standard side chain protecting groups: Pbf for Arg; Trt for Asn, Cys, Gln and His; tBu for Asp, Glu, Ser, Thr and Tyr; Boc for Lys and Trp; no protecting groups were needed for Ala, Gly, Ile, Leu, Met, Phe, Pro and Val.

3. Chromatographic methods

a. Thin Layer Chromatography (TLC)

Analytical TLC was performed on silica TLC Al foils (silica gel matrix with fluorescent indicator 254 nm, thickness: 500 μm, Sigma-Aldrich). Compounds were visualized by fluorescence quenching detection (Bioblock VL-4C UV-Lamp, 6 W, 254 nm and 365 nm) and/or dipping into standard staining solutions based on phosphomolybdic acid and cerium molybdate.

b. Preparative absorbance flash column chromatography

Flash chromatography was performed on silica gel (60 Å, 230-400 mesh, 40-63 μm, Sigma-Aldrich). Separations were performed manually using glass columns of different sizes.

c. Analytical Ultra Performance Liquid Chromatography (UPLC)

i. UPLC-ESI

Analytical Ultra Performance Liquid Chromatography (UPLC) was carried out on a Waters *ACQUITY UPLC* coupled with a *SQD Acquity* electrospray mass detector and a photodiode array (PDA) detector (190–500 nm, 80Hz). Compounds were separated on a Waters *ACQUITY UPLC HSS* T3 1.8 µm particle size, 2.1 x 50 mm column or Waters *ACQUITY UPLC BEH* C_{18} 1.7 µm, 2.1 x 50 mm column using 0.1 %v/v formic acid in water (solvent A) and 0.1 %v/v formic acid in acetonitrile (solvent B) as the two mobile phases. Data were processed with the MassLynx 4.1 – XP software. The two gradients used (A and B) are shown below.

Gradient A

Time (min)	Flow (mL/min)	% Solvent A	% Solvent B
0	1	95	5
3.5	1	5	95
4	1	5	95
4.10	1	95	5
5.10	1	95	5

Gradient B

Time (min)	Flow (mL/min)	% Solvent A	% Solvent B
0	1	95	5
0.80	1	80	20
2	1	72.5	27.5
4	1	20	80
4.10	1	95	5
5.10	1	95	5

ii. UPLC-ESI-TOF

The affidody exchange reaction mixtures were analyzed by Dr. Jean-Marc Strub on a Waters *nanoACQUITY UPLC* coupled to a *Waters SYNAPT*, an hybrid quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer equipped with a Z-spray ion source and a lock mass system. Peptides were separated on a Waters *ACQUITY UPLC BEH130* C_{18} column 1.7 µm particle size, 75 µm x 250 mm using 0.1 %v/v formic acid in water (solvent A)

and 0.1 %v/v formic acid in acetonitrile (solvent B).

d. Preparative High Performance Liquid Chromatography (HPLC)

Preparative High Performance Liquid Chromatography (HPLC) was carried out on a Waters *AutoPurification* system, coupled to a Waters *SQD 3100* electrospray mass detector and a Waters 2489 UV/Visible detector. Purifications were performed either using 0.1 %v/v formic acid in water (solvent A) and 0.1 %v/v formic acid in methanol (solvent B) as the two mobile phases on a Waters *SunFire Prep C18 QBD* 5 µm 19x150 mm column or using 0.05 %v/v ammonia in water (solvent C) and 0.05 %v/v ammonia in methanol (solvent D) as the two mobile phases on a Waters *XBridge Prep C18 QBD* 5 µm 19x150 mm column (typical gradients shown below).

Time (min)	Flow (mL/min)	% solvent A (or C)	% solvent B (or D)
0	23.28	100	0
4.95	23.28	100	0
29.03	23.28	5	95
37.06	23.28	5	95
37.86	23.28	100	0
45.89	23.28	100	0

4. Structure determination

a. Nuclear Magnetic Resonance (NMR)

^{1}H NMR spectra were recorded either on a *Bruker Avance 400* spectrometer at 400 MHz, on a *JEOL ECX-400* spectrometer at 400 MHz or on a *Bruker Avance 500* spectrometer at 500 MHz and ^{13}C NMR spectra either at 100 MHz or 125 MHz. Residual solvent peaks were taken as reference ($CDCl_3$: 7.26 ppm for ^{1}H NMR and 77.36 for ^{13}C NMR; DMSO-d_6: 2.50 ppm for ^{1}H NMR and 40.45 ppm for ^{13}C NMR). In the case of compounds solubilized in D_2O, spectra were internally referenced to the CH_3- signals of *t*BuOH (1.24 ppm for ^{1}H NMR and 30.29 ppm for ^{13}C NMR). Spin multiplicities are reported as a singlet (s), doublet (d), doublet of doublets (dd), triplet (t) with coupling constants (J) given in Hz, or multiplet (m). Broad peaks are marked as br. When specified, ^{1}H and ^{13}C resonances were assigned with the aid of additional information

from respectively 2D NMR proton COrrelation SpectroscopY (COSY) or Heteronuclear Multiple-Quantum Correlation (HMQC) experiments.

b. Mass Spectrometry (MS)

i. ESI-MS

ESI-MS spectra were acquired using the Waters *SQD* mass detector directly connected to the output of an UPLC-ESI or HPLC column.

ii. ESI-TOF

ESI-TOF spectra were acquired using the Waters *SYNAPT* mass detector connected to the output of the *nanoACQUITY UPLC* column.

iii. MALDI-TOF

MALDI-TOF measurements were carried out on an *AutoflexTM MALDI-TOF* mass spectrometer (Bruker Daltonics GmbH). Externally calibration was performed with standard peptide/protein calibration mixtures that contained 7 peptides/proteins (Bruker Daltonics GmbH) covering the 1000-4000 m/z or the 5000-17500 m/z mass range. Mass data collected during MALDI-TOF analyzes were processed using the software tool *flexAnalysis* (Bruker Daltonics GmbH). Sample preparation was performed with the dried droplet method using a mixture of 1 µL of sample with 1 µL of matrix solution (10 mg/mL of 2,5-dihydroxybenzoic acid in water/acetonitrile/TFA 50/50/0.1) dried at room temperature.

c. Infrared spectroscopy

IR spectra were recorded on a Fourier transform infrared spectrometer *Vertex 70* (Bruker), equipped with a diamond ATR accessory (Smiths detections).

d. Polarimetry measurements

Optical rotatory powers were measured on a *Perkin-Elmer 241* polarimeter at 20 °C with a wavelength of 589 nm (sodium lamp). Optical rotatory power values were obtained with the Biot law $[\alpha]_D^{20} = \frac{\alpha \cdot 100}{C \cdot l}$ where α is the deviation angle of the polarized light at 20 °C, C is the sample concentration (g/100 mL) and l is the length of the cell (dm).

5. Synthesis and characterization of organic compounds

a. Chapter 2

Compound 1

1

To a solution of triphenylmethyl chloride (5.39 g, 68.99 mmol) in $CHCl_3$ (20 mL), 2-mercaptoethanol (16.1 mmol, 1.26 g) was added and the reaction was left stirring under argon for 24 h. The solvent was then evaporated and the crude recrystallized with cyclohexane and $CHCl_3$ (100 mL, 20:1) to obtain compound **1** (4.19 g, 81 %) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz, 25°C): δ = 7.45-7.15 (m, 15H), 3.37 (t, *J* = 6.1 Hz, 2H), 2.47 (t, *J* = 6.1 Hz, 2H), 1.57 (s, 1H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25°C): δ = 144.72, 129.54, 127.89, 126.72, 66.64, 60.81, 35.18 ppm; ESI-MS: m/z calcd for $C_{21}H_{20}OS$ $[M+H]^+$ 321.13, found 321.54.

Peptides incorporating the dynamic unit N-(2-thioethyl)-cysteine (Daa1)

Peptide synthesis was performed on a Fmoc-Rink-amide-MBHA resin (0.74 mmol/g, 0.5 mmol scale) using a combination of conventional and microwave-assisted SPPS.

Swelling of the resin (in the microwave synthesizer): the resin was let swell in DMF (15 mL) for 30 min.

Fmoc deprotection (in the microwave synthesizer): two successive treatments of the resin with a 20 %v/v piperidine solution in DMF (15 mL) at 70 ºC and 55 W for 3 min.

Amino acid coupling (in the microwave synthesizer): four-fold molar excess of each Fmoc-aa-OH (10 mL of a 0.2 M solution in DMF), HBTU (4 mL of a 0.5 M solution in DMF) and DIEA (2 mL of a 2 M solution in NMP) at 70 ºC and 35 W (microwave power) for 5 min. Fmoc-Cys(Trt)-OH was coupled for 2 min without microwave heating and then at 50 ºC with 25 W for 4 min.

Successive Fmoc deprotection and amino acid coupling steps led to the synthesis of sequences **P2** and **P3**, respectively NH_2-Cys(Trt)-Ala-Lys(Boc)-Leu-Leu-● and NH_2-Cys(Trt)-Ala-Phe-Lys(Boc)-Phe-●.

Sulfonylation (in the parallel synthesizer): a solution of 2-nitrobenzenesulfonylchloride (2 mmol, 443 mg) and 2,6-lutidine (3 mmol, 347 µL) in CH_2Cl_2 (5 mL) was added to the resin and

let shake for 3 h.

Alkylation (in the parallel synthesizer): a solution of **1** (2.5 mmol, 800 mg), triphenylphosphine (2.5 mmol, 655 mg) and DIAD (2.5 mmol, 492 μL) in a dry THF/CH_2Cl_2 mixture (5 mL and 6.5 mL respectively) was added to the washed resin under argon atmosphere and let shake for 5 h. This reaction was then repeated once more for 12 hours.

Removal of NBS (in the parallel synthesizer): a solution of 2,2'-(ethyldioxy)diethanethiol (10 mmol, 1.63 mL) and DBU (5 mmol, 748 μL) in DMF (5 mL) was added to the resin and let shake for 2 h. The reaction was repeated a second time for 2 h.

For peptides **P1-Daa1-P2** and **P1-Daa1-P3**:

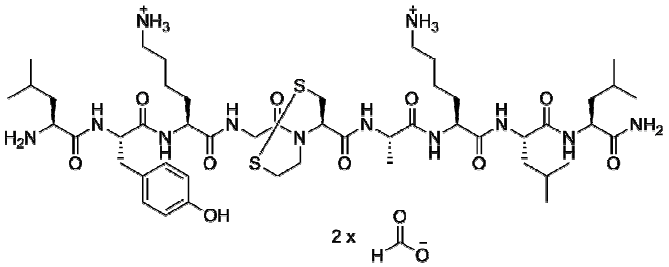
- Fmoc-Gly-OH coupling (in the parallel synthesizer): a solution of Fmoc-Gly-OH (4 mmol, 1.19 g), DIC (4 mmol, 620 μL) and OxymaPure (4 mmol, 568 mg) in DMF (5 mL) was let shake for 72 h at rt under argon atmosphere.
- Successive Fmoc deprotection and amino acid coupling steps (in the microwave synthesizer) of Fmoc-Lys(Boc)-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Leu-OH (**P1**).

Cleavage (in the parallel synthesizer): the resin was suspended in a 10 mL solution of TFA/TIS/2-2'(ethyldioxy)diethanethiol/H_2O: 92.5/2.5/2.5/2.5 vol% for 1 h, then filtered and washed with TFA (5 mL). The combined filtrates were concentrated *in vacuo* and then added to cold diethylether (200 mL). The precipitate was separated by centrifugation, washed with cold diethylether (2 x 40 mL) and dried under argon to afford crude peptides.

Oxidation: crude peptides were dissolved in milliQ H_2O (1 mL/50 mg), the pH was adjusted to 7-8 with $(NH_4)_2CO_3$ and compressed air was bubbled in the resulting solution for 30 min. Prior to purification, the pH was lowered to 4-5 with formic acid.

Purification: preparative HPLC using 0.1 %v/v formic acid in water (solvent A) and 0.1 %v/v formic acid in methanol (solvent B) as mobile phases on the preparative SunFire Prep C18 column. Fractions were automatically collected using the ESI-MS signal of the peptide as a trigger and then individually analyzed via UPLC-ESI. The pure fractions were then combined and the organic solvent removed under vacuum prior to lyophilization.

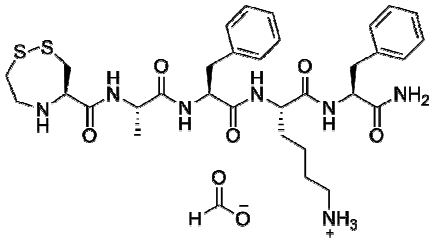
- *P1-Daa1-P2ox*



P1-Daa1-P2ox

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.53 (bs, 2H), 7.14 (d, *J* = 8.6 Hz, 2H), 6.83 (d, *J* = 8.6 Hz, 2H), 4.59 (dd, *J* = 9.7 Hz, 6.8 Hz, 1H), 4.40-4.15 (m, 7H), 4.05-3.93 (m, 2H), 3.88-3.76 (m, 1H), 3.55-3.32 (m, 2H), 3.21-2.87 (m, 8H), 1.85-1.53 (m, 17H), 1.50-1.28 (m, 7H), 0.99-0.84 (m, 18H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 177.98, 175.94, 175.13, 174.41, 173.54, 172.89, 172.21, 171.65, 170.76, 163.90, 163.62, 155.31, 131.31, 128.20, 118.21, 116.24, 115.89, 56.19, 54.17, 53.72, 53.57, 53.07, 52.75, 52.34, 50.89, 40.58, 40.42, 40.16, 39.88, 39.82, 36.81, 31.33, 30.91, 26.96, 26.92, 25.04, 25.01, 24.51, 22.95, 22.80, 22.55, 22.29, 22.00, 21.48, 21.21, 17.16 ppm; UPLC (gradient A): t_r = 0.83 min; ESI-MS: m/z calcd for $C_{49}H_{84}N_{12}O_{10}S_2$ $[M+2H]^{2+}$ 533.35, found 533.63.

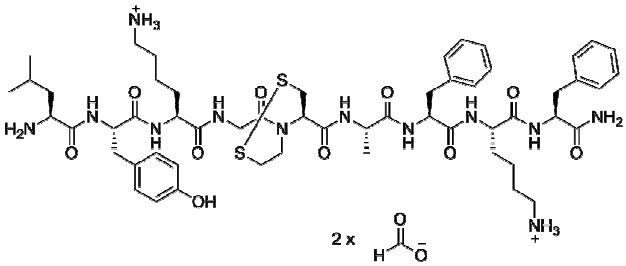
- *Daa1-P3ox*



Daa1-P3ox

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.45 (s, 1H), 7.42-7.15 (m, 10H), 4.59-4.50 (m, 2H), 4.27-4.16 (m, 2H), 3.86 (dd, *J* = 10.75 Hz, 4.27 Hz, 1H), 3.52-3.42 (m, 1H), 3.34 (dd, *J* = 13.48 Hz, *J* = 4.27 Hz, 1H), 3.22-2.88 (m, 9H), 2.81 (dd, *J* = 13.48 Hz, 10.75 Hz, 1H), 1.68-1.50 (m, 4H), 1.31 (d, *J* = 7.17 Hz, 3H), 1.26-1.14 (m, 2H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 176.19, 174.88, 173.39, 173.30, 171.54, 137.07, 136.61, 129.96, 129.88, 129.50, 129.48, 127.94, 62.84, 55.51, 55.43, 53.88, 50.42, 49.37, 39.82, 37.67, 37.64, 31.21, 26.94, 22.48, 17.12 ppm; UPLC (gradient A): t_r = 0.86 min; ESI-MS: m/z calcd for $C_{32}H_{45}N_7O_5S_2$ $[M+2H]^{2+}$ 336.68, found 336.83.

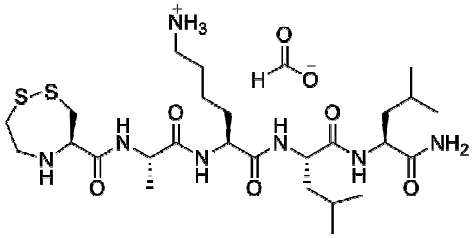
- *P1-Daa1-P3ox*



P1-Daa1-P3ox

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.46 (s, 1H), 7.41-7.18 (m, 10H), 7.13 (d, *J* = 8.6 Hz, 2H), 6.83 (d, *J* = 8.6 Hz, 2H), 4.67-4.50 (m, 4H), 4.37-4.16 (m, 5H), 4.04-3.95 (m, 2H), 3.81-3.69 (m, 1H), 3.43-3.27 (m, 2H), 3.23-2.87 (m, 13H), 1.82-1.50 (m, 12H), 1.41-1.27 (m, 4H), 1.22-1.11 (m, 2H), 0.99-0.88 (m, 6H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 176.20, 175.57, 173.64, 173.61, 173.52, 172.90, 172.10, 171.48, 163.91, 163.63, 155.30, 137.14, 136.70, 131.31, 129.96, 129.93, 129.54, 129.48, 128.22, 128.04, 127.89, 118.20, 116.23, 115.89, 56.14, 55.53, 55.34, 54.16, 53.58, 52.39, 50.83, 41.45, 40.72, 39.82, 39.79, 38.88, 37.59, 37.28, 36.78, 31.31, 31.04, 26.97, 26.91, 24.52, 22.55, 22.53, 22.30, 22.02, 16.59 ppm; UPLC (gradient A): t_r = 0.92 min; ESI-MS: m/z calcd for $C_{55}H_{80}N_{12}O_{10}S_2$ $[M+2H]^{2+}$ 567.29, found 567.26.

- *Daa1-P2ox*



Daa1-P2ox

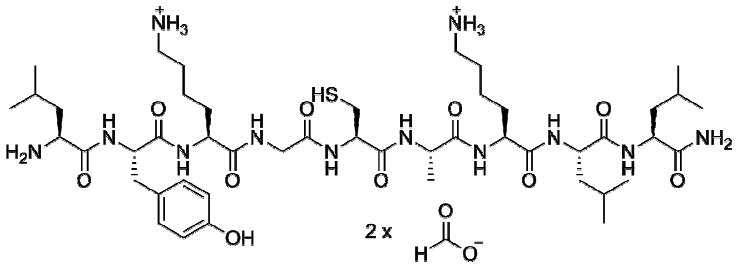
^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.47 (bs, 1H), 4.41-4.22 (m, 4H), 4.11-4.00 (m, 1H), 3.62-3.45 (m, 2H), 3.23-3.09 (m, 2H), 3.07-2.89 (m, 4H), 1.86-1.51 (m, 10H), 1.51-1.40 (m, 2H), 1.40-1.31 (d, ^{3}J = 7.50 Hz, 3H), 1.08-0.69 (m, 12H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 178.00, 175.64, 175.13, 174.31, 63.07, 54.12, 53.01, 52.76, 50.53, 49.54, 40.42, 40.14, 39.88, 30.97, 26.97, 25.04, 25.02, 22.93, 22.81, 22.76, 21.45, 21.24, 17.17 ppm; UPLC (gradient A): t_r = 0.73 min; ESI-MS: m/z calcd for $C_{26}H_{49}N_7O_5S_2$ $[M+H]^{+}$ 302.67, found 302.93.

Peptides incorporating cysteine (Cys)

Peptide synthesis was performed on a Fmoc-Rink-amide-MBHA resin (0.74 mmol/g, 0.5 mmol scale) using microwave-assisted SPPS. Swelling of the resin, Fmoc deprotection, amino acid coupling, cleavage and purification were performed according to the procedure reported above (for the synthesis of peptides incorporating Daa1).

The peptides were isolated in their reduced form and their purity assessed by UPLC-ESI. In order to prevent oxidation during the NMR acquisition, we lowered the pH of D_2O by adding an excess of formic acid. However, and even after degassing the solvent, we could not prevent the partial oxidation of **Cys-P3** (disulfide dimer), resulting in additional peaks in the ^{1}H and ^{13}C NMR spectra.

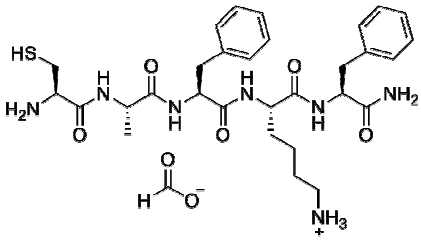
- *P1-Cys-P2*



P1-Cys-P2

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.32 (s, formate), 7.15 (d, *J* = 8.0 Hz, 2H), 6.83 (d, *J* = 8.0 Hz, 2H), 4.62 (dd, *J* = 8.9 Hz, 7.2 Hz, 1H), 4.52 (t, *J* = 6.5 Hz), 4.38-4.20 (m, 5H), 4.00 (t, *J* = 7.4 Hz, 1H), 3.89 (s, 2H), 3.10-2.86 (m, 8H), 1.86-1.52 (m, 18H), 1.50-1.27 (m, 7H), 0.97-0.82 (m, 18H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 177.20, 174.73, 174.31, 173.50, 173.33, 172.34, 171.57, 170.73, 169.98, 154.54, 130.55, 127.42, 115.43, 55.29, 55.23, 53.46, 53.23, 52.23, 51.98, 51.53, 49.90, 42.19, 39.79, 39.62, 39.35, 39.07, 39.01, 35, 95, 30.23, 26.20, 26.18, 25.33, 24.24, 24.23, 23.71, 22.14, 22.01, 21.98, 21.78, 21.55, 21.12, 20.66, 20.44, 16.30 ppm; UPLC (gradient A): t_r = 0.68 min; ESI-MS: m/z calcd for $C_{47}H_{82}N_{12}O_{10}S$ $[M+2H]^{2+}$ 504.31, found 504.28.

- *Cys-P3*



Cys-P3

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.29 (s, formate), 7.43-7.10 (m, 10H), 4.56-4.45 (m, 2H), 4.44-4.31 (m, 1H), 4.25-4.13 (m, 1H), 3.36-3.27 (m, 0.5H), 3.19-2.85 (m, 7.5H), 1.68-1.49 (m, 4H), 1.37-1.31 (m, 3H), 1.28-1.11 (m, 2H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 173.20, 173.07, 168.39, 168.24, 137.07, 137.05, 136.61, 136.59, 129.95, 129.90, 129.86, 129.51, 129.49, 129.47, 127.95, 55.70, 55.60, 55.46, 55.44, 54.70, 53.84, 53.82, 52.17, 50.37, 50.26, 39.83, 39.80, 37.99, 37.93, 37.80, 37.71, 37.65, 31.37, 31.26, 26.97, 26.94, 25.76, 22.46, 22.43, 17.55, 17.16 ppm; UPLC (gradient A): t_r = 0.82 min; ESI-MS: m/z calcd for $C_{30}H_{43}N_7O_5S$ $[M+H]^+$ 614.31, found 614.05.

b. Chapter 3

Peptides incorporating the dynamic unit N-methyl-cysteine (Daa2)

Peptide synthesis was performed on a Fmoc-Rink-amide-MBHA resin (0.74 mmol/g, 0.5 mmol scale) using a combination of conventional and microwave-assisted SPPS, according to the procedure described above (for the synthesis of peptides incorporating Daa1) with few exceptions.

Swelling of the resin (in the microwave synthesizer): see above.

Fmoc deprotection (in the microwave synthesizer): see above.

Amino acid coupling (in the microwave synthesizer): see above.

Sulfonylation (in the parallel synthesizer): see above.

Methylation (in the parallel synthesizer): methyl iodide (10 mmol, 622 µL) in a 1 M solution of TBAF in THF (5 mL) was added to the resin under argon and let shake for 1 h. The reaction was repeated a second time with a fresh solution.

Removal of NBS (in the parallel synthesizer): see above

For peptides **P1-Daa2-P2** and **P1-Daa2-P3**:

- Fmoc-Gly-OH (**P1**) coupling (in the microwave synthesizer): a solution of

Fmoc-Gly-OH (2.0 mmol, 595 mg), HATU (2.0 mmol, 760 mg) and DIEA (2.7 mmol, 470 µL) in DMF (8 mL) was added to the microwave reaction vessel; the coupling reaction was performed at 75 °C for 20 min at 30 W and repeated a second time with a fresh solution.

- Successive Fmoc deprotection and amino acid coupling steps (in the microwave synthesizer) of Fmoc-Lys(Boc)-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Leu-OH.

For peptides **P1'-Daa2-P2** and **P1''-Daa2-P2**:

- Fmoc-Val-OH (**P1'**) or Fmoc-Lys(Boc)-OH (**P1''**) coupling (in the microwave synthesizer): a solution of Fmoc-aa-OH (Fmoc-Val-OH: 4.0 mmol, 1.36 g; Fmoc-Lys(Boc)-OH: 4.0 mmol, 1.87 g), HATU (4.0 mmol, 1.52 g) and DIEA (5.5 mmol, 958 µL) in DMF (8 mL) was added to the microwave reaction vessel; the coupling reaction was performed at 75 °C for 20 min at 30 W and repeated a second time with a fresh solution.
- Successive Fmoc deprotection and amino acid coupling steps (in the microwave synthesizer) of Fmoc-Lys(Boc)-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Leu-OH.

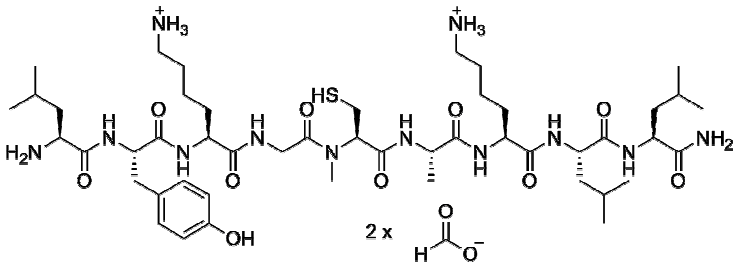
Cleavage (in the parallel synthesizer): see above.

Oxidation: none, the peptides were purified in their reduced form.

Purification: see above.

In order to prevent oxidation during the NMR acquisition, we lowered the pH of D_2O by adding an excess of formic acid. However, oxidation still occurred in some cases, resulting in additional peaks in the ^1H and ^{13}C NMR spectra.

- *P1-Daa2-P2*

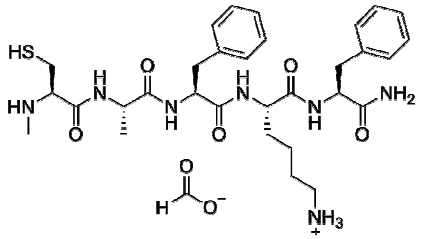


P1-Daa2-P2

^1H NMR (D_2O, 500 MHz, 25 °C): δ = 8.29 (s, formate), 7.15 (d, *J* = 8.1 Hz, 2H), 6.81 (d, *J* = 8.1 Hz, 2H), 5.04 (dd, *J* = 8.3 Hz, 6.4 Hz, 1H), 4.63-4.57 (m, 1H), 4.40-4.22 (m, 5H), 4.08-3.97 (m,

2H), 3.27-2.11 (m, 12H), 1.89-1.53 (m, 18H), 1.53-1.30 (m, 6H) 1.02-0.81 (m, 18H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 177.99, 175.72, 175.14, 174.37, 173.68, 172.97, 171.63, 171.48, 170.73, 155.33, 131.33, 128.13, 116.22, 60.72, 56.22, 54.25, 53.65, 53.04, 52.77, 52.34, 50.90, 40.57, 40.41, 40.16, 39.87, 39.81, 31.54, 31.29, 30.95, 26.98, 26.92, 25.03, 25.01, 24.52, 22.94, 22.79, 22.54, 22.27, 22.02, 21.43, 21.20, 17.06 ppm; UPLC (gradient A): t_r = 0.72 min; ESI-MS: m/z calcd for $C_{48}H_{84}N_{12}O_{10}S$ $[M+2H]^{2+}$ 511.32, found 511.67.

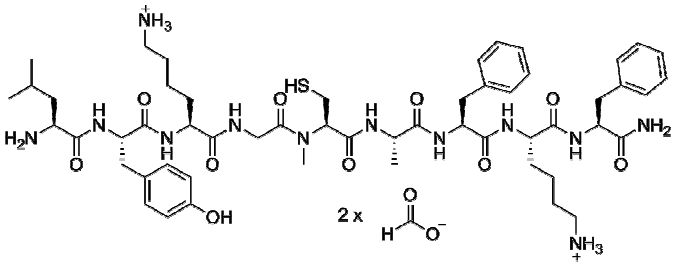
- *Daa2-P3*



Daa2-P3

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.25 (s, formate), 7.43-7.13 (m, 10H), 4.56-4.49 (m, 2H), 4.36 (q, *J* = 7.0 Hz, 1H, 4.21 (dd, *J* = 7.5 Hz, 6.6 Hz, 1H), 4.09 (dd, *J* = 6.2 Hz, 4.6 Hz, 1H), 3.19-3.08 (m, 2H), 3.05-2.89 (m, 6H), 2.72 (s, 3H, 3 H-C(1)), 1.68 -1.52 (m, 4H), 1.35 (d, ^3J = 7.0 Hz, 3H), 1.30-1.15 (m, 2H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 172.50, 172.44, 166.65, 136.28, 135.80, 129.17, 129.06, 128.71, 128.70, 127.16, 61.70, 54.92, 54.66, 53.03, 49.64, 39.03, 37.01, 36.85, 31.37, 30.47, 26.15, 23.86, 21.66, 16.35 ppm; UPLC (gradient B): t_r = 0.88 min; ESI-MS: m/z calcd for $C_{31}H_{45}N_7O_5S$ $[M+H]^+$ 628.33, found 628.12.

- *P1-Daa2-P3*

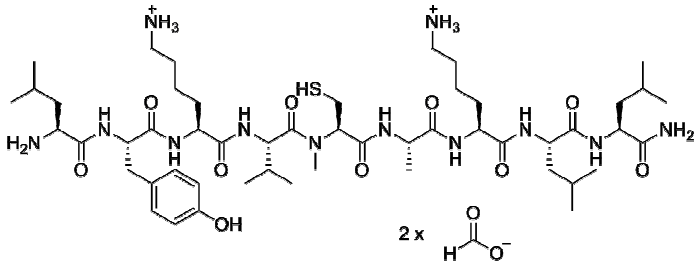


P1-Daa2-P3

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.28 (s, formate), 7.41-7.22 (m, 8H), 7.21-7.08 (m, 4H), 6.86-6.75 (d, *J* = 8.6 Hz, 2H), 4.96 (dd, *J* = 9.0 Hz, 6.3 Hz, 1H), 4.59 (dd, *J* = 9.9 Hz, 6.6 Hz, 1H), 4.57-4.48 (m, 2H), 4.39-4.14 (m, 3H), 4.11-3.94 (m, 2H), 3.23-2.68 (m, 16H), 1.85-1.47

(m, 10H), 1.46-1.11 (m, 8H), 1.03-0.86 (m, 6H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 176.21, 175.32, 173.67, 173.52, 173.43, 172.96, 171.64, 171.43, 170.72, 155.32, 137.11, 136.64, 131.33, 129.95, 129.91, 129.50, 129.48, 128.15, 127.94, 127.91, 116.22, 60.96, 56.21, 55.53, 55.36, 54.01, 53.64, 52.34, 50.75, 42.14, 40.57, 39.83, 39.80, 37.60, 37.49, 31.77, 31.29, 31.11, 26.94, 26.92, 24.52, 22.94, 22.54, 22.52, 22.27, 22.02, 16.84 ppm; UPLC (gradient B): t_r = 0.97 min; ESI-MS: m/z calcd for $C_{54}H_{80}N_{12}O_{10}S$ $[M+2H]^{2+}$ 545.34, found 545.79.

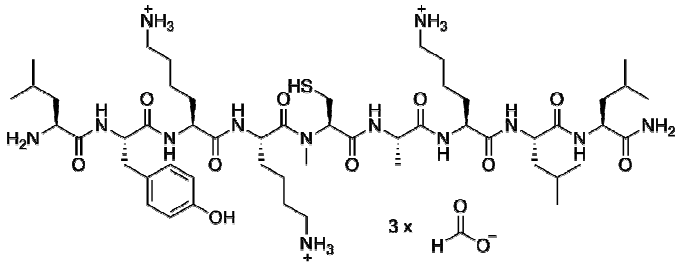
- *P1'-Daa2-P2*



P1'-Daa2-P2

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.49 (s, formate), 7.17 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.06 (dd, *J* = 8.6 Hz, 6.9 Hz, 1H), 4.67-4.62 (m, 1H), 4.53 (d, 1H, *J* = 8.5 Hz), 4.45-4.25 (m, 6H), 3.88 (t, *J* = 6.7 Hz, 1H), 3.24 (s, 3H), 3.10-2.91 (m, 8H), 2.14-2.05 (m, 1H), 1.87-1.79 (m, 2H), 1.78-1.59 (m, 15H), 1.51-1.36 (m, 6H), 1.08-0.89 (m, 24H) ppm; ^{13}C NMR (D_2O, 125 MHz, 25 °C): δ = 178.00, 175.72, 175.22, 175.08, 174.48, 173.00, 172.77, 171.50, 155.25, 131.16, 128.26, 116.17, 60.74, 56.22, 55.83, 54.45, 53.10, 52.79, 52.69, 50.89, 40.44, 40.17, 39.87, 32.83, 30.90, 27.00, 25.02, 22.98, 22.80, 22.62, 22.46, 21.98, 21.62, 21.47, 21.17, 18.69 ppm; UPLC (gradient A): t_r = 0.84 min; ESI-MS: m/z calcd for $C_{51}H_{90}N_{12}O_{10}S$ $[M+2H]^{2+}$ 532.34, found 532.67.

- *P1"-Daa2-P2*



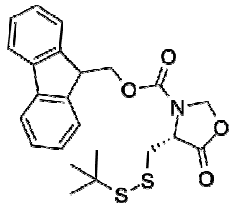
P1"-Daa2-P2

^{1}H NMR (D_2O, 500 MHz, 25 °C): δ = 8.22 (s, formate), 7.14-7.04 (m, 2H), 6.83-6.73 (m, 2H),

4.94-4.86 (m, 2H), 4.67-4.18 (m, 7.5H), 4.08-3.90 (m, 2H), 3.41 (d, *J* = 6.1 Hz, 1H), 3.20-3.10 (m, 1.5H), 3.10-2.78 (m, 9H), 2.77-2.64 (m, 2H), 2.07-1.87 (m, 1H), 1.87-1.50 (m, 19H), 1.49-1.24 (m, 8H), 1.00-0.74 (m, 16H) ppm; ^{13}C NMR (D_2O, 500 MHz, 25 °C): δ = 177.96, 177.89, 175.58, 175.10, 174.95, 174.50, 174.29, 174.20, 174.16, 173.01, 172.67, 171.35, 170.78, 170.76, 167.00, 163.87, 163.59, 155.28, 155.27, 131.20, 131.17, 120.53, 118.21, 116.18, 116.11, 115.89, 113.57, 52.98, 52.75, 52.33, 40.57, 40.47, 40.44, 39.86, 39.81, 32.50, 27.28, 27.13, 27.01, 25.06, 25.02, 24.49, 22.96, 22.92, 22.89, 22.86, 22.83, 22.78, 22.59, 22.37, 22.35, 21.89, 21.60, 21.47, 21.32, 21.22, 17.49, 17.21 ppm; UPLC (gradient B): t_r = 0.65 min; ESI-MS: m/z calcd for $C_{58}H_{89}N_{13}O_{10}S$ $[M+2H]^{2+}$ 546.85, found 547.26.

c. Chapter 4

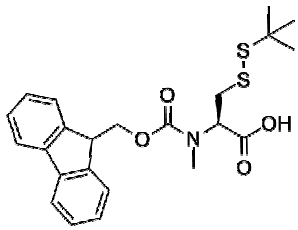
Compound 2



2

To a room temperature solution of Fmoc-L-Cys(StBu)-OH (5.0 g, 12 mmol) in toluene (117 mL) was added p-formaldehyde (1.80 g, 60 mmol) and camphorsulfonic acid (185 mg, 0.81 mmol). The reaction was heated in a Dean-Stark apparatus under reflux (oil bath at 120 °C). After 18 h the solvent was evaporated under vacuum and the crude oil was purified by flash column chromatography (SiO_2, cyclohexane/EtOAc 10:1 → 4:1) to afford compound **2** (5.02 g, yield 97%) as a white viscous gum. R_f (cyclohexane/EtOAc 4:1) = 0.5; ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 7.78 (d, *J* = 7.5 Hz, 2H), 7.57 (d, *J* = 7.5 Hz, 2H), 7.45-7.31 (m, 4H), 5.52-5.18 (br m, 2H), 4.77-4.34 (br m, 2.5H), 4.32-4.19 (br m, 1H), 4.05 (br s, 0.5H), 3.55 (br s, 0.5H), 3.22 (br s, 0.5H), 3.00 (br s, 0.5H), 2.72 (br s, 0.5H), 1.29 (s, 9H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 170.62, 151.95, 143.24, 141.14, 127.78, 127.06, 124.58, 119.92, 78.31, 77.20, 67.65, 55.24, 48.20, 46.86, 29.44 ppm; ESI-MS: m/z calcd for $C_{23}H_{25}NO_4S_2$ $[M+H]^+$ 444.13, found 444.38.

Compound 3



3

To a stirring solution of compound **2** (900 mg, 2.0 mmol) in $CHCl_3$ (10 mL) was added TES (3.3 mL, 20.3 mmol) and TFA (5 mL, 67.3 mmol). The reaction was stirred at room temperature for 18 h and then concentrated under vacuum. The resulting oil was then dissolved in CH_2Cl_2 and concentrated under vacuum (3 x 20 mL). Purification by flash column chromatography (SiO_2, cyclohexane/EtOAc 10:1 → 4:1) afforded compound **3** (700 mg, 77%) as a white solid in a 2:1 conformer ratio. R_f = 0.32 (EtOAc); ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 7.81-7.73 (m, 2H), 7.66-7.55 (m, 2H); 7.45-7.36 (m, 2H), 7.35-7.28 (m, 2H), 4.74 (dd, *J* = 10.4, 4.7, 0.66H), 4.68 (dd, *J* = 10.4, 4.8, 0.33H), 4.61-4.41 (m, 2H), 4.31 (t, *J* = 7.0, 0.66H), 4.25 (t, *J* = 5.5, 0.33H), 3.34 (dd, *J* = 14.0, 4.6, 0.66H), 3.18 (dd, *J* = 14.0, 10.6, 0.66H), 3.11-3.07 (m, 0.33H), 3.05 (s, 2H), 2.96 (s, 1H), 2.69 (dd, *J* = 14.0, 10.24, 0.33H), 1.36 (s, 6H), 1.33 (s, 3H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 175.33, 175.28, 156.62, 156.10, 143.81, 143.71, 141.26, 127.83, 127.65, 127.05, 125.06, 124.84, 124.79, 119.91, 68.00, 67.84, 59.91, 58.81, 48.19, 47.11, 47.02, 39.23, 38.94, 33.87, 33.07, 29.88; ESI-MS: m/z calcd for $C_{23}H_{27}NO_4S_2$ $[M+2H]^{2+}$ 446.15, found 446.20.

Affibody helices

Peptide synthesis was performed on a ChemMatrix-Rink-amide resin (0.51 mmol/g, 0.1 mmol scale) using microwave-assisted SPPS.

Swelling of the resin: the resin was let swell in DMF (10 mL) for 30 min.

Fmoc deprotection: two successive treatments of the resin with a 5 %w/v piperidine solution in DMF (10 mL) containing 0.1 M HOBt at 70 ºC and 55 W for 3 min.

Amino acid coupling: ten-fold molar excess of each Fmoc-aa-OH (5 mL of a 0.2 M solution in DMF), HBTU (2 mL of a 0.5 M solution in DMF) and DIEA (1 mL of a 2 M solution in NMP) at 70 ºC and 35 W (microwave power) for 5 min. Double couplings were performed in the case of Fmoc-Arg(Pbf)-OH and Fmoc-His(Trt)-OH was coupled for 2 min without microwave heating and then at 50 ºC with 25 W for 4 min.

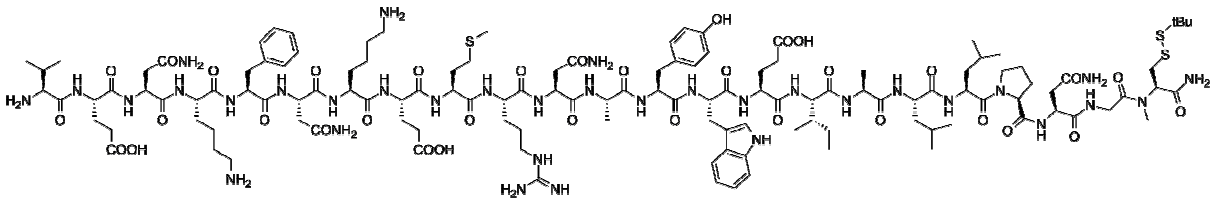
Coupling of compound **3**: a solution of **3** (178 mg, 0.4 mmol), HATU (152 mg, 0.4 mmol) and DIEA (94 µL, 0.54 mmol) in DMF (4 mL) was added to the resin. The coupling was performed and repeated a second time.

Coupling of Fmoc-Gly-OH: a solution of Fmoc-Gly-OH (297 mg, 1 mmol), HATU (380 mg, 1 mmol) and DIEA (235 µL, 1.1 mmol) in DMF (4 mL) was added to the microwave reaction vessel; the coupling reaction was performed at 75 °C for 20 min at 30 W and repeated a second time with a fresh solution.

Cleavage (in the parallel synthesizer): the resin was suspended in a 5 mL solution of TFA/phenol/H_2O/TIS: 88/5/5/2 vol% for 2 h, then filtered and washed with TFA (5 mL). The combined filtrates were concentrated *in vacuo* and then added to cold diethylether (200 mL). The precipitate was separated by centrifugation, washed with cold diethylether (2 x 40 mL) and dried under argon to afford crude peptides.

Purification: preparative HPLC using either 0.1 %v/v formic acid in water (solvent A) and 0.1 %v/v formic acid in methanol (solvent B) as the two mobile phases on a Waters *SunFire Prep C18 QBD* 5 µm 19x150 mm column or using 0.05 %v/v ammonia in water (solvent C) and 0.05 %v/v ammonia in methanol (solvent D) as the two mobile phases on a Waters *XBridge Prep C18 QBD* 5 µm 19x150 mm column. Fractions were automatically collected using the ESI-MS signal of the peptide as a trigger and then individually analyzed via UPLC-ESI. The pure fractions were then combined and the organic solvent removed under vacuum prior to lyophilization.

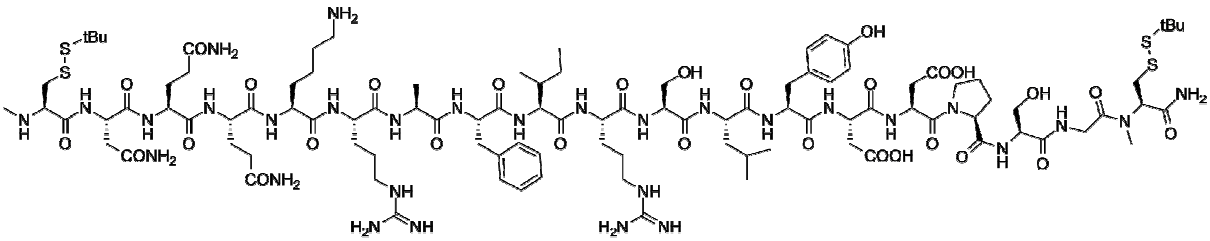
- *Helix 1 Her2*



H1(StBu)-Her2

UPLC (gradient A): t_r = 1.25 min; ESI-MS: m/z calcd for $C_{127}H_{198}N_{34}O_{34}S_3$ $[M+3H]^{3+}$ 947.81, found 948.43; MALDI-TOF: m/z calcd for $C_{127}H_{198}N_{34}O_{34}S_3$ $[M+H]^+$ 2839.40 found: 2840.44.

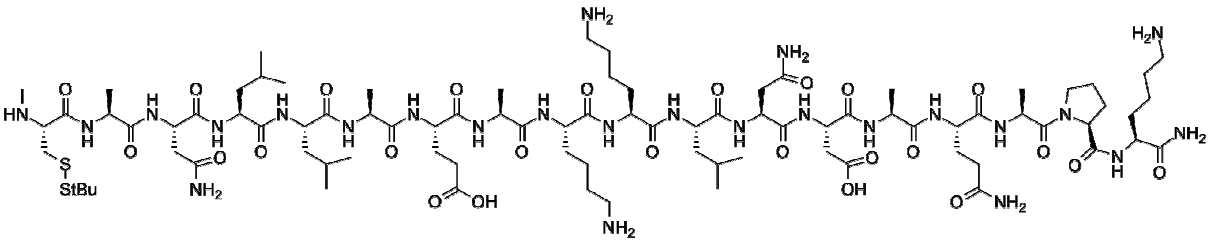
- *Helix 2 Her2*



H2(StBu)$_2$-Her2

UPLC (gradient A): t_r = 0.96 min; ESI-MS: m/z calcd for $C_{102}H_{166}N_{30}O_{29}S_4$ $[M+3H]^{3+}$ 802.39, found 802.73; MALDI-TOF: m/z calcd for $C_{102}H_{166}N_{30}O_{29}S_4$ $[M+H]^+$ 2403.13, found 2404.35.

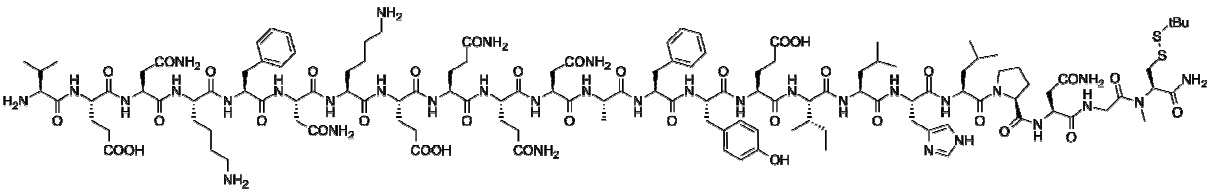
- *Helix 3*



H3(StBu)

UPLC (gradient B): t_r = 0.90 min; ESI-MS: m/z calcd for $C_{86}H_{151}N_{25}O_{25}S_2$ $[M+2H]^{2+}$ 1000.55, found 1000.40; MALDI-TOF: m/z calcd for $C_{86}H_{151}N_{25}O_{25}S_2$ $[M+H]^+$ 1998.08, found 1999.02.

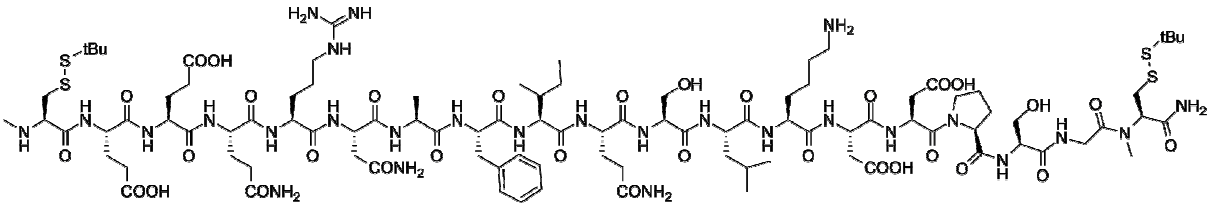
- *Helix 1 IgG*



H1(StBu)-IgG

UPLC (gradient B): t_r = 1.25 min; ESI-MS: m/z calcd for $C_{127}H_{194}N_{34}O_{36}S_2$ $[M+2H]^{2+}$ 1419.20, found 1419.80.

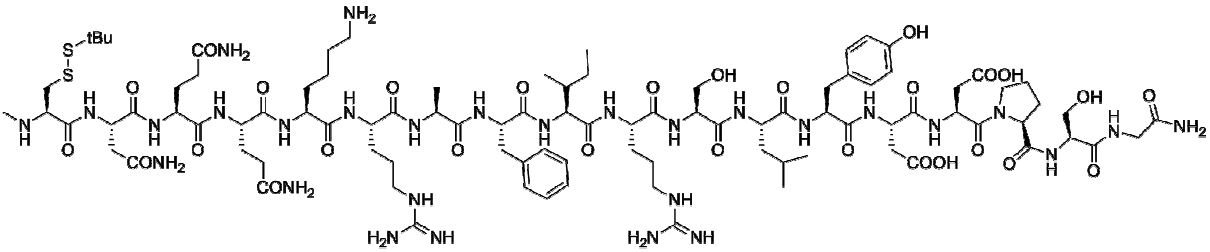
- *Helix 2 IgG*



H2(StBu)$_2$-IgG

UPLC (gradient B): t_r = 1.05 min; ESI-MS: m/z calcd for $C_{97}H_{159}N_{27}O_{32}S_4$ $[M+2H]^{2+}$ 1172.54, found 1173.68.

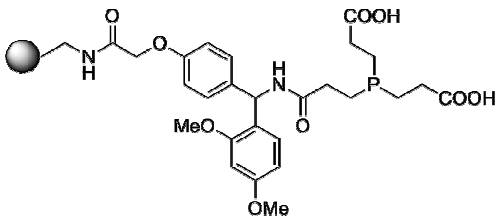
- *Helix 2a Her2*



H2a-Her2

UPLC (gradient A): t_r = 0.74 min; ESI-MS: m/z calcd for $C_{94}H_{151}N_{29}O_{28}S_2$ $[M+2H]^{2+}$ 1100.25, found 1101.25.

Compound 4



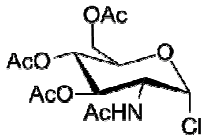
4

H-Rink Amide ChemMatrix resin (514 mg, loading 0.51 mmol/g) was swollen with CH_2Cl_2 (20 mL) for 30 min and then DMF (5 mL) for 30 min. A solution of HBTU (99 mg, 262 mmol) in DMF (3 mL) was slowly added to a stirred solution of DIEA (87 µL, 524 mmol) and TCEP·HCl (75 mg, 262 mmol) in DMF (4 mL). After 5 min this solution was added to the swelled resin and the final mixture shaken for 24 h on an orbital shaker. The resin was washed several times with DMF (3 x 5 mL), HCl 1M (5 mL), phosphate 200 mM/EDTA 0.2 M buffer at pH 7.4 (5

mL), methanol (3 x 5 mL), diethyl ether (5 mL) and CH_2Cl_2 (2 x 5 mL) and dried under vacuum. IR (ATR): 3557, 2865, 1731, 1655, 1611, 1588, 1538, 1508, 1455, 1348, 1296, 1246, 1209, 1902, 946, 840 cm^{-1}.

d. Chapter 5

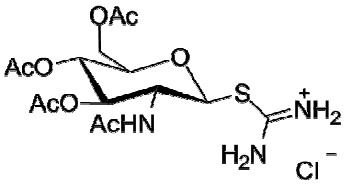
Compound 5



5

N-acetyl-D-glucosamine (5.0 g, 22.6 mmol) was suspended in acetyl chloride (30 mL) and left stirring at room temperature under argon. After 24 h the solution was diluted with CH_2Cl_2 (100 mL) and poured into ice water (100 mL). After phase separation, the aqueous phase was re-extracted with CH_2Cl_2 (3 x 50 mL). The combined organic phases were then washed with a saturated solution of $NaHCO_3$ (100 mL), dried over $MgSO_4$, filtered and concentrated under vacuum. Purification by flash column chromatography (SiO_2, EtOAc/cyclohexane 7:3), afforded compound **5** (3.38 g, 47%) as a white solid. R_f (EtOAc/cyclohexane 7:3) = 0.43; ^1H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 6.19 (d, *J* = 3.7 Hz, 1H), 5.87-5.78 (m, 1H), 5.36-5.29 (m, 1H), 5.22 (t, *J* = 9.8 Hz, 1H), 4.58-4.50 (m, 1H), 4.33-4.24 (m, 2H), 4.17-4.11 (m, 1H), 2.11 (s, 3H), 2.06 (s, 6H), 1.99 (s, 3H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 171.4, 170.5, 170.1, 169.1, 93.6, 70.9, 70.1, 66.9, 61.1, 53.4, 23.0, 20.6, 20.5 ppm.

Compound 6

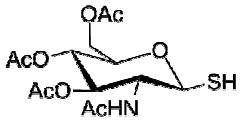


6

Compound **5** (3.79 g, 10.36 mmol) and thiourea (1.34 g, 17.62 mmol) were dissolved in dry acetone (57 mL) under argon. The reaction was heated to reflux (60 °C, oil bath). After 1h the white precipitate was filtered and washed with acetone. The filtrate was returned to reflux and

the process repeated 4 more times (45 min each time) to afford compound **6** (4.35 g, 95%) as a white solid. ^{1}H NMR (DMSO-d6, 400 MHz, 25 °C): δ = 9.44 (s, 2H), 9.24 (s, 2H), 8.43 (d, *J* = 9.3 Hz, 1H), 5.70 (d, *J* = 10.4 Hz, 1H), 5.09 (t, *J* = 9.7 Hz, 1H), 4.90 (t, *J* = 9.8 Hz, 1H), 4.26-4.09 (m, 2H), 4.06-3.90 (m, 2H), 1.98 (s, 3H), 1.95 (s, 3H), 1.90 (s, 3H), 1.77 (s, 3H) ppm; ^{13}C NMR (DMSO-d6, 100 MHz, 25 °C): δ = 170.0, 169.8, 169.6, 169.3, 80.6, 74.7, 72.7, 67.9, 61.5, 51.2, 22.5, 20.6, 20.4, 20.3 ppm.

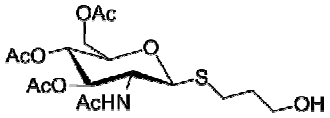
Compound 7



7

To a stirring mixture of CH_2Cl_2 (30 mL) and H_2O (20 mL) were added compound **6** (4.13 g, 9.34 mmol) and $Na_2S_2O_4$ (2.13 g, 11.18 mmol) under argon. The resulting mixture was heated under reflux (45 °C, oil bath) for 2.5 h and then brought to room temperature. The phases were separated, and the aqueous phase extracted with CH_2Cl_2 (2 x 50 mL). The combined organic phases were then washed with H_2O (1 x 50 mL) and brine (1 x 50 mL), dried over $MgSO_4$, filtered and concentrated under vacuum to afford compound **7** (2.86 g, 84%) as a white solid. R_f (EtOAc/EtOH 9:1) = 0.8; ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 5.78-5.61 (m, 1H), 5.16-5.05 (m, 2H), 4.59 (t, *J* = 9.7 Hz, 1H), 4.24 (dd, *J* = 12.5 Hz, 4.7 Hz, 1H), 4.17-4.08 (m, 2H), 3.74-3.64 (m, 1H), 2.57 (d, *J* = 9.4 Hz, 1H), 2.10 (s, 3H), 2.05 (s, 3H), 2.03 (s, 3H), 1.99 (s, 3H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 171.2, 170.8, 170.4, 169.2, 80.3, 76.2, 73.4, 68.0, 62.1, 56.8, 23.3, 20.8, 20.7, 20.6 ppm.

Compound 8

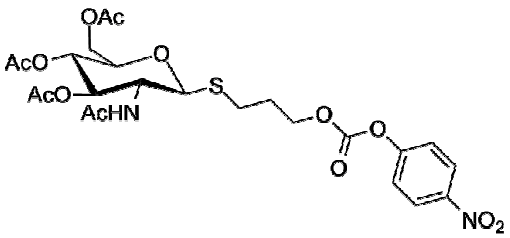


8

A solution of compound **7** (100 mg, 0.28 mmol) and AIBN (30 mg, 0.18 mmol) in a 1:1 allyl alcohol/$CHCl_3$ mixture (2.8 mL) was heated at reflux (50 °C, oil bath) under argon for 2 h. The solvent was then removed under vacuum and purification by flash column chromatography (SiO_2, EtOAc/EtOH 19:1) afforded compound **8** (97 mg, 84%) as a white solid. R_f (EtOAc) =

0.2; $[\alpha]_D^{20}$ = -81.4 (c = 0.5 in CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 5.92 (d, *J* = 9.5 Hz, N-*H*, 1H), 5.15 (t, *J* = 9.5 Hz, 1H, *H-3*), 5.07 (t, *J* = 9.7 Hz, 1H, *H-4*), 4.55 (d, *J* = 10.4 Hz, 1H, *H-1*), 4.25-4.10 (m, 3H, *H-2*, *H-6*, *H-6'*), 3.77-3.65 (m, 3H, *H-5*, -C*H*$_2$-OH), 2.92-2.78 (m, 2H, -S-C*H*$_2$-), 2.08, 2.03, 2.03, 1.98 (4 x s, 12H, -COC*H*$_3$), 1.92-1.70 (m, 2H, -C*H*$_2$-) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 171.33, 170.96, 170.78, 169.51 (4 x -*C*OCH$_3$), 84.62 (*C-1*), 76.18 (*C-5*), 73.88 (*C-3*), 68.47 (*C-4*), 62.38 (*C-6*), 60.44 (-*C*H$_2$OH), 53.08 (*C-2*), 31.89 (-*C*H$_2$-), 25.99 (-S-*C*H$_2$-), 23.43, 20.93, 20.87, 20.78 (4 x -CO*C*H$_3$) ppm; ESI-MS: m/z calcd for $C_{17}H_{27}NO_9S$ $[M+H]^+$ 422.15, found 422.44.

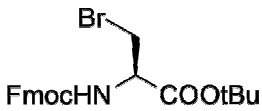
Compound 9



9

Compound **8** (80 mg, 0.19 mmol), p-nitrobenzenchloroformate (57 mg, 0.28 mmol) and DMAP (4.6 mg, 0.04 mmol) were suspended in dry CH_2Cl_2 (4 mL) under argon. Pyridine (30 µL, 0.38 mmol) was then added dropwise and the reaction was left stirring at room temperature. After 2 h the reaction was diluted with CH_2Cl_2 (50 mL) and washed with HCl 5% (1 x 10 mL), NH_4Cl ss (2 x 10 mL) and brine (2 x 10 mL). The organic phase was then dried over $MgSO_4$, filtered and concentrated under vacuum. Purification by flash column chromatography (SiO_2, EtOAc/EtOH 4:1) afforded compound **9** (110 mg, 99%) as white solid. R_f (EtOAc/EtOH 19:1) = 0.68; $[\alpha]_D^{20}$ = -58.4 (c = 0.5 in CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 8.28 (d, *J* = 9.3 Hz, 2H, Ar-*H*), 7.39 (d, *J* = 9.3 Hz, 2H, Ar-*H*), 5.60 (d, *J* = 9.2 Hz, 1H, N-*H*), 5.18 (t, *J* = 9.5 Hz, 1H, *H-3*), 5.11 (t, *J* = 9.7 Hz, 1H, *H-4*), 4.59 (d, *J* = 10.3 Hz, 1H, *H-1*), 4.39 (t, *J* = 6.2 Hz, 2H, -C*H*$_2$-O-), 4.26-4.04 (m, 3H, *H-2*, *H-6*, *H-6'*), 3.74-3.64 (m, 1H, *H-5*), 2.96-2.72 (m, 2H, -S-C*H*$_2$-), 2.18-1.68 (m, 14H, -C*H2*-, 4 x -COC*H*$_3$) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 171.01, 170.72, 170.41, 169.42 (4 x –*C*OCH3), 155.60 (*C*(Ar)-O-), 152.48 (-O-*C*O-O-), 145.41 (*C*(Ar)-NO$_2$), 125.20, 121.72 (2 x *C*(Ar)-H), 84.48 (*C-1*), 75.93 (*C-5*), 73.67 (*C-3*), 68.54 (*C-4*), 67.80 (-*C*H$_2$-O-), 62.37 (*C-6*), 53.13 (*C-2*), 28.74 (-S-*C*H$_2$-), 26.24 (-*C*H2-), 23.28, 20.78, 20.74, 20.67 (4 x –CO*C*H$_3$) ppm; ESI-MS: m/z calcd for $C_{24}H_{30}N_2O_{13}S$ $[M+H]^+$ 587.15, found 587.45.

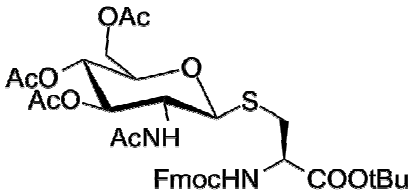
Compound 10



10

PPh_3 (137 mg, 0.52 mmol) was added portionwise to a stirring solution of Fmoc-Ser-OtBu (100 mg, 0.26 mmol) and CBr_4 (147 mg, 0.44 mmol) in dry CH_2Cl_2 (2.7 mL) at 0 °C. After 20 min the mixture was concentrated under vacuum and purified via flash column chromatography (SiO_2, cyclohexane/EtOAc 4:1) to afford compound **10** (60 mg, 52%) as a white solid. R_f (cyclohexane/EtOAc 4:1) = 0.4; ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 7.79 (d, *J* = 7.5 Hz, 2H), 7.63 (d, *J* = 7.4 Hz, 2H), 7.45-7.39 (m, 2H), 7.38-7.31 (m, 2H), 5.78 (d, *J* = 7.2 Hz, 1H), 5.75-4.68 (m, 1H), 4.48-4.35 (m, 2H), 4.26 (t, *J* = 7.2 Hz, 1H), 3.81 (dd, *J* = 10.5, 3.0 Hz, 1H), 3.78 (dd, *J* = 10.5, 3.3 Hz, 1H), 1.54 (s, 9H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 167.11, 155.48, 143.72, 143.61, 141.24, 127.70, 127.04, 119.95, 83.52, 67.26, 54.46, 47.01, 34.31, 27.90 ppm.

Compound 11



11

To a vigorously stirred solution of **7** (119 mg, 0.33 mmol) and **10** (115 mg, 0.26 mmol) in EtOAc (5.2 mL), was added a $NaHCO_3$ solution (pH 8.7, 5.2 mL), followed by tetrabutylammonium hydrogen sulfate (350 mg, 1.03 mmol). After 5 h the solution was diluted with CH_2Cl_2 (60 mL), washed with saturated aqueous $NaHCO_3$ (3 x 20 mL) and brine (2 x 10 mL). The organic phase was dried over $MgSO_4$, filtered and concentrated in vacuo. Purification by flash column chromatography chromatography (SiO_2, cyclohexane/EtOAc 2:1 → 1:1 → 2:3)afforded compound **11** (180 mg, 96%) as a white solid. R_f (cyclohexane/EtOAc 2:3) = 0.35; ^{1}H NMR ($CDCl_3$, 400 MHz, 25 °C): δ = 7.76 (d, *J* = 7.5 Hz, 2H), 7.66-7.57 (m, 2H), 7.44-7.36 (m, 2H), 7.36-7.28 (m, 2H), 6.02-5.89 (m, 2H), 5.16 (t, *J* = 9.7 Hz, 1H), 5.07 (t, *J* = 9.7 Hz, 1H), 4.69 (d, *J* = 10.4 Hz, 1H), 4.51-4.40 (m, 2H) 4.39-4.29 (m, 1H), 4.28-4.02 (m, 4H), 3.70-3.62 (m, 1H), 3.30 (dd, *J* = 14.1, 2.7 Hz, 1H), 2.98-2.81 (m, 1H), 2.04 (s, 3H), 2.02 (s, 3H), 2.01 (s,

3H), 1.87 (s, 3H), 1.47 (s, 9H) ppm; ^{13}C NMR ($CDCl_3$, 100 MHz, 25 °C): δ = 170.81, 170.58, 170.33, 169.33, 169.18, 151.06, 143.70, 143.56, 141.14, 127.66, 127.04, 125.03, 124.94, 119.91, 83.36, 82.73, 75.91, 73.66, 68.29, 66.99, 62.14, 53.85, 52.81, 46.98, 27.85, 22.98, 20.55, 20.47 ppm; ESI-MS: m/z calcd for $C_{36}H_{44}N_2O_{12}S$ $[M+H]^+$ 729.27, found 729.44.

6. Exchange reaction protocols

a. Chapters 2 and 3

Phosphate buffers were prepared at pH 6, 7, 8 and 9: the amount of NaH_2PO_4 required to give a 0.2 M solution in 20 mL of water (551 mg) was dissolved in water (10 mL). The resulting acidic solution was titrated to pH 6, 7, 8, or 9 using 2 M NaOH and its volume adjusted to 20 mL. Buffers were degassed using freeze/thaw cycles and placed under argon prior to use.

Exchange reaction between peptides incorporating N-(2-thioethyl)-cysteine (Daa1)

Setup. Each of the two peptides (formate salts) was dissolved in water to give a 5 mM solution. 250 µL of each of these two solutions were mixed, together with 250 µL of a TCEP solution (2 eq., 10 mM, 7.1 mg in 2.5 mL of water neutralized with 8.4 mg of $NaHCO_3$) and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides and TCEP was then diluted under argon with 500 µL of 0.2 M phosphate buffer (pH 6, 7, 8 or 9) containing 25 mM DTT (5.8 mg in 1.5 mL) and left stirring at room temperature.

Analysis. A 1 mL solution containing H_2O_2 and an internal UV standard (3,5-dimethoxybenzoic acid) was prepared as follows: 23 µL H_2O_2 (30 % solution), 20 µL UV standard solution (5 mM in a 0.2 M pH 7 phosphate buffer) and 957 µL water. For each time point an aliquot of the reaction (25 µL) was diluted with 175 µL of this solution, then filtered and injected in the UPLC-ESI (2 x 2 µL injections).

Control exchange reaction between peptides incorporating cysteine (Cys)

Setup. **Cys-P3** (formate salt) and **P1-Cys-P2** (diformate salt) were dissolved in water to give two 5 mM solutions. 250 µL of each of these two solutions were mixed together and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides was diluted under argon with 500 µL of 0.2 M phosphate buffer (pH 7) containing 25 mM of DTT (5.8 mg in 1.5 mL) and left stirring at room temperature.

Analysis. A 1 mL solution containing TCEP and an internal UV standard (3,5-dimethoxybenzoic acid) was prepared as follows: 960 µL of a 10 mM TCEP solution in water and 40 µL UV standard solution (5 mM in a 0.2 M pH 7 phosphate buffer) . For each time point an aliquot of the reaction (25 µL) was diluted with 175 µL of this solution, then filtered and injected in the UPLC-ESI (2 x 2 µL injections).

Exchange reaction between peptides incorporating N-methyl-cysteine (Daa2)

Setup.

- *Between P1-Daa2-P2 and Daa2-P3*

  P1-Daa2-P2 (formate salt) and **Daa2-P3** (formate salt) were dissolved in water to give two 5 mM solutions. 250 µL of each of these two solutions were mixed together and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides was diluted under argon with 500 µL of 0.2 M phosphate buffer (pH 6, 7 or 8) containing 25 mM of DTT (5.8 mg in 1.5 mL) and left stirring at room temperature or at 37 °C.

- *Between P1'-Daa2-P2 and Daa2-P3*

  P1'-Daa2-P2 (formate salt) and **Daa2-P3** (formate salt) were dissolved in water to give two 5 mM solutions. 250 µL of each of these two solutions were mixed together and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides was diluted under argon with 250 µL of 0.2 M phosphate buffer (pH 7) containing 125 mM of DTT (28.9 mg in 1.5 mL) and left stirring at 37 °C.

- *Between P1"-Daa2-P2 and P1-Daa2-P3*

  P1"-Daa2-P2 (formate salt) and **P1-Daa2-P3** (formate salt) were dissolved in water to give two 5 mM solutions. 750 µL of each of these two solutions were mixed together and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides was diluted under argon with 500 µL of 0.2 M phosphate buffer (pH 8) containing 50 mM of DTT (11.6 mg in 1.5 mL) and left stirring at 37 °C.

Analysis. A 1 mL solution containing TCEP and an internal UV standard (3,5-dimethoxybenzoic acid) was prepared as follows: 960 µL of a 10 mM TCEP solution in water and 40 µL UV standard solution (5 mM in a 0.2 M pH 7 phosphate buffer) . For each time point an aliquot of the reaction (25 µL) was diluted with 175 µL of this solution, then filtered and injected in the UPLC-ESI (2 x 2 µL injections).

b. Chapter 4

Phosphate buffers were prepared at pH 7.4: the amount of NaH_2PO_4 required to give a 0.2 M solution in 20 mL of water (551 mg) was dissolved in water (10 mL). The resulting acidic solution was titrated to pH 7.4 using 2 M NaOH and its volume adjusted to 20 mL. Buffers were degassed using freeze/thaw cycles and placed under argon prior to use.

Phosphate buffers (0.2 M) containing NaCl (150 mM) and NaN_3 (0.02 %w/v) were prepared at pH 7.4: NaH_2PO4 (551 mg) and EDTA (117 mg) were dissolved in water (10 mL). The resulting acidic solution was titrated to pH 7.4 using 2 M NaOH. NaCl (175 mg) and NaN_3 (4 mg) were then dissolved in the buffer solution and the final volume was adjusted to 20 mL. The buffer solution was degassed using freeze-thaw cycles and placed under argon prior to use.

Ammonium bicarbonate buffers were prepared at pH 7.4: the amount of NH_4HCO_3 required to give a 0.2 M solution in 20 mL of water (158 mg) was dissolved in water (10 mL). The resulting basic solution was titrated to pH 7.4 using formic acid (10 %v/v) and its volume adjusted to 20 mL. Buffers were degassed using freeze/thaw cycles and placed under argon prior to use.

Ammonium bicarbonate buffers (0.2 M) containing NaCl (150 mM) and NaN_3 (0.02 %w/v) were prepared at pH 7.4: NH_4HCO_3 (158 mg) and EDTA (117 mg) were dissolved in water (10 mL). The resulting basic solution was titrated to pH 7.4 using formic acid (10 %v/v). NaCl (175 mg) and NaN_3 (4 mg) were then dissolved in the buffer solution and the final volume was adjusted to 20 mL. The buffer solution was degassed using freeze-thaw cycles and placed under argon prior to use.

Exchange reaction among the affibody helices H1-HER2, H2-HER2 and H3, H1-IgG, H2-IgG and H3 or between H1-HER2 and H2a-HER2.

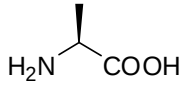
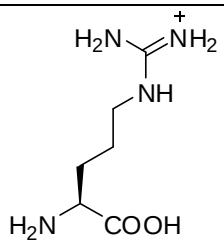
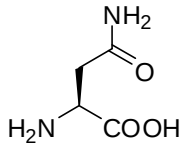
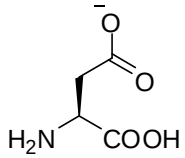
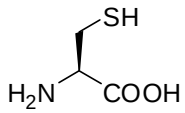
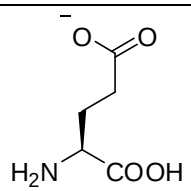
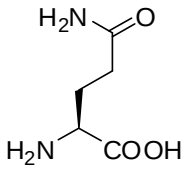
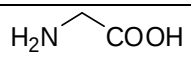
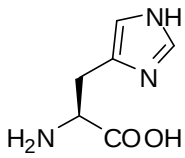
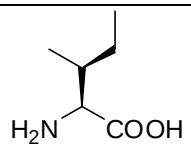
Setup. Each peptide was dissolved in water to give three (or two) 500 μM solutions. Equivalent volumes of these three (or two) solutions were mixed together in a vial and immediately frozen and lyophilized. The lyophilized powder of the mixed peptides was then diluted under argon (final peptide concentration from 53 μM to 1 mM) with the appropriate buffer at pH 7.4 containing DTT (from 1.3 mM to 625 mM). The exchange reactions were then left stirring in a vial and placed into a Schlenk tube, that was equipped with a Young PTFE valve and contained a 250 mM DTT solution in degassed buffer at the bottom.

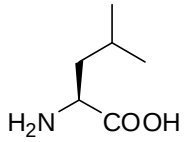
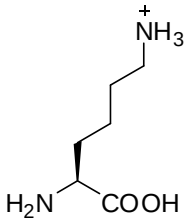
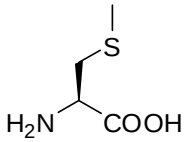
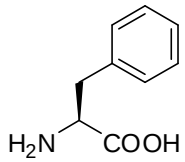
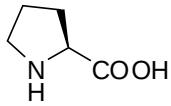
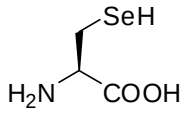
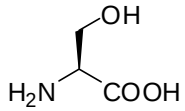
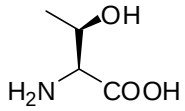
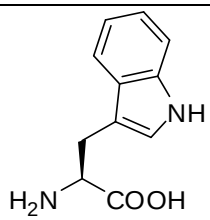
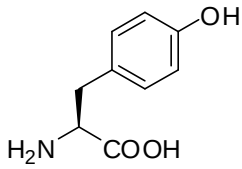
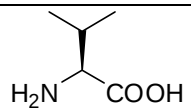
Analysis. Aliquots of the exchange reaction were treated according to different protocols and depending on the chosen analysis method:

- UPLC-ESI: 20 μL aliquots were injected as such or after dilution with a 0.5 %v/v formic acid solution in water (30 μL, injections of 20 μL).
- MALDI-TOF: 30 μL aliquots of the exchange reaction mixture were diluted with 20 μL of a formic acid solution (0.5 %v/v in water). 1 μL of this mixture was then diluted 1000 fold with a 0.1 %v/v TFA solution. 1 μL of this final solution was then co-deposed on a metal plate with a solution of the matrix, CHCA or DHB, and analyzed.
-UPLC-ESI-TOF: 1 μL aliquots of the reaction mixture were diluted with 100 μL of a 0.1 %v/v TFA solution in water; 2 μL of the resulting solution were then injected.

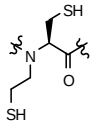
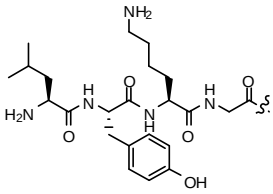
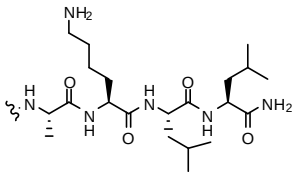
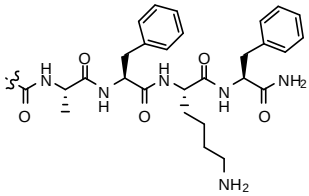
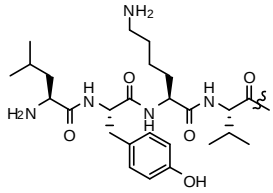
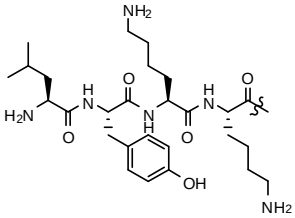
ANNEX

Annex 1 – Structures and abbreviations of amino acids

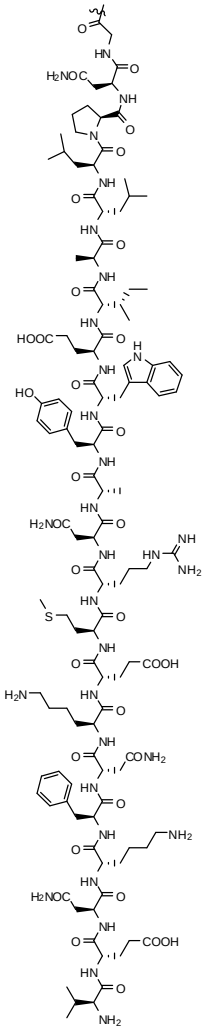
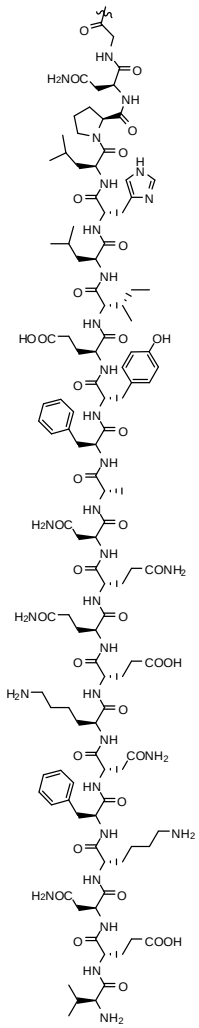
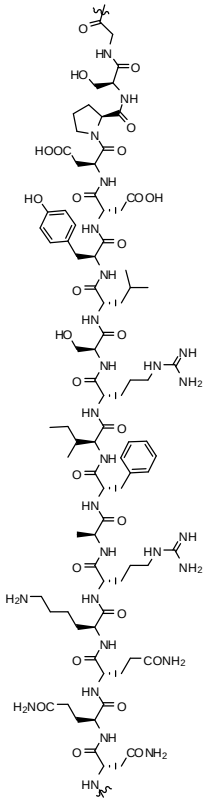
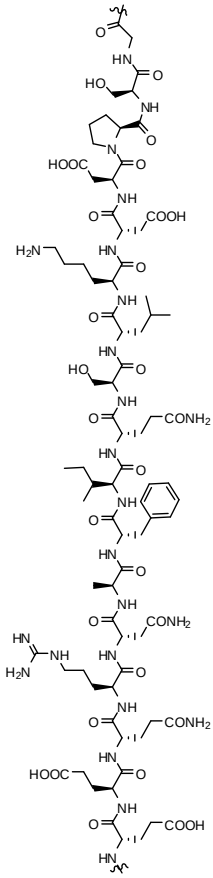
Full name	Abbreviations	Structure
Alanine	Ala, A	
Arginine	Arg, R	
Asparagine	Asn, N	
Aspartic acid	Asp, D	
Cysteine	Cys, C	
Glutamic acid	Glu, E	
Glutamine	Gln, Q	
Glycine	Gly, G	
Histidine	Hys, H	
Isoleucine	Ile, I	

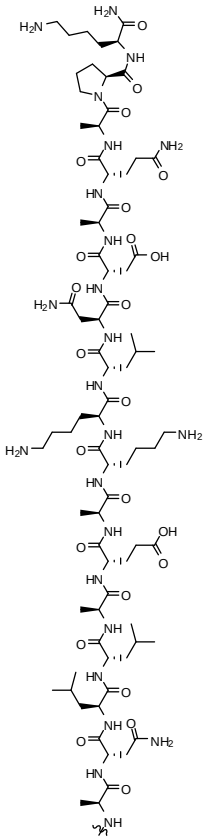
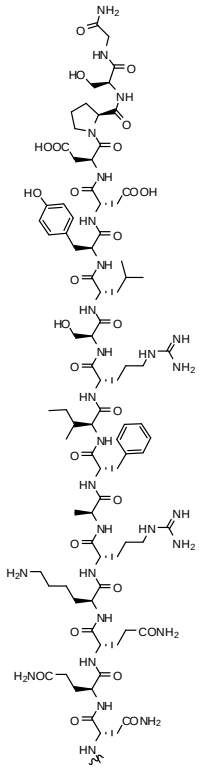
Leucine	Leu, L	
Lysine	Lys, K	
Methionine	Met, M	
Phenylalanine	Phe, F	
Proline	Pro, P	
Selenocysteine	Sec, U	
Serine	Ser, S	
Threonine	Thr, T	
Tryptophan	Trp, W	
Tyrosine	Tyr, Y	
Valine	Val, V	

Annex 2 – Structures and abbreviations of specific amino acids and peptides used in Chapters 2 and 3

Abbreviation	Structure
-Daa1-	N-(2-thioetyl)-cysteine 
-Daa2-	N-methyl-cysteine
P1-	LYKG- 
-P2	-AKLL 
-P3	-AFKF 
P1'-	LYKV- 
P1''-	LYKK- 

Annex 3 – Structures and abbreviations of the affibody helices used in Chapter 4

Abbreviation		Structure
H1-	HER2	VENKFNKEMRNAYWEIALLPNG- 
	IgG	VENKFNKEQQNAFYEILHLPNG- 
-H2-	HER2	-NQQKRAFIRSLYDDPSG- 
	IgG	-EEQRNAFIQSLKDDPSG- 

Abbreviations	Structure
-H3 (HER2 and IgG)	-ANLLAEAKKLNDAQAPK 
-H2a (HER2)	-NQQKRAFIRSLYDDPSG 

Annex 4 – Further characterization of the peptides from Chapter 2

- *P1-Daa1-P2ox*

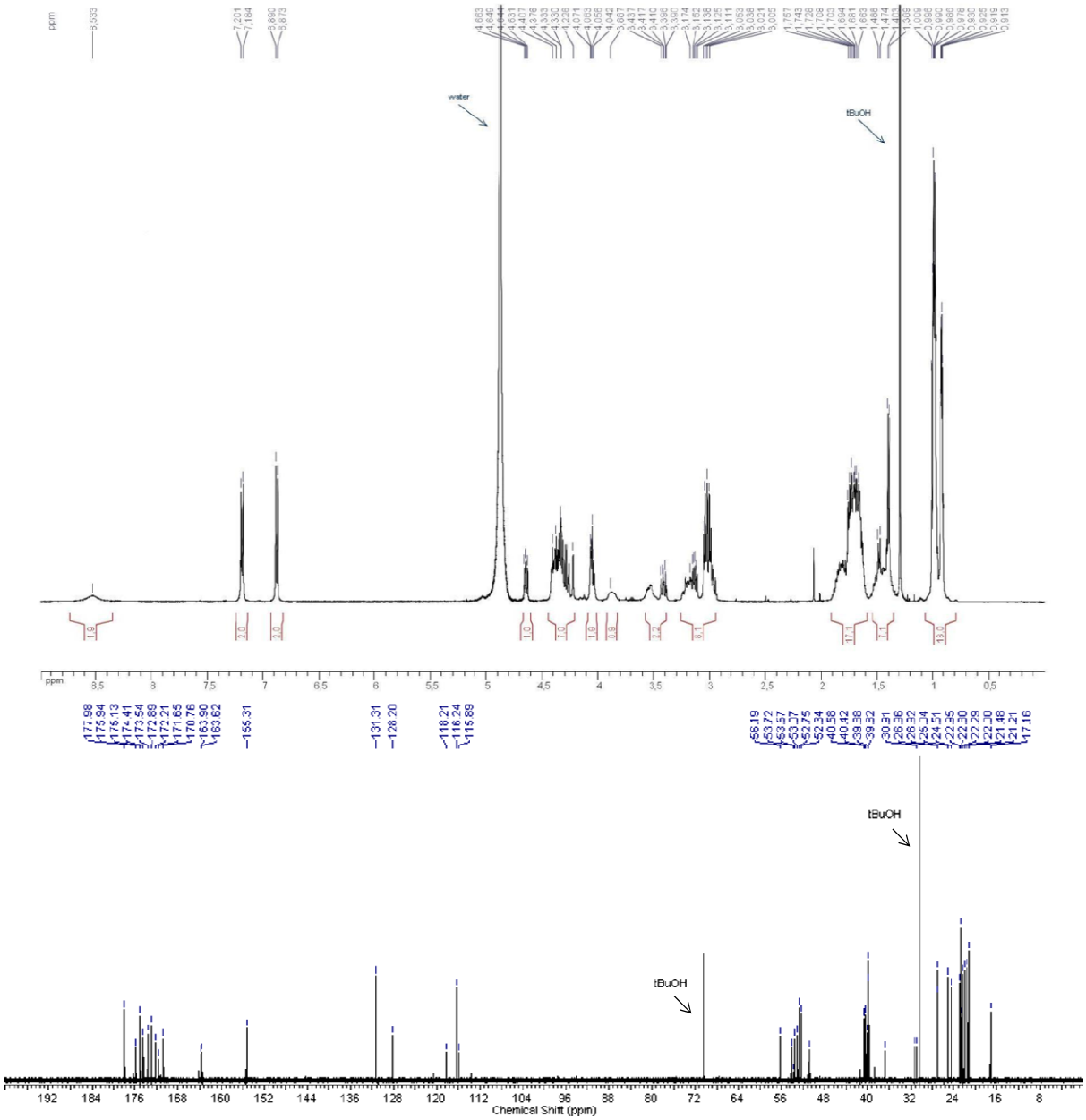


Figure 60 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1-Daa1-P2ox**.

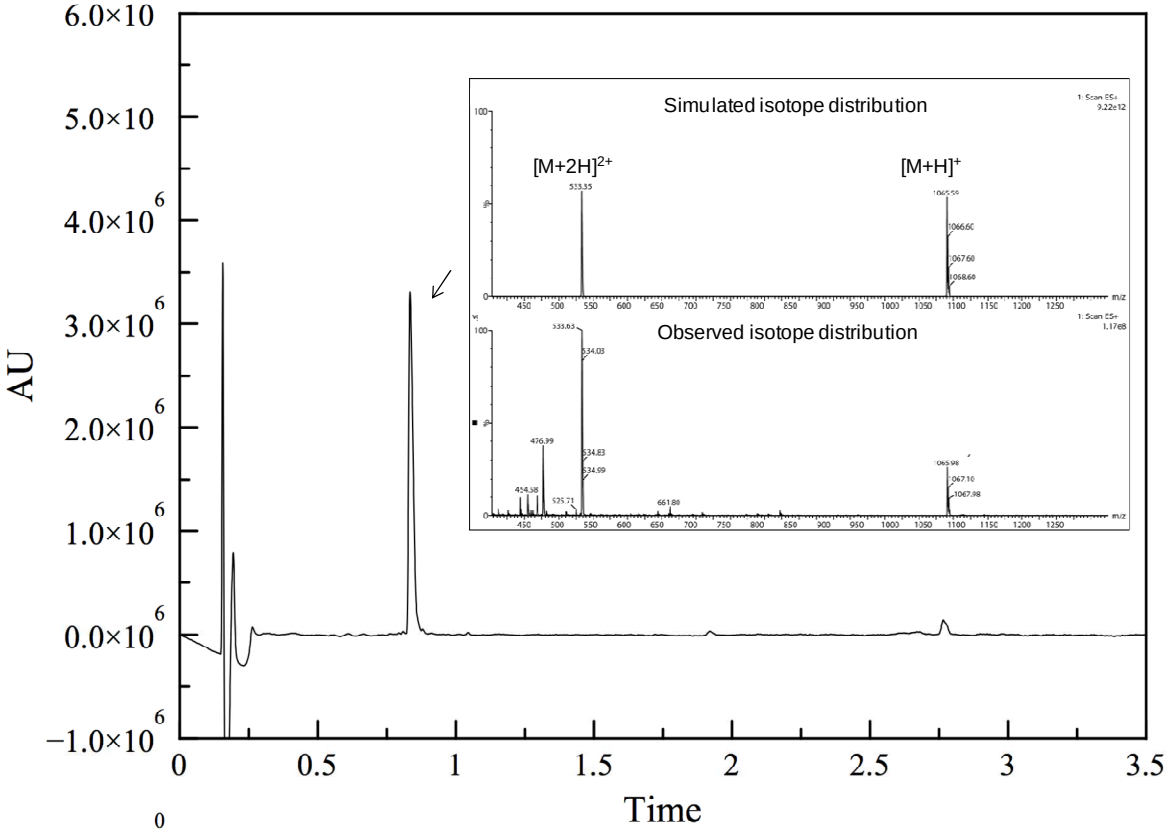


Figure 61 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1-Daa1-P2ox** and relative ESI-MS spectrum.

- *Daa1-P3ox*

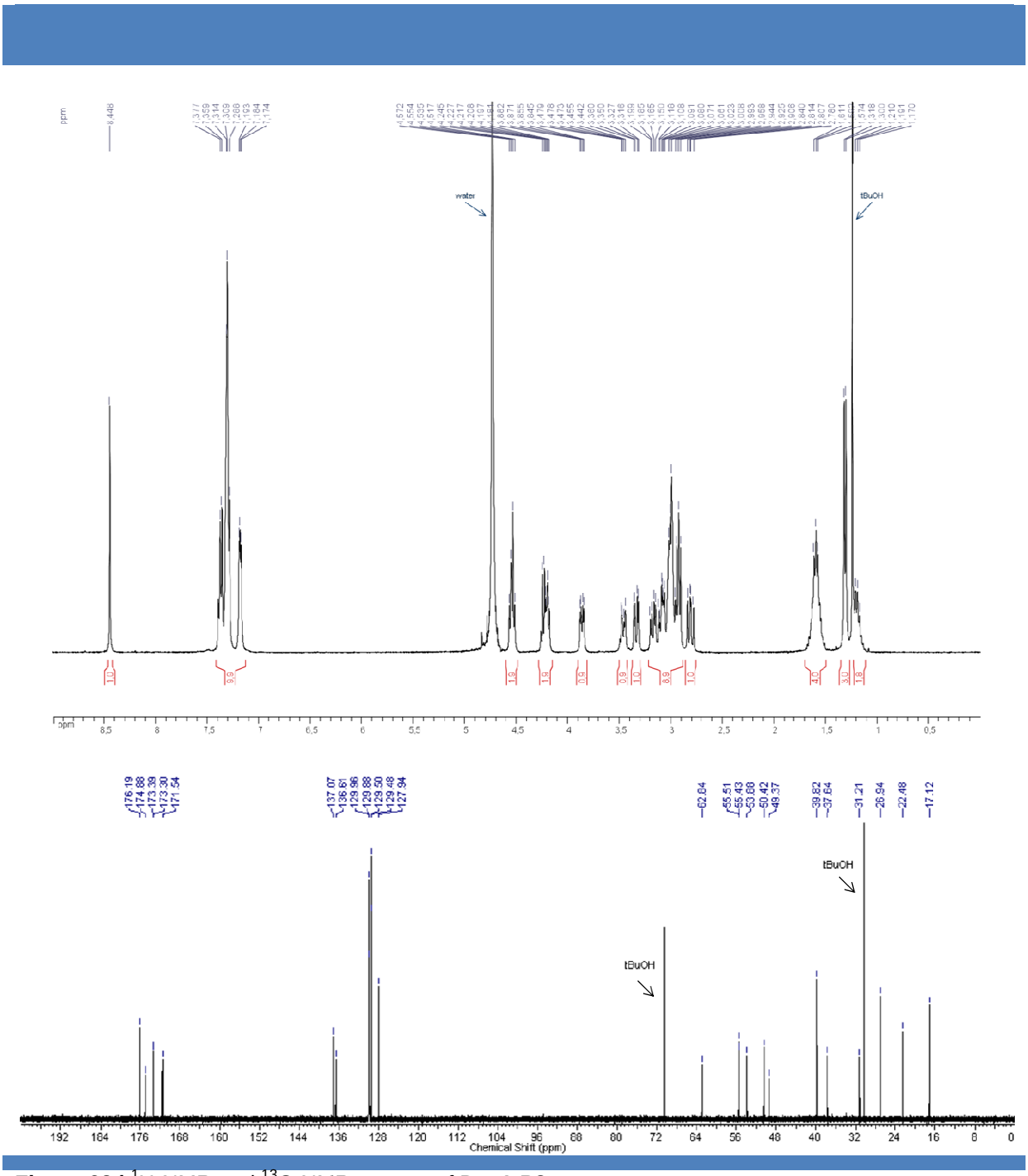


Figure 62 | ^{1}H-NMR and ^{13}C-NMR spectra of **Daa1-P3ox**.

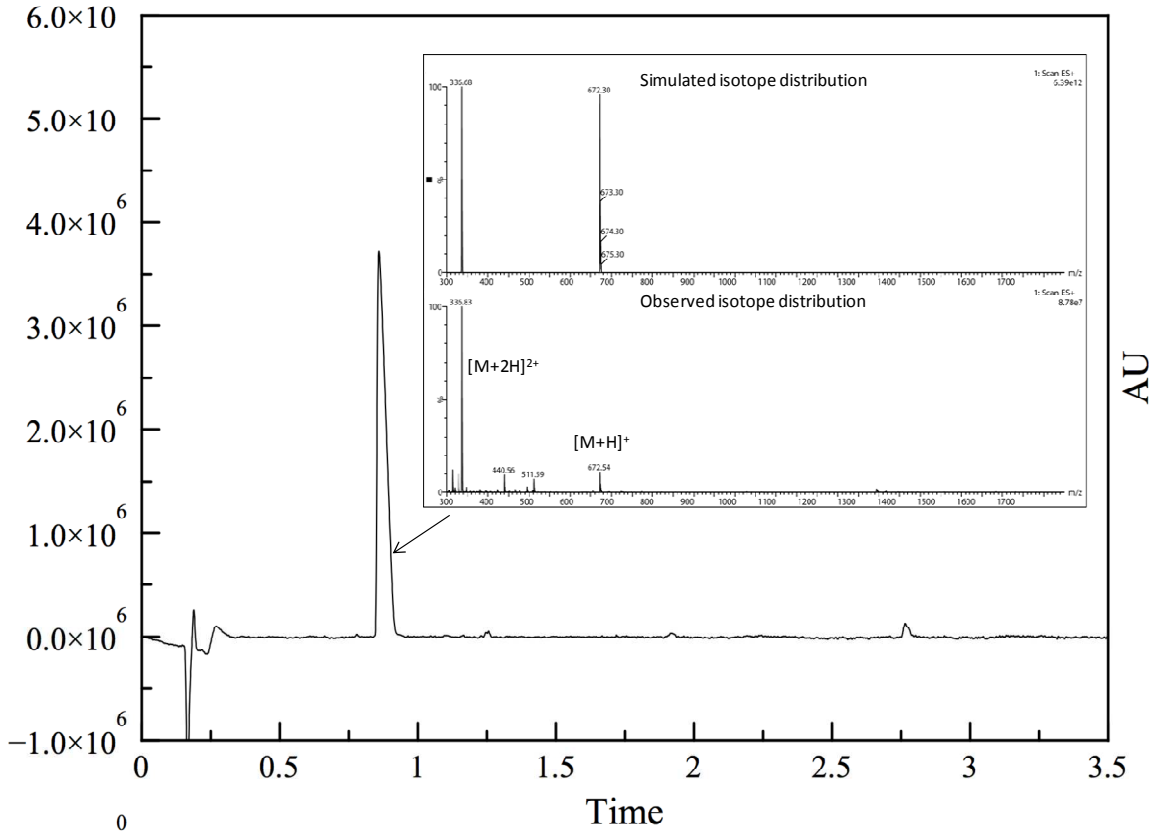


Figure 63 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **Daa1-P3ox** and relative ESI-MS spectrum.

- *P1-Daa1-P3ox*

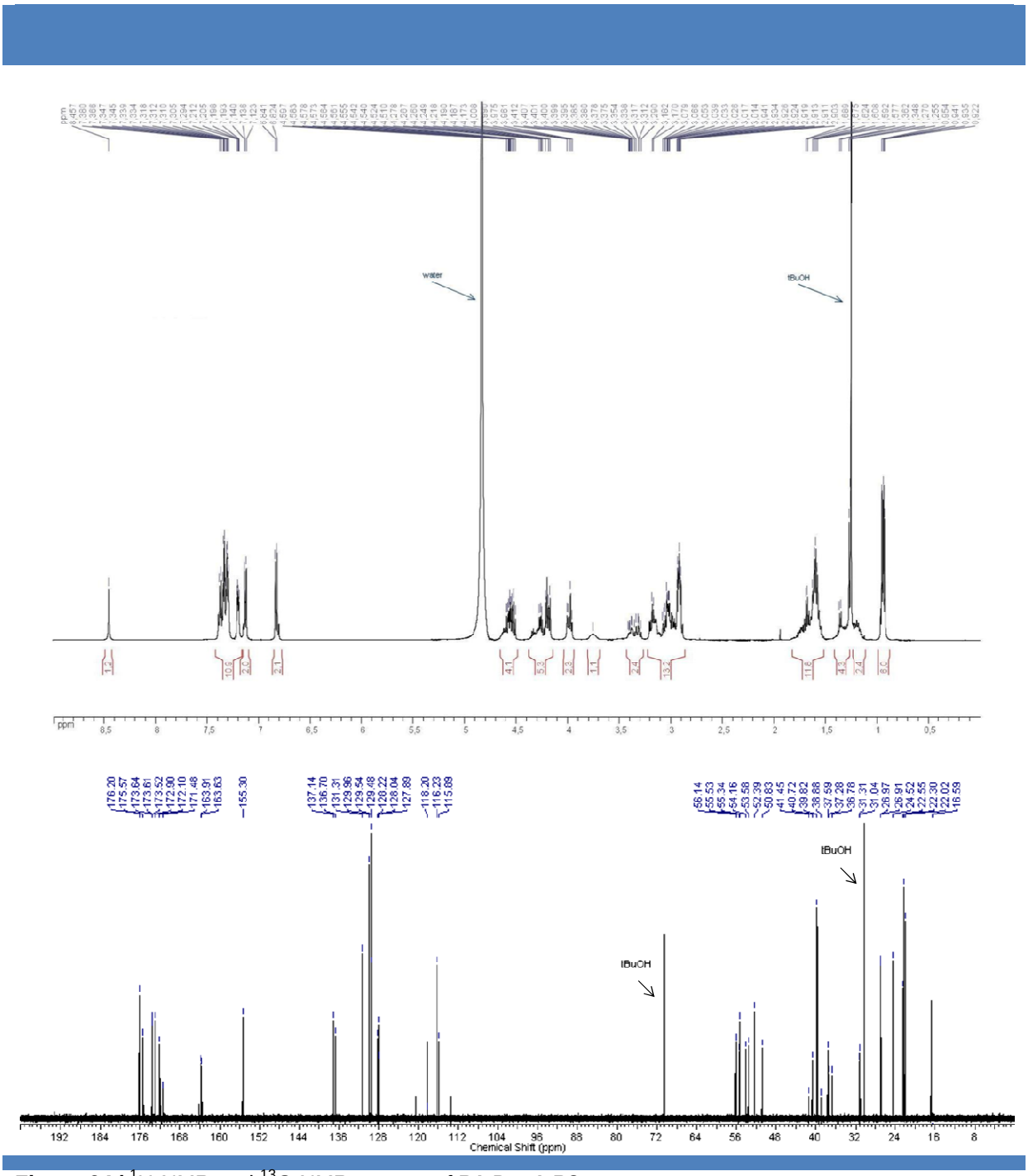


Figure 64 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1-Daa1-P3ox**.

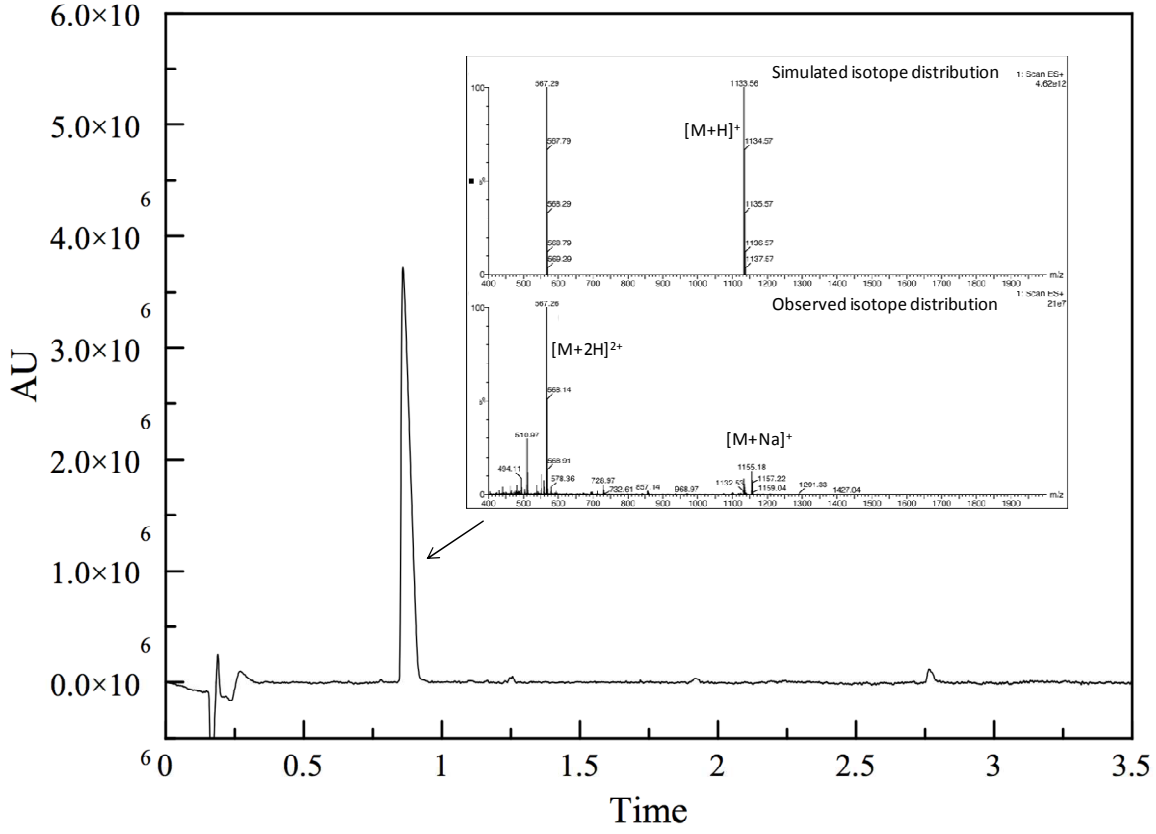


Figure 65 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1-Daa1-P3ox** and relative ESI-MS spectrum.

- *Daa1-P2ox*

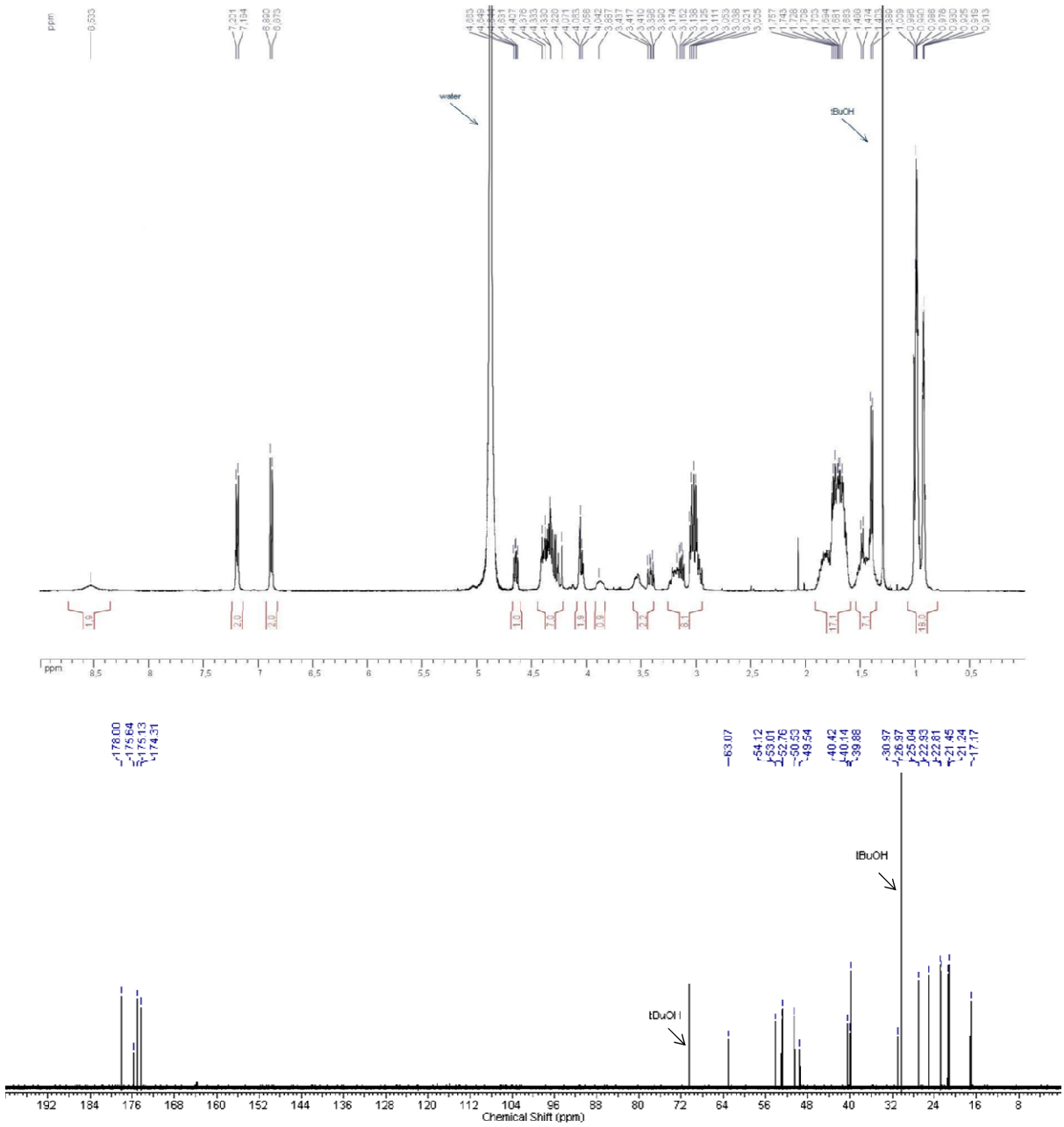


Figure 66 | ^{1}H-NMR and ^{13}C-NMR spectra of **Daa1-P2ox**.

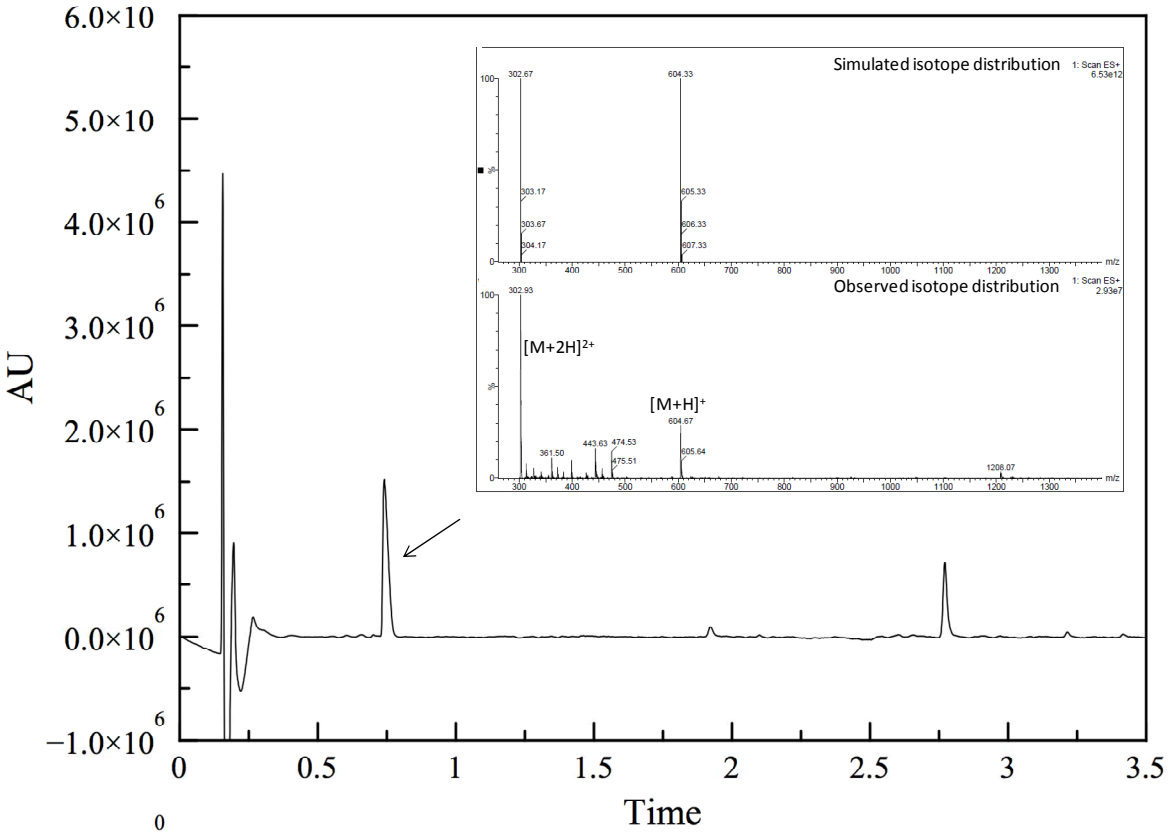


Figure 67 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **Daa1-P2ox** and relative ESI-MS spectrum.

- *P1-Cys-P2*

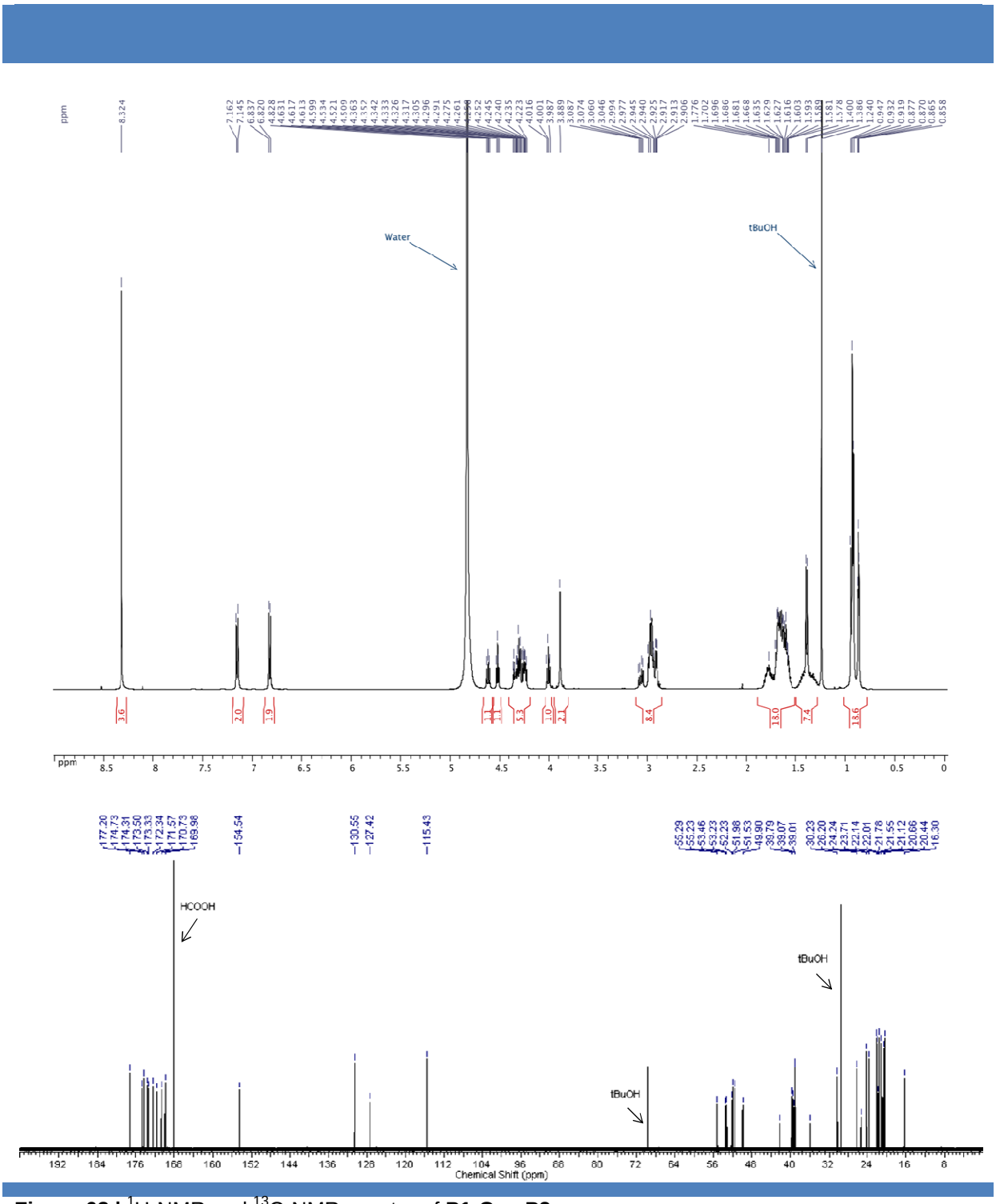


Figure 68 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1-Cys-P2**.

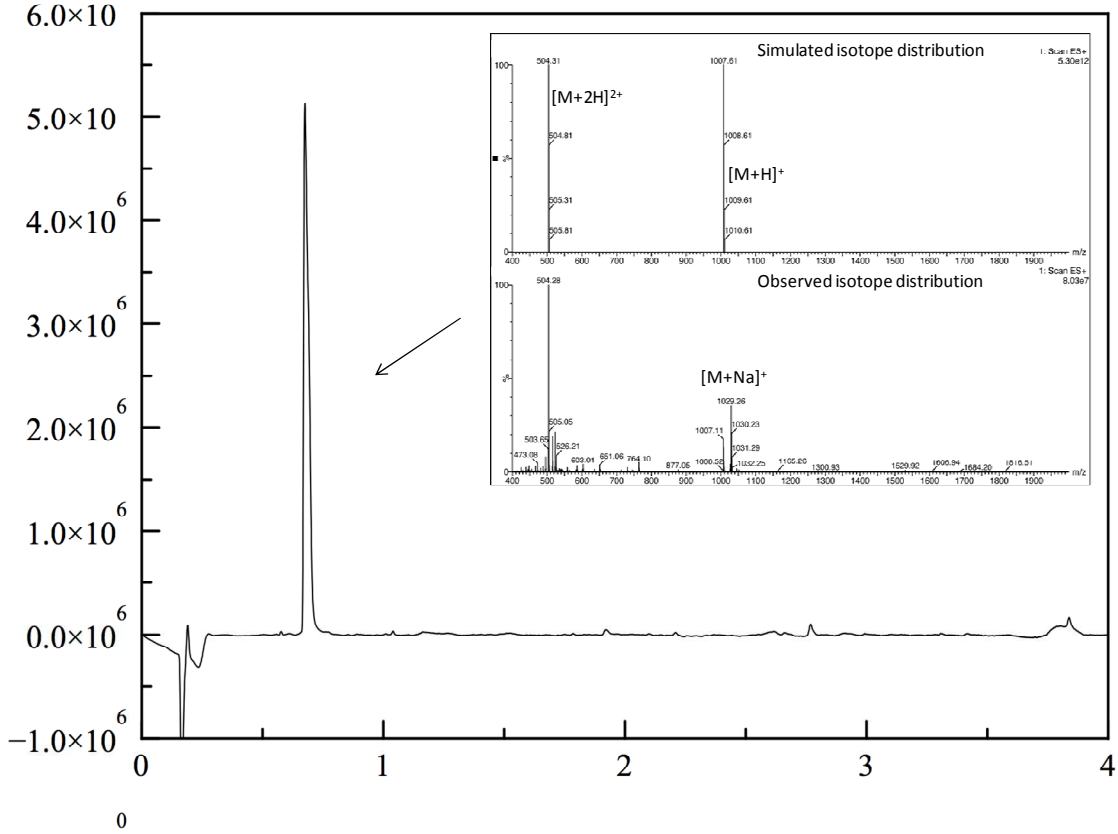


Figure 69 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1-Cys-P2** and relative ESI-MS spectrum.

- *Cys-P3*

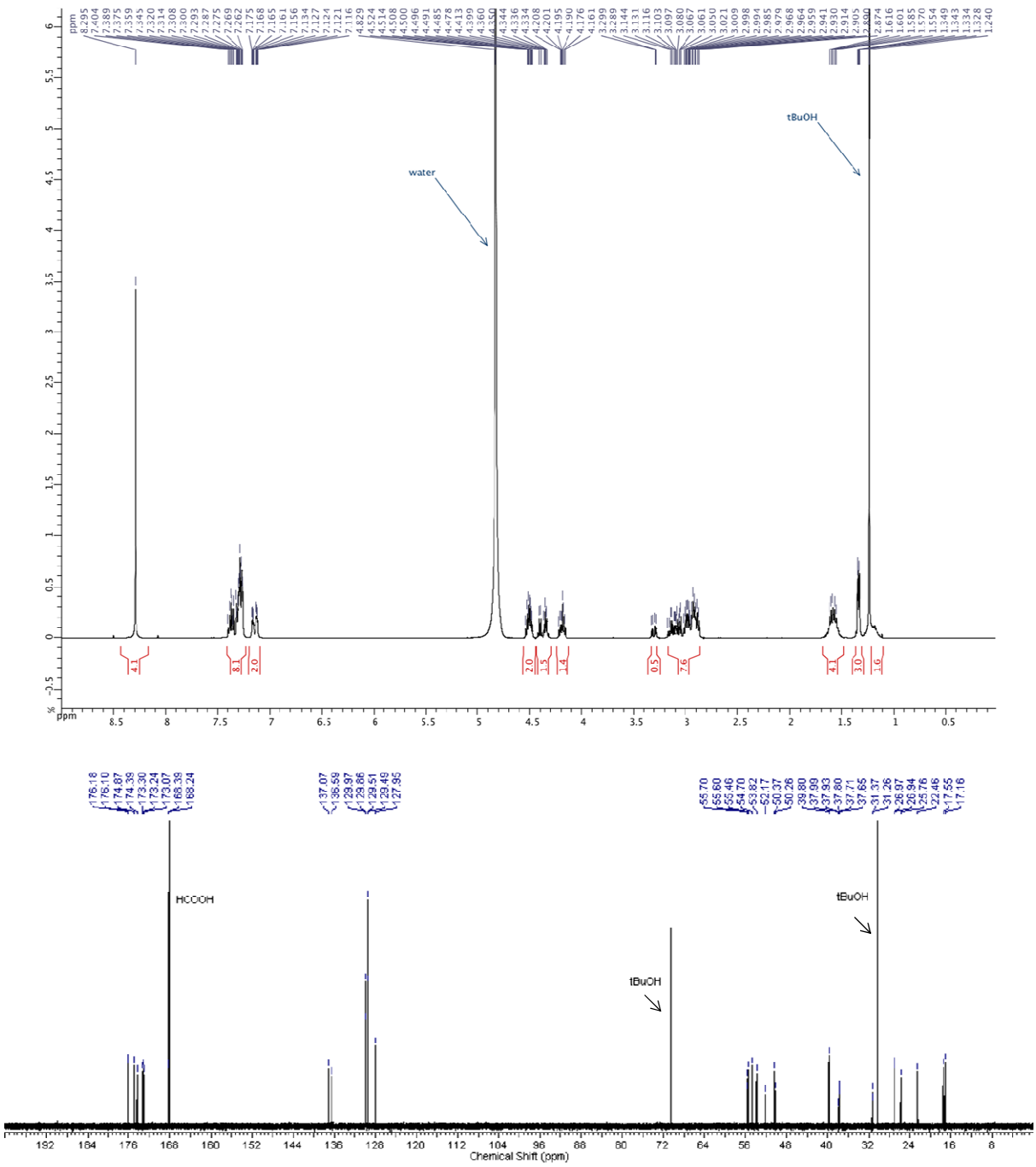


Figure 70 | ^{1}H-NMR and ^{13}C-NMR spectra of **Cys-P3**.

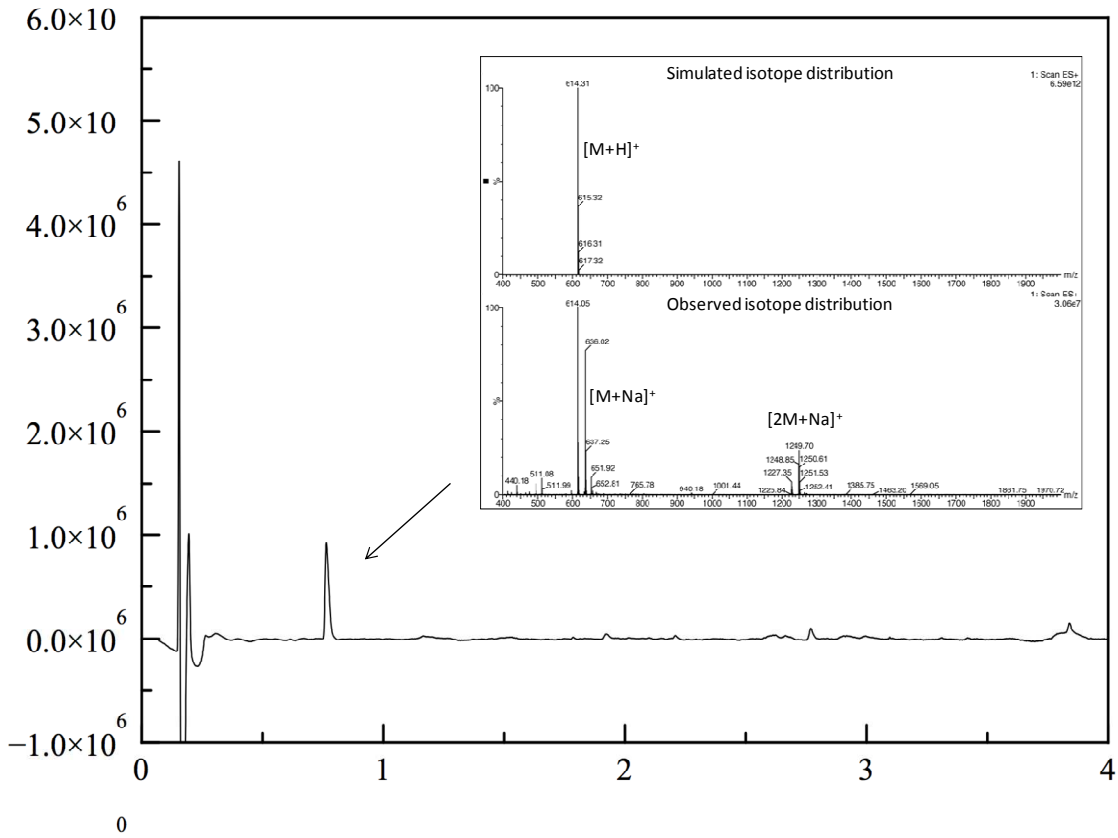


Figure 71 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **Cys-P3**. and relative ESI-MS spectrum.

Annex 5 – Further characterization of the peptides from Chapter 3

- *P1-Daa2-P2*

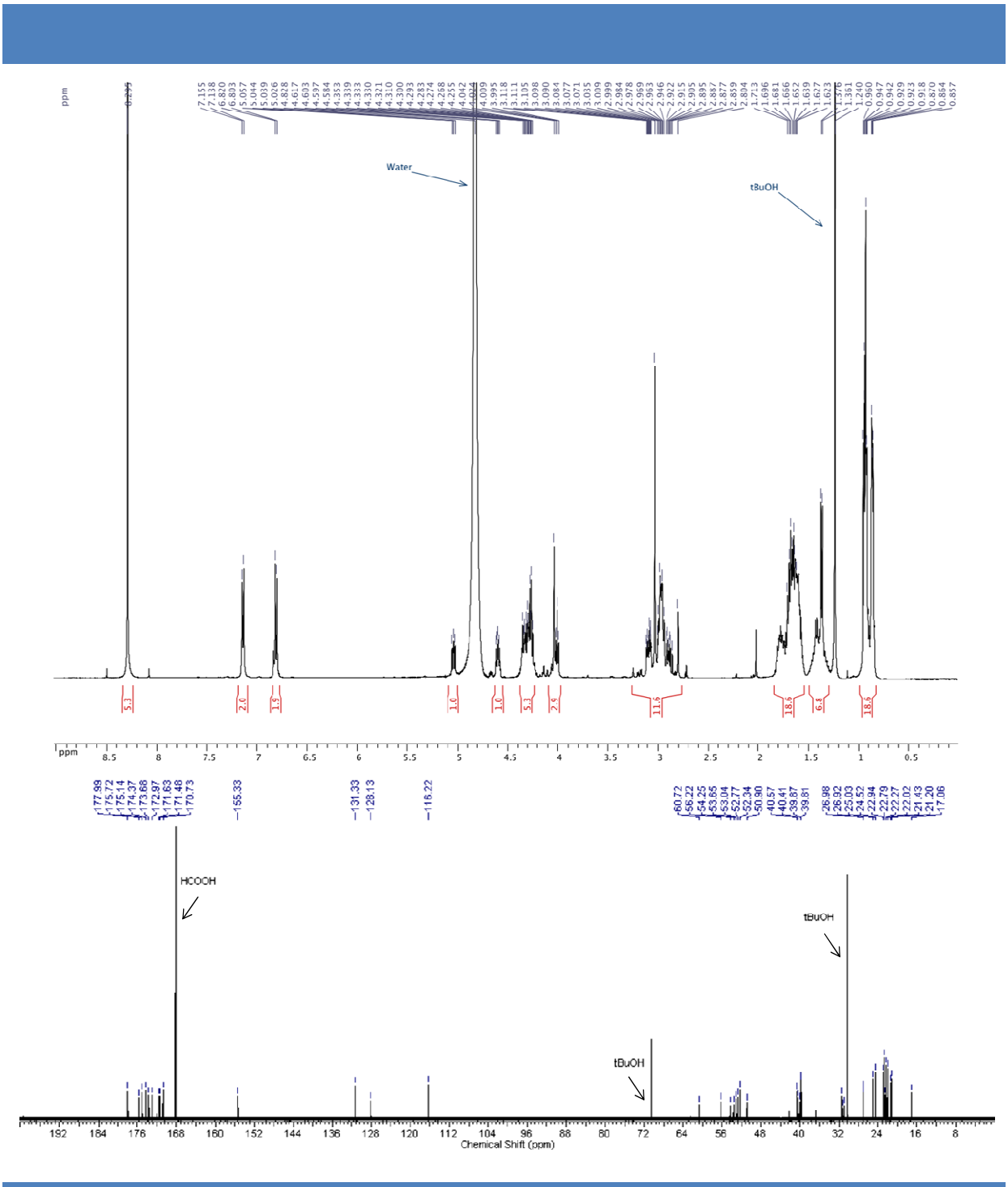
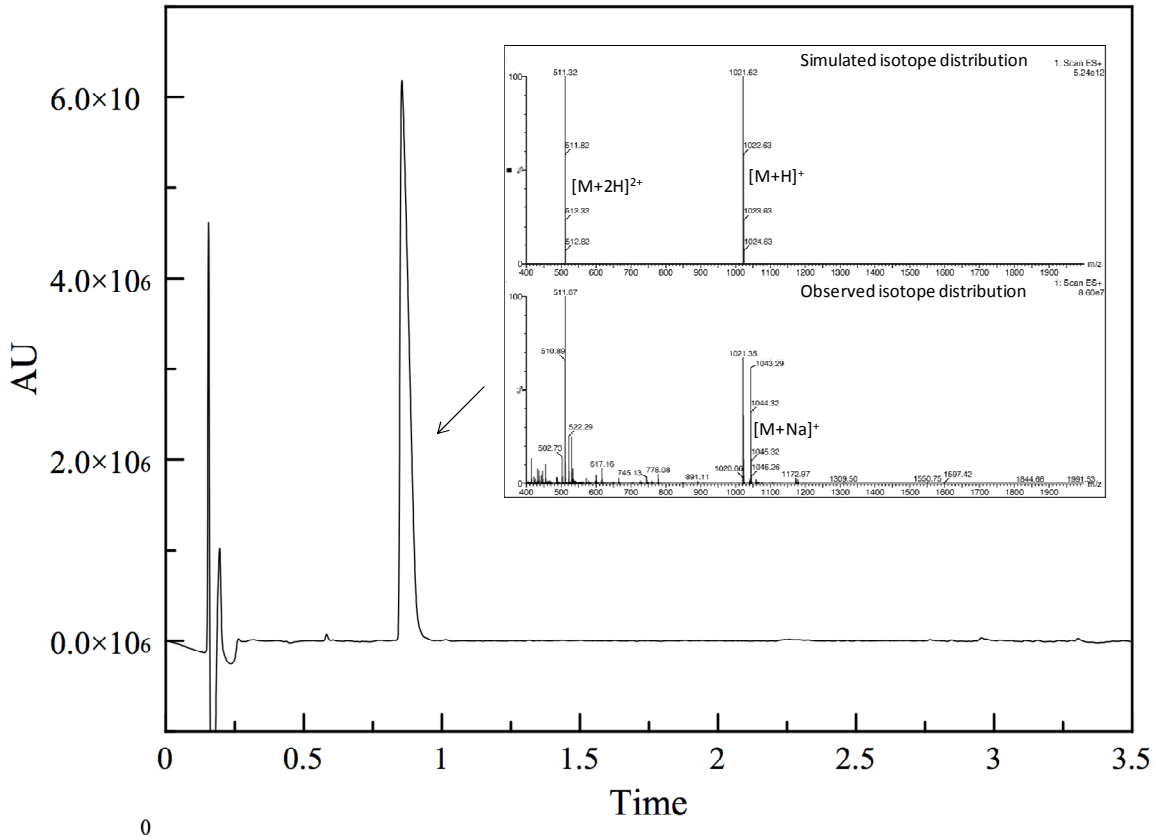


Figure 72 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1-Daa2-P2**.



0

Figure 73 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1-Daa2-P2** and relative ESI-MS spectrum.

- *Daa2-P3*

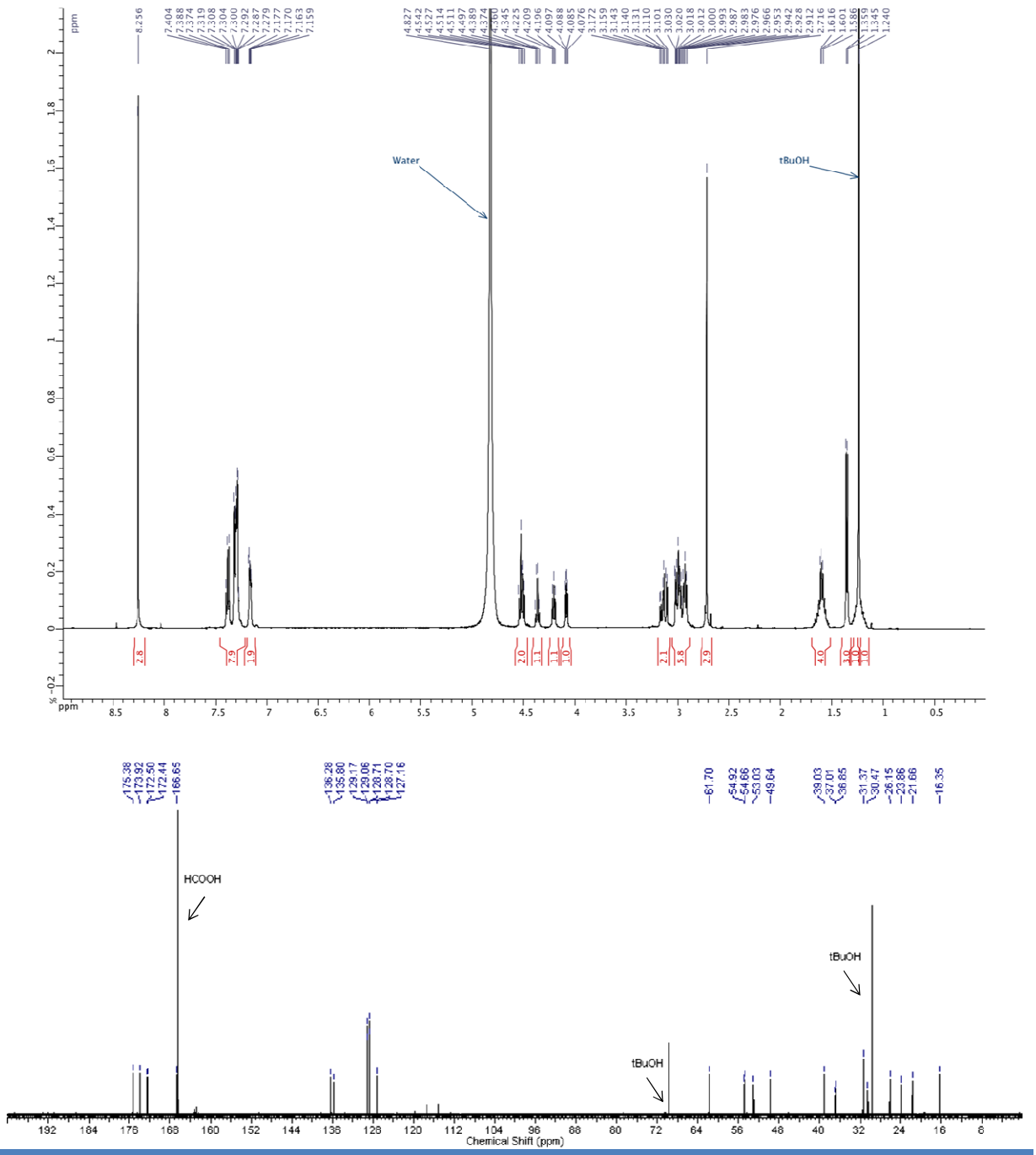


Figure 74 | ^{1}H-NMR and ^{13}C-NMR spectra of **Daa2-P3**.

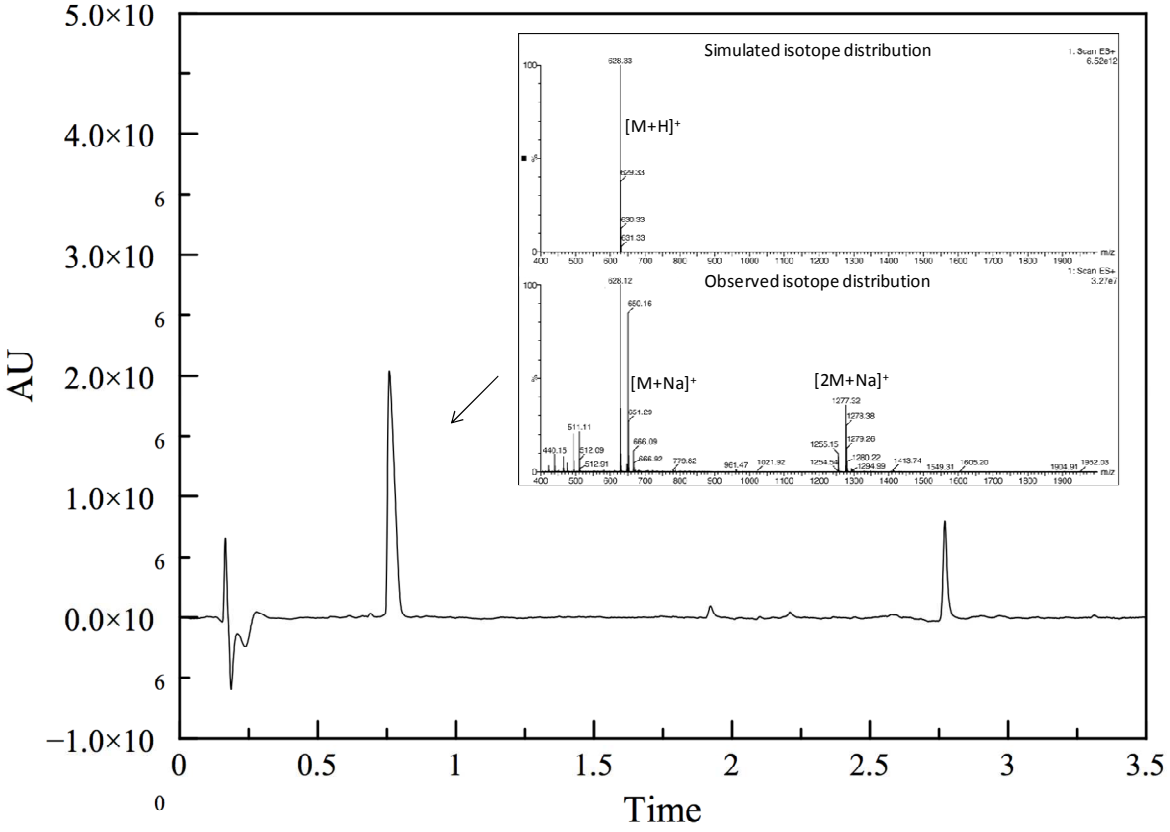


Figure 75 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **Daa2-P3** and relative ESI-MS spectrum.

- *P1-Daa2-P3*

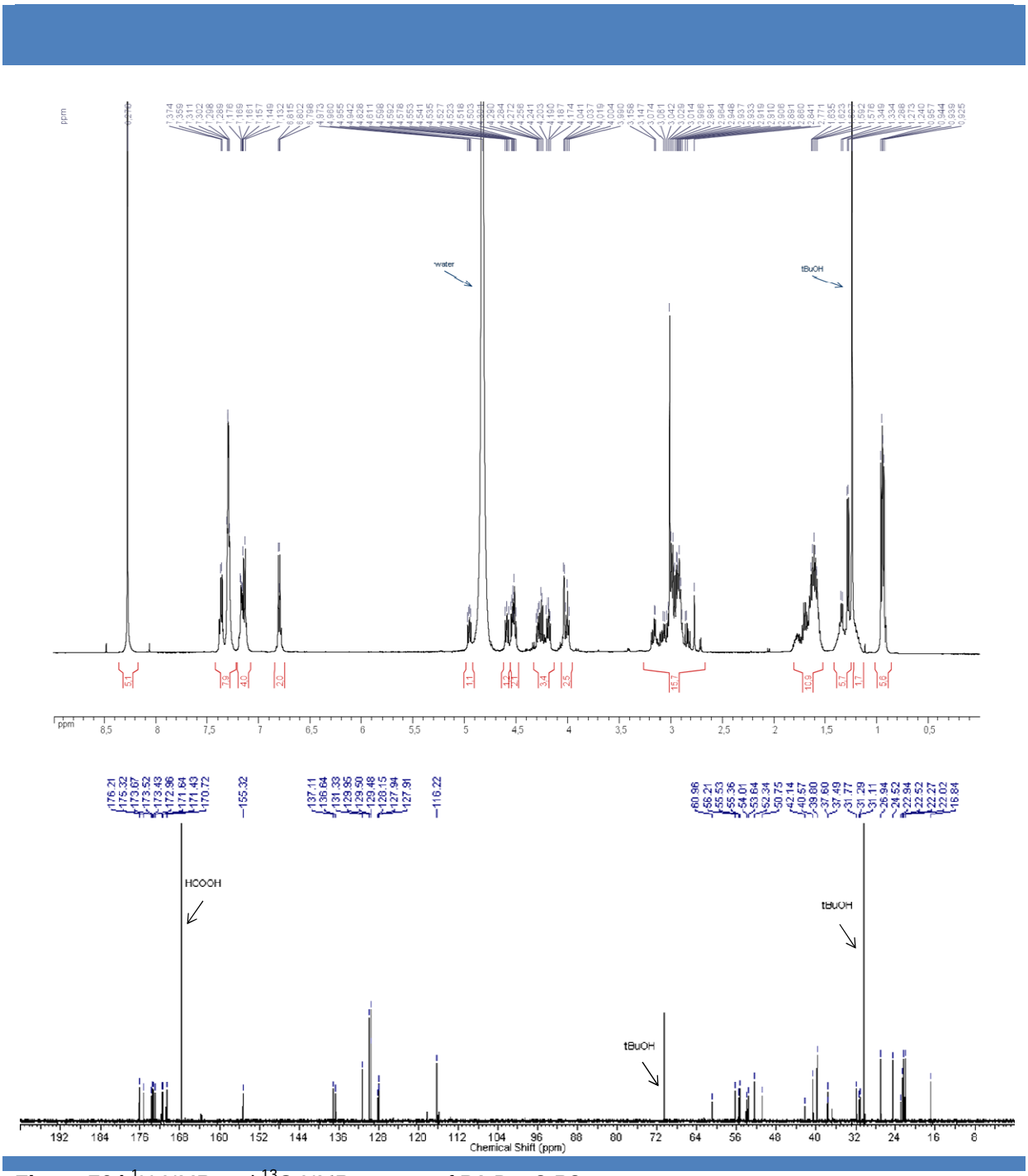


Figure 76 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1-Daa2-P3**.

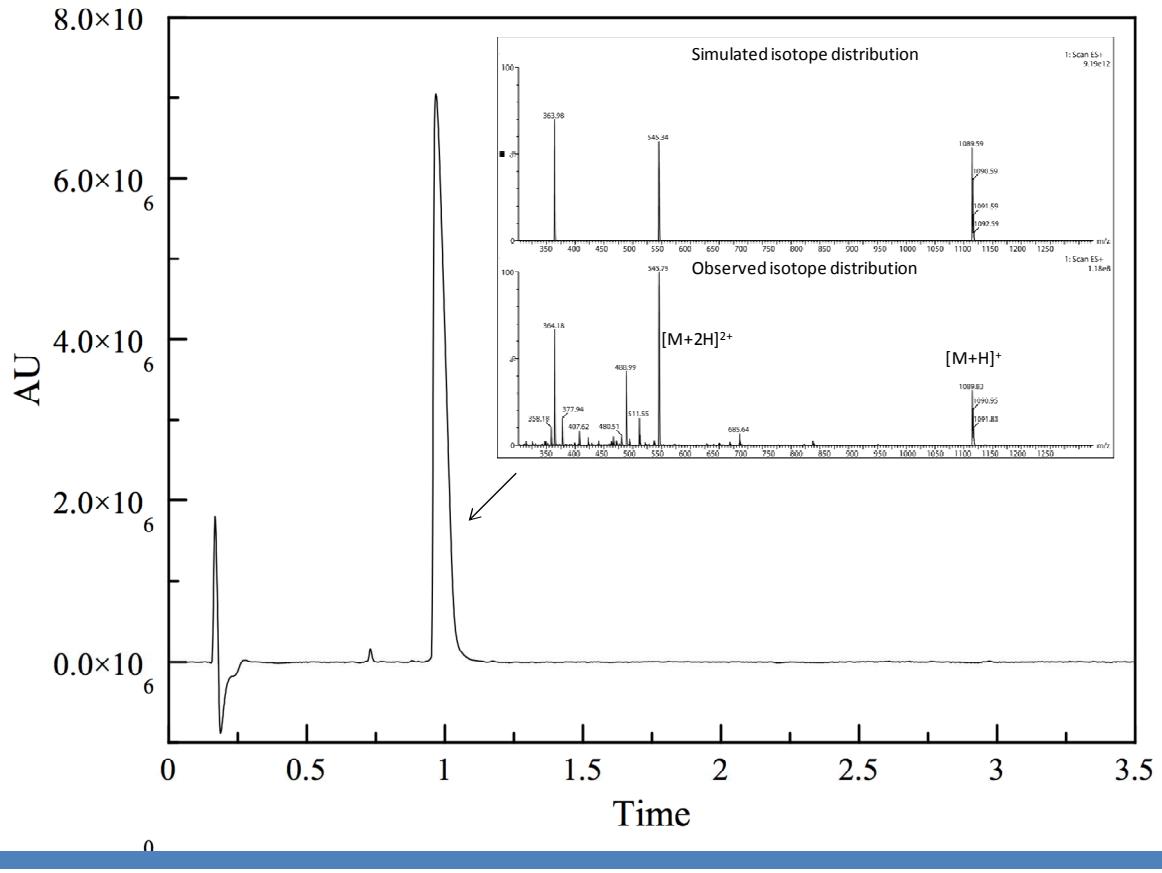


Figure 77 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1-Daa2-P3** and relative ESI-MS spectrum.

- *P1'-Daa2-P2*

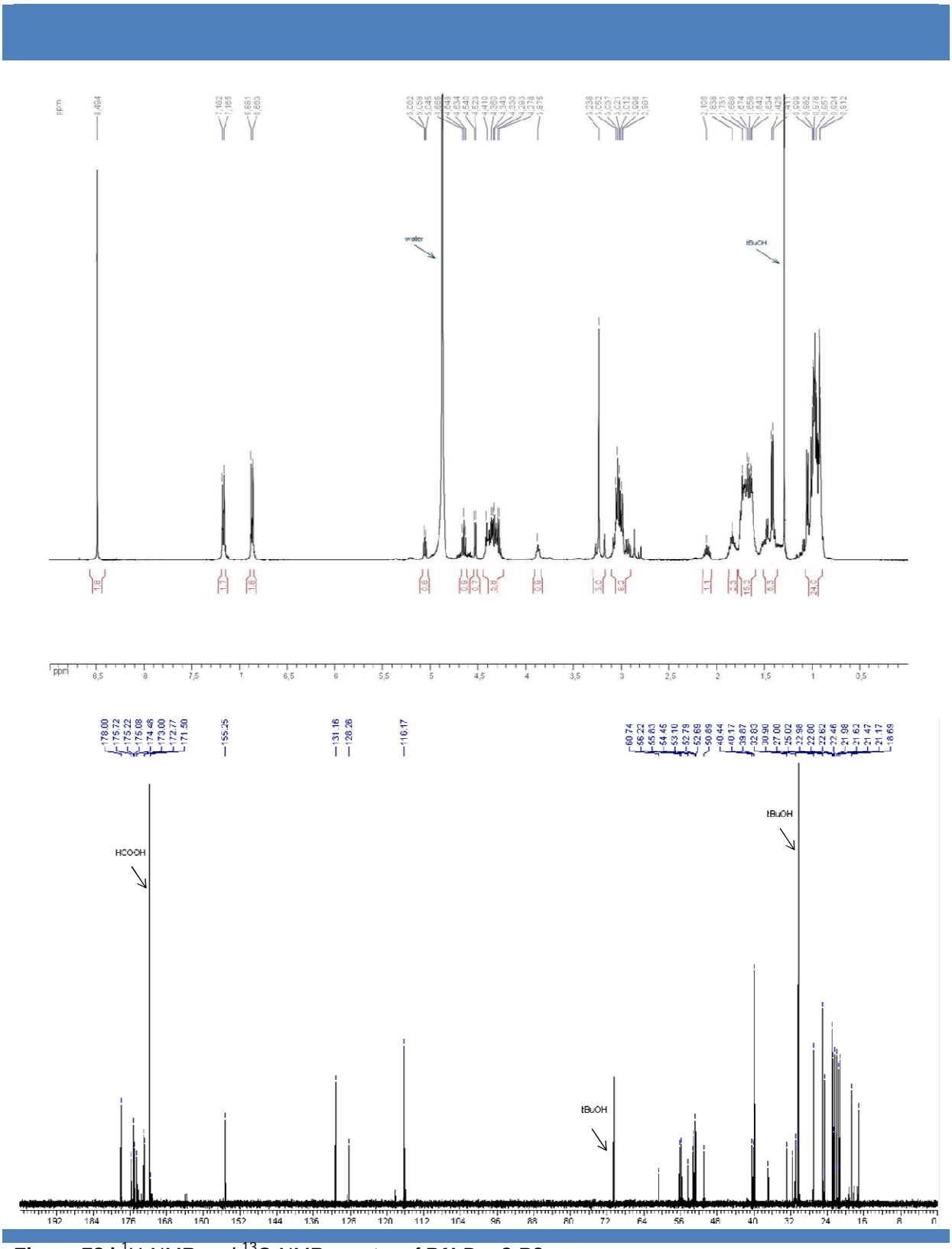


Figure 78 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1'-Daa2-P2**.

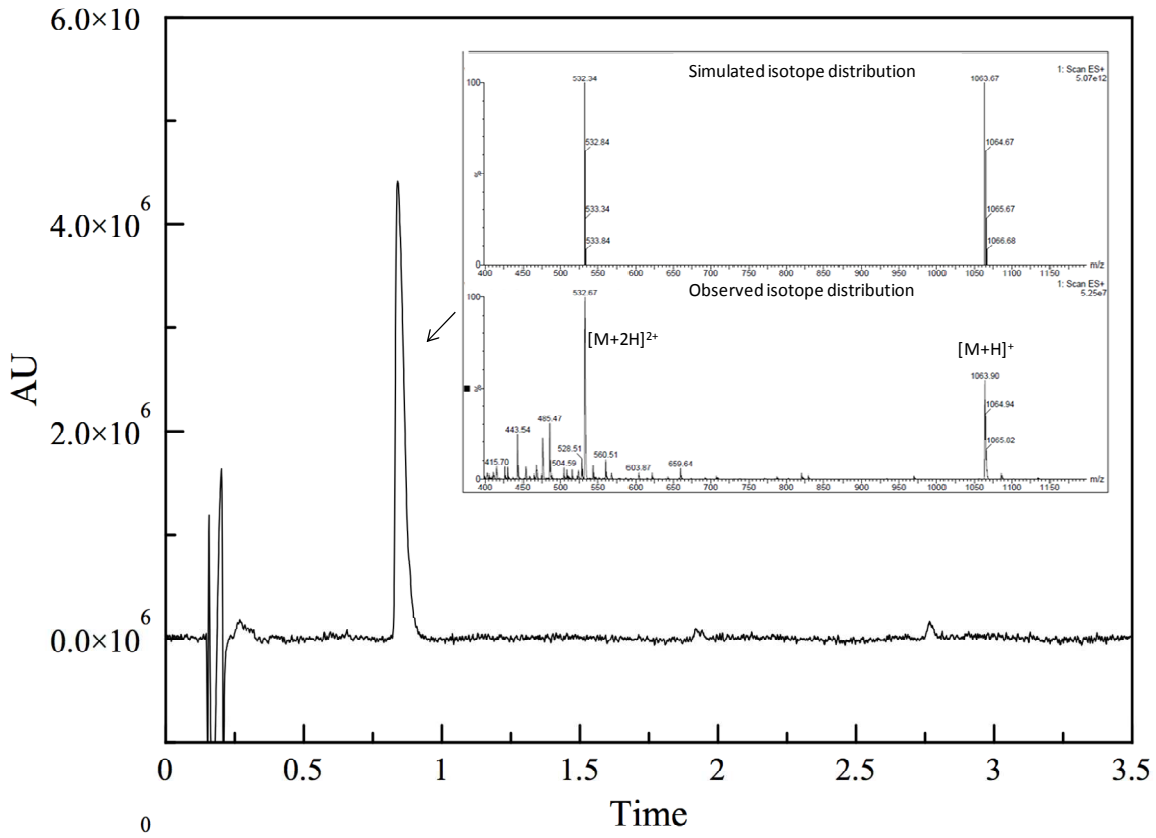


Figure 79 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1'-Daa2-P2** and relative ESI-MS spectrum.

- *P1”-Daa2-P2*

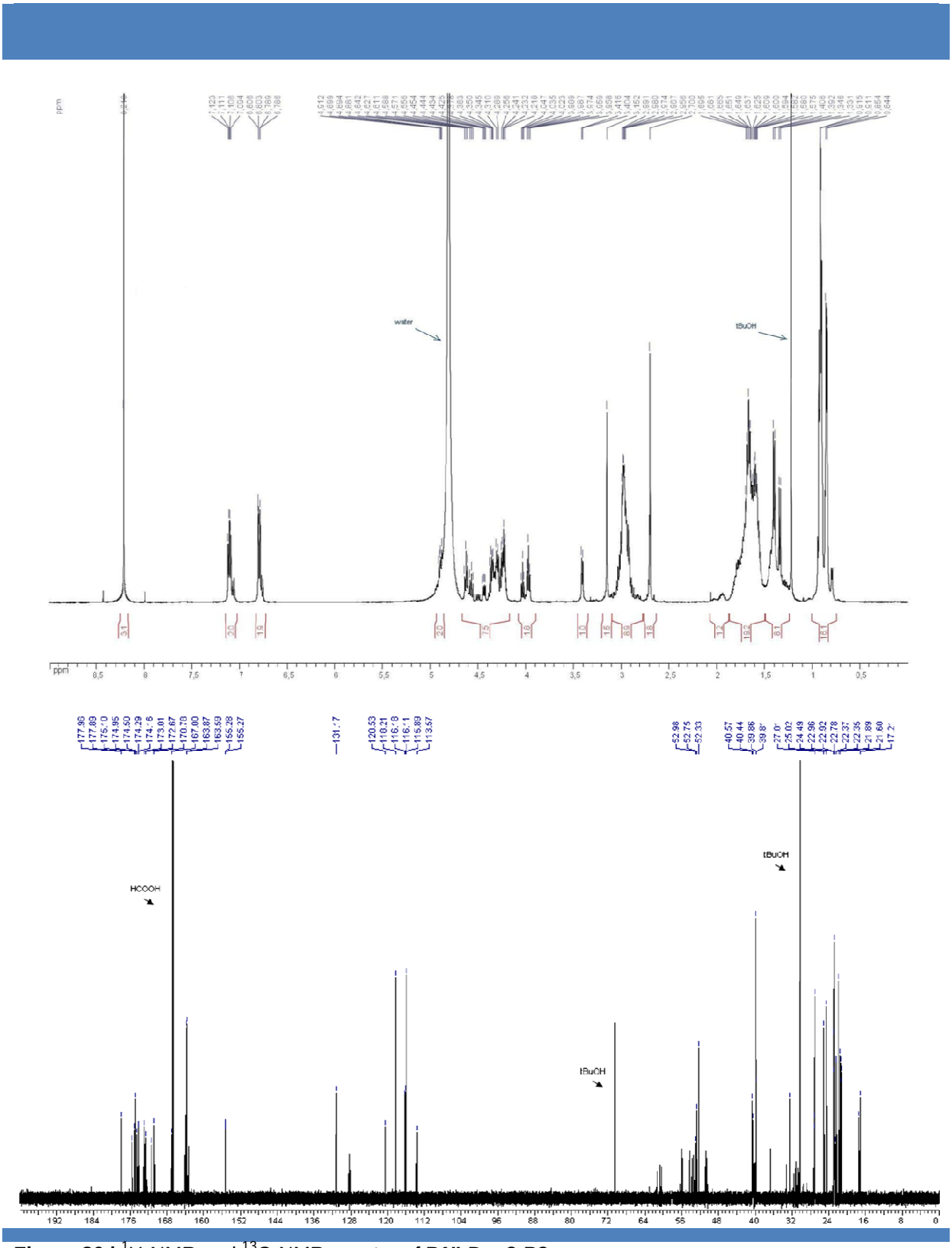


Figure 80 | ^{1}H-NMR and ^{13}C-NMR spectra of **P1”-Daa2-P2**.

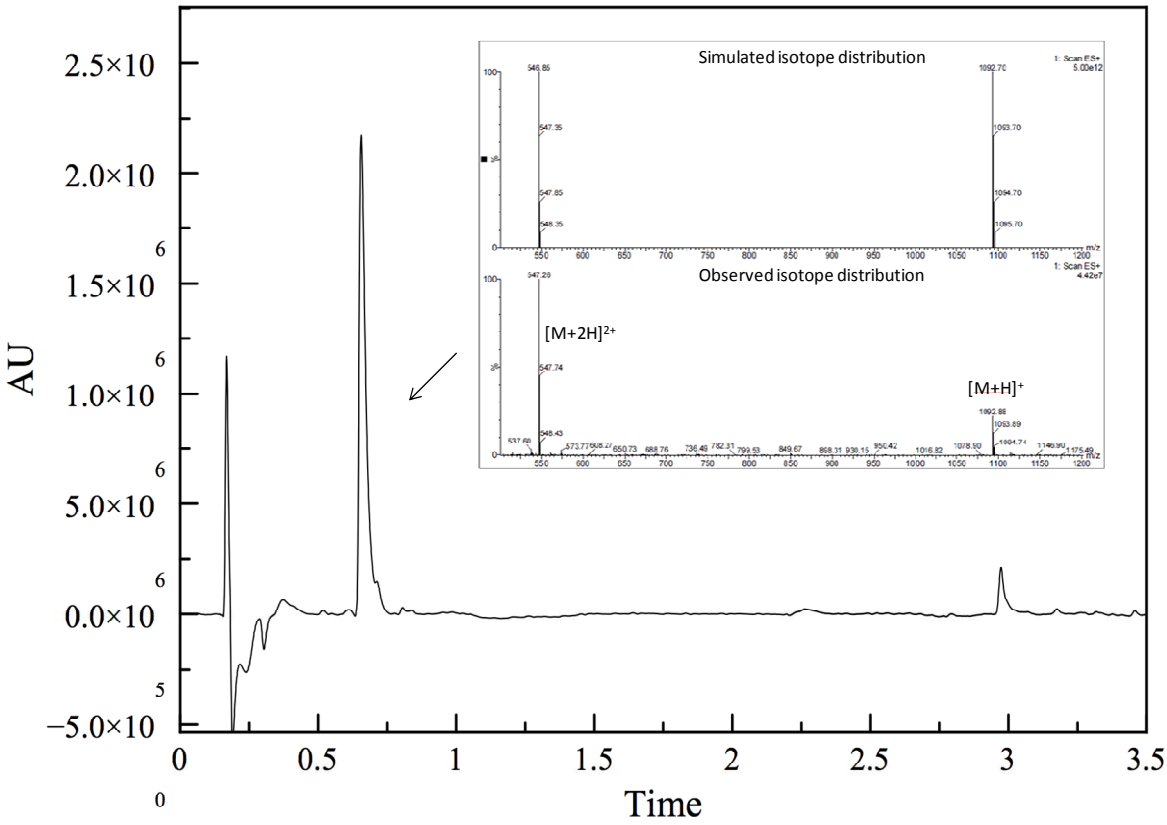


Figure 81 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **P1''-Daa2-P2** and relative ESI-MS spectrum.

Annex 6 – Further characterization of the peptides from Chapter 4

- *H1(StBu)-HER2*

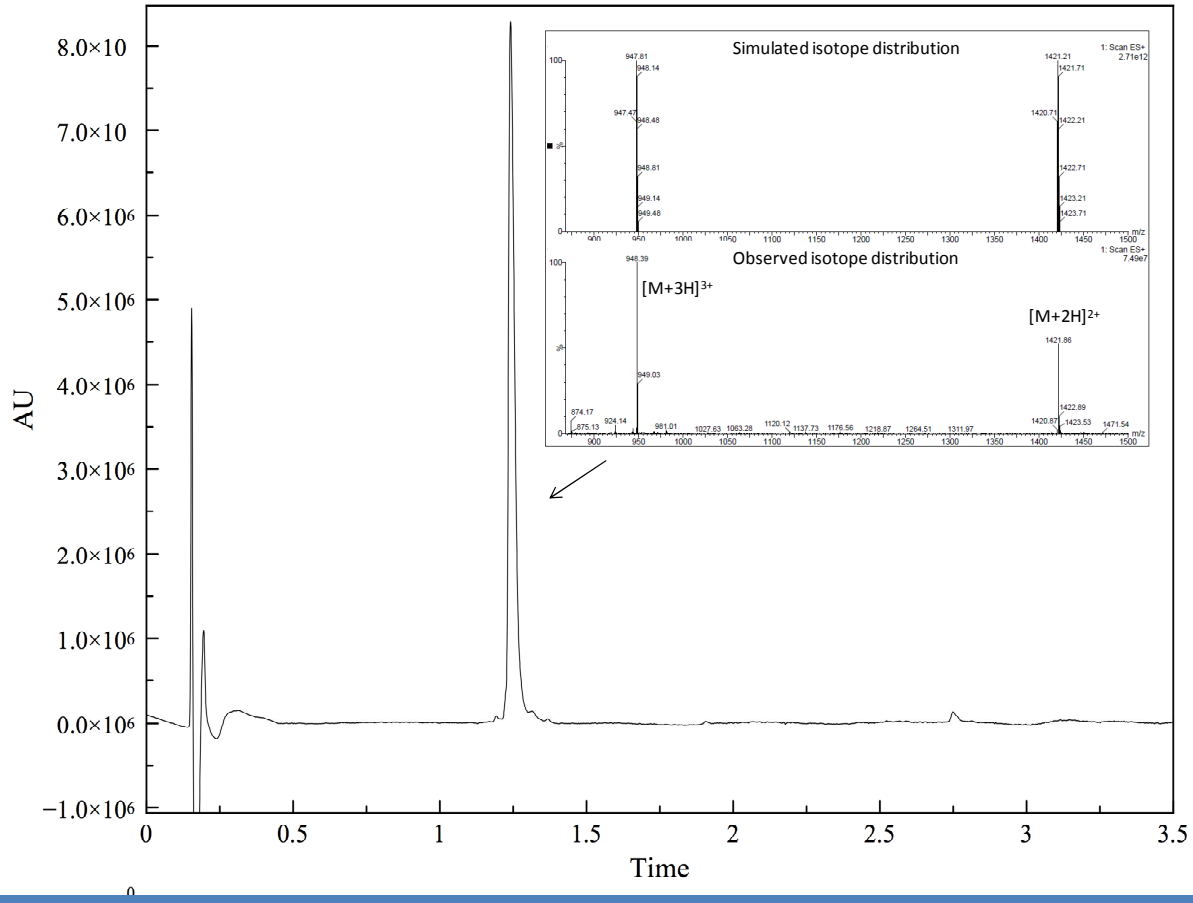


Figure 82 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H1(StBu)-HER2** and relative ESI-MS spectrum.

- *H2(StBu)$_2$-HER2*

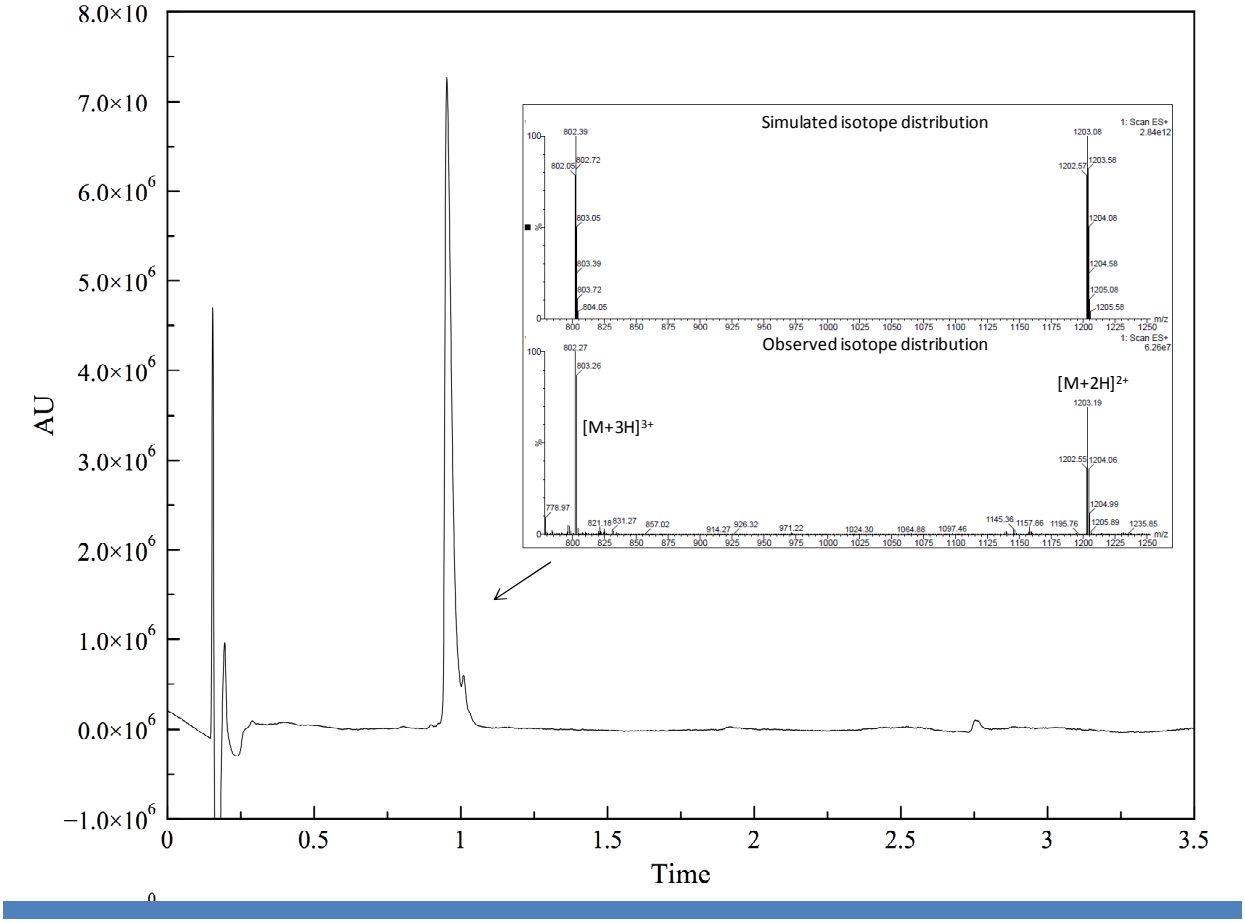


Figure 83 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H2(StBu)$_2$-HER2** and relative ESI-MS spectrum.

- *H3(StBu)*

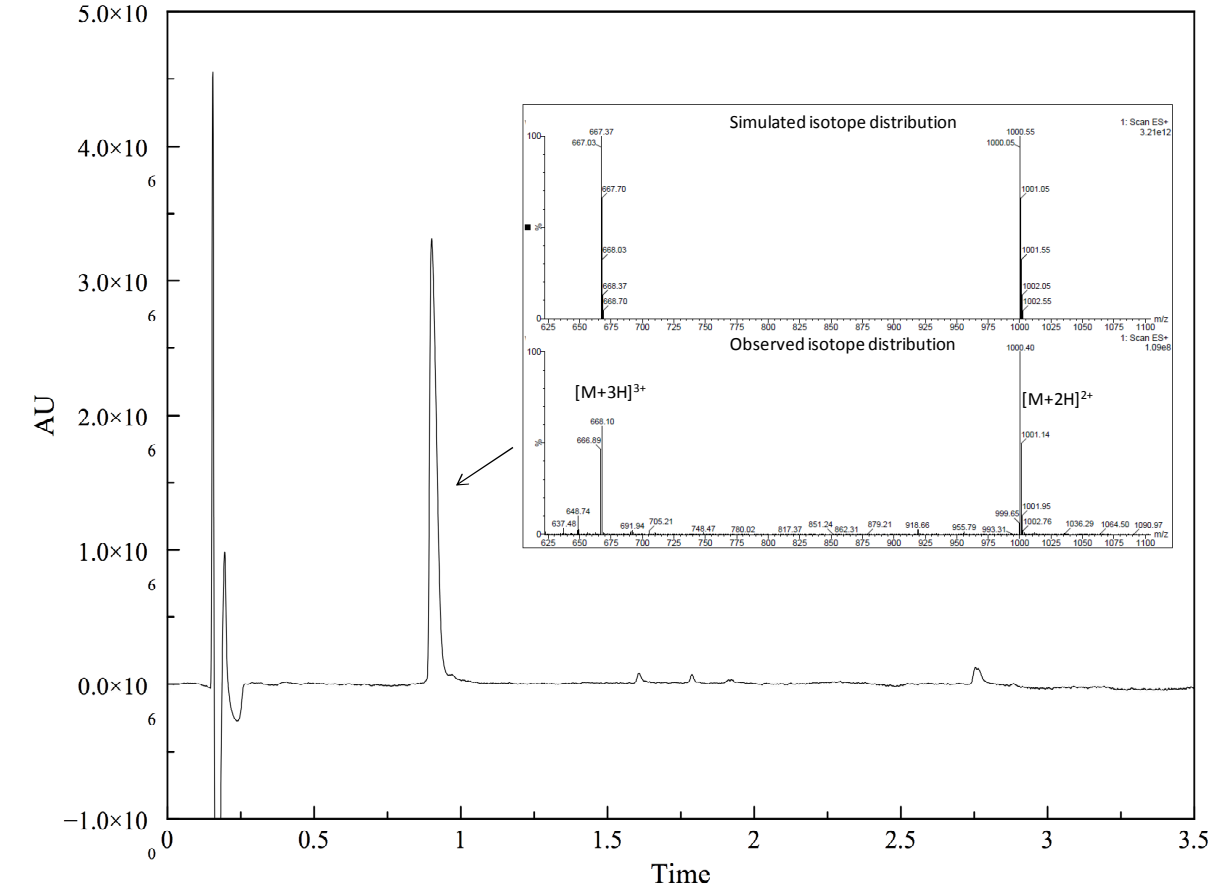


Figure 84 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H3(StBu)** and relative ESI-MS spectrum.

- *H1(StBu)-IgG*

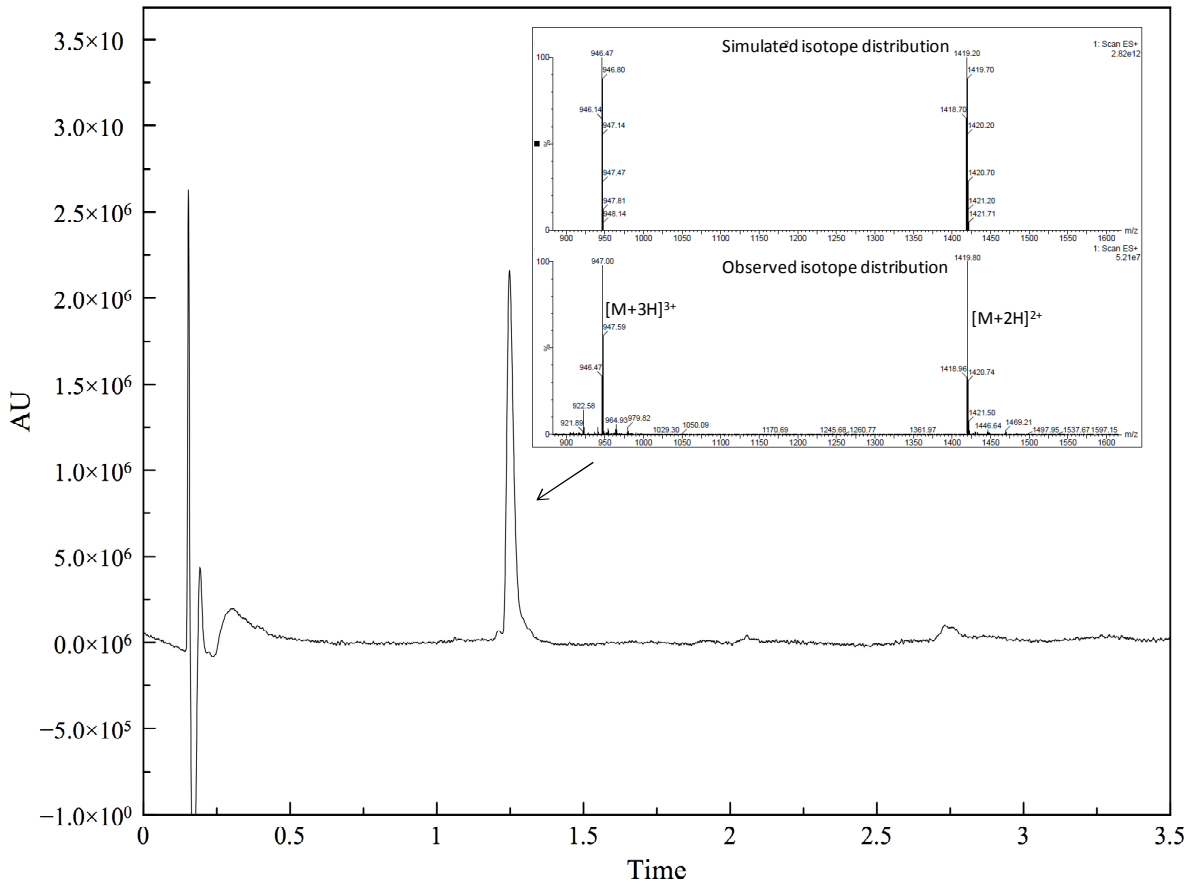


Figure 85 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H1(StBu)-IgG** and relative ESI-MS spectrum.

- *H2(StBu)$_2$-IgG*

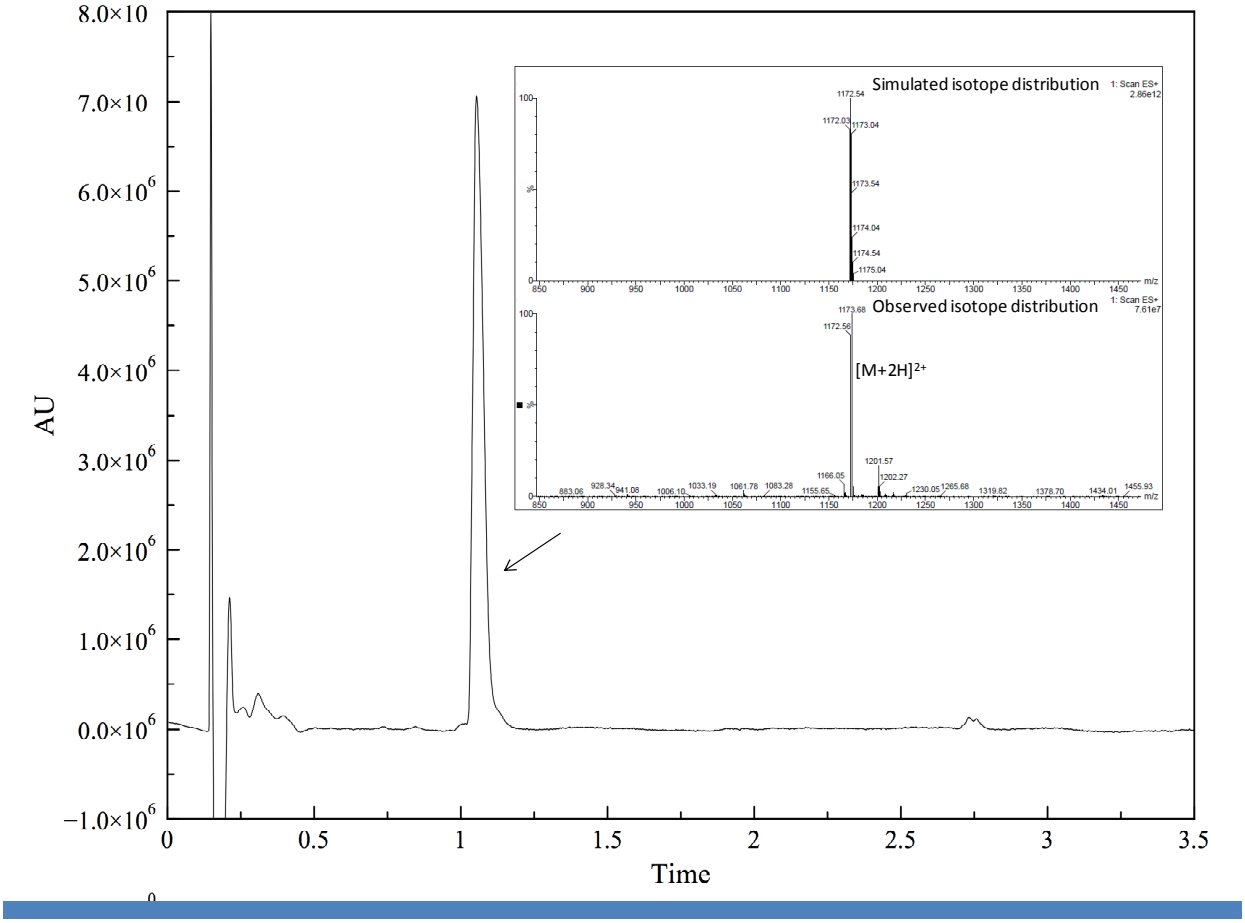


Figure 86 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H2(StBu)$_2$-IgG** and relative ESI-MS spectrum.

- *H2a(StBu)-HER2*

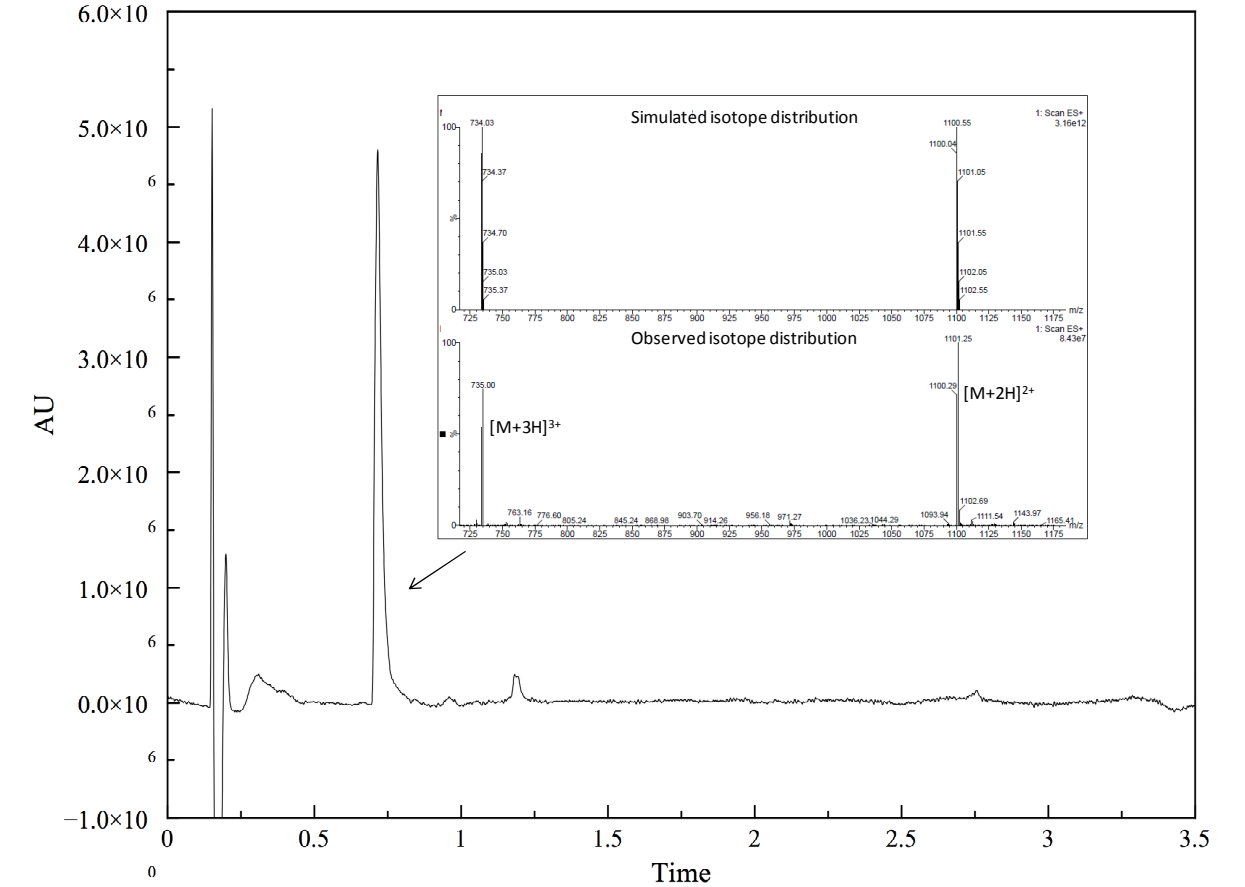


Figure 87 | Reverse phase chromatogram (UPLC-ESI) recorded with UV and MS detection of **H2a-HER2** and relative ESI-MS spectrum.

Annex 7 – Further characterization of the organic compounds from Chapter 5

- *Compound 8*

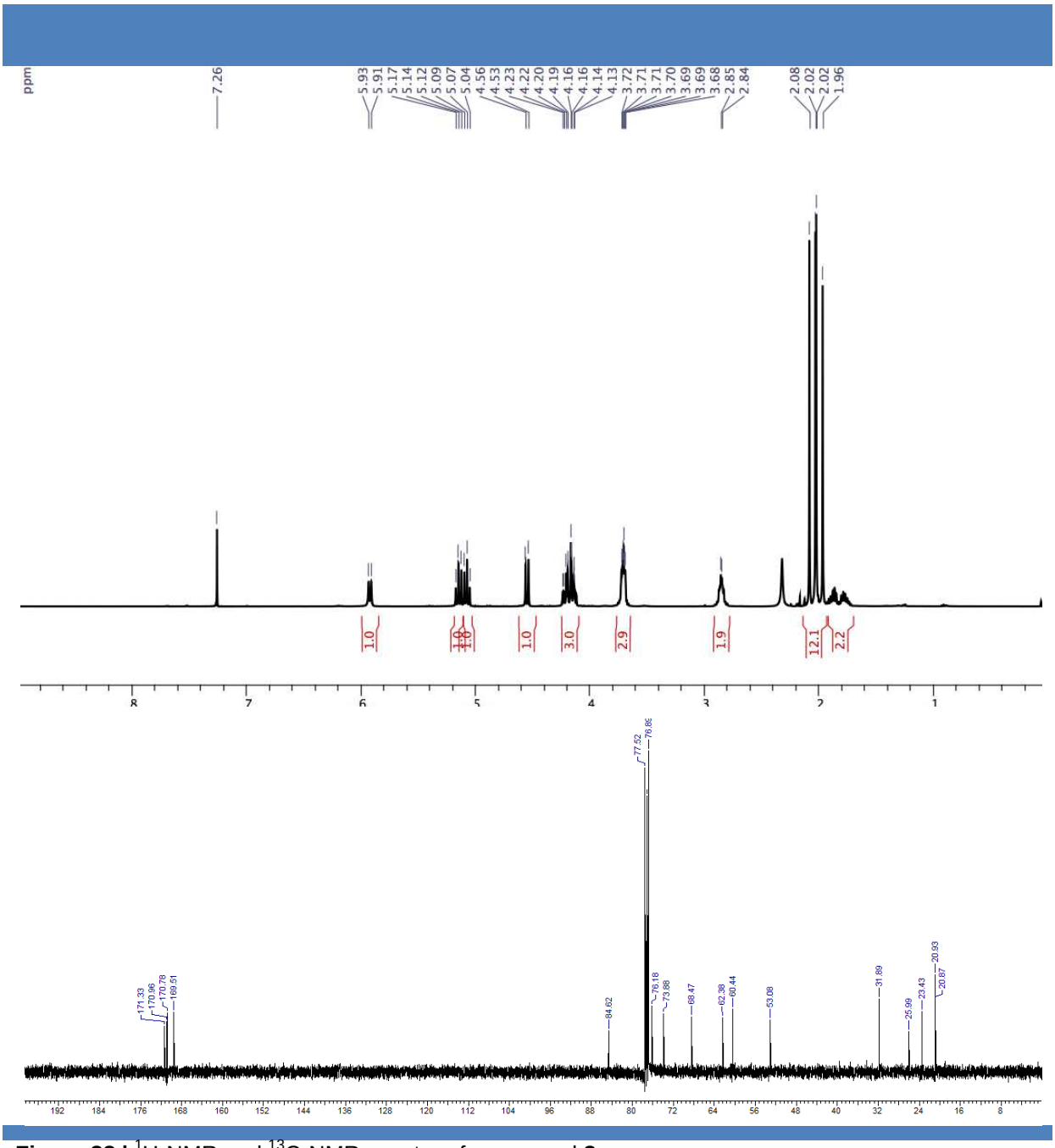


Figure 88 | ^{1}H-NMR and ^{13}C-NMR spectra of compound **8**.

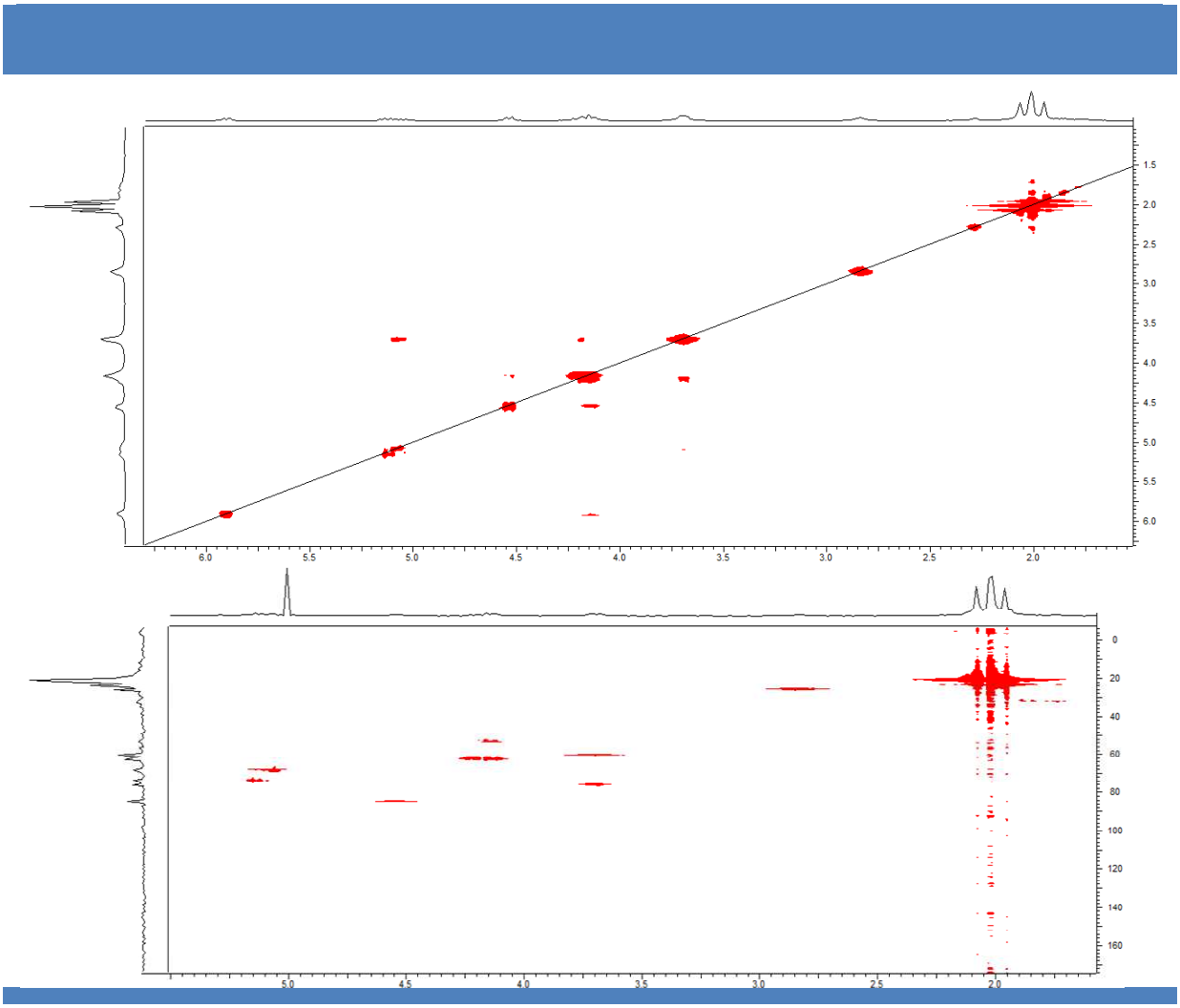


Figure 89 | Zoom in of COSY and HMQC spectra of compound **8**.

- *Compound 9*

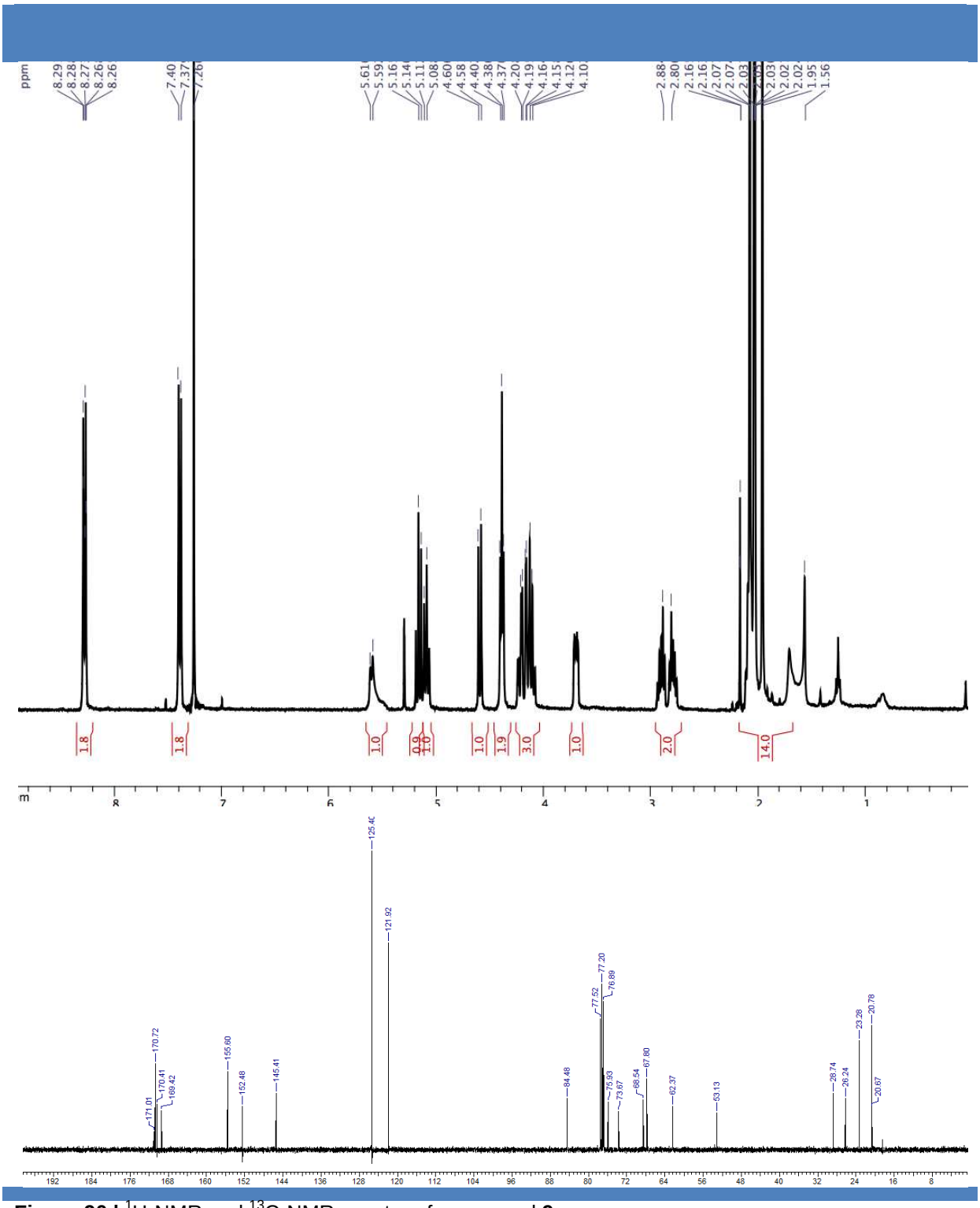


Figure 90 | ^{1}H-NMR and ^{13}C-NMR spectra of compound **9**.

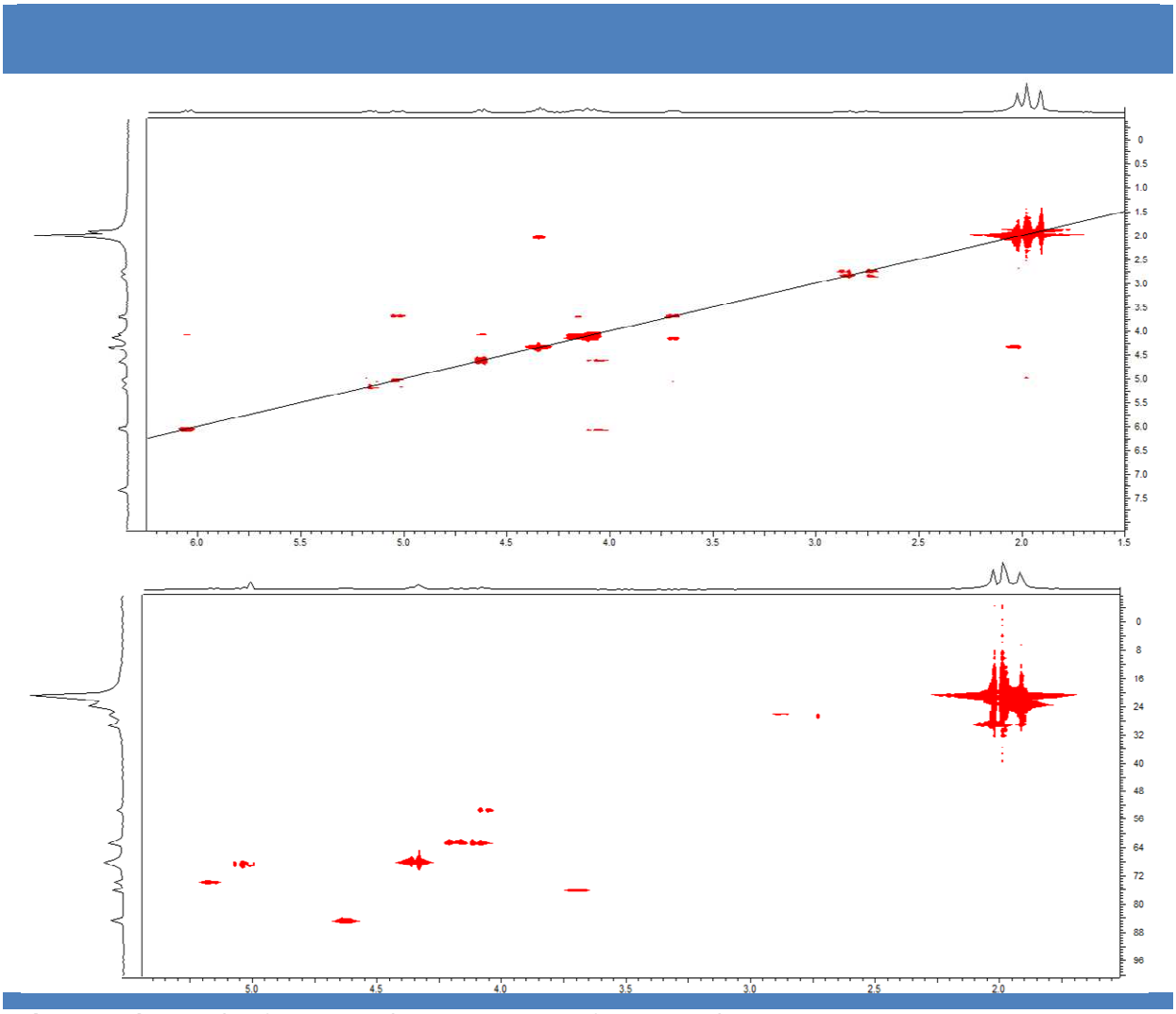


Figure 91 | Zoom in of COSY and HMQC spectra of compound **9**.

Valentina GARAVINI

Native Chemical Ligation for the Design of Dynamic Combinatorial Peptides

Résumé

Utiliser la liaison peptidique dans des systèmes dynamiques covalents est très difficile en raison de sa stabilité intrinsèque. Dans ce travail, une nouvelle méthodologie pour échanger fragments peptidiques dans des conditions biocompatibles est décrite. Légères modifications du groupe amine d'un résidu de cystéine en peptides modèle permettent l'activation spécifique de cette jonction peptidique pour des réactions d'échange covalent. Grâce à un mécanisme de ligation chimique native réversible, fragments peptidiques sont échangés en solution aqueuse à pH physiologique et en présence de dithiothréitol (DTT), avec des demi-temps d'équilibration de 2 à 10 heures. Différentes possibles applications biologiques de cette nouvelle réaction réversible à peptides et glycopeptides sont aussi proposées.

Peptides - Chimie combinatoire dynamique – Liaison réversible - Ligation chimique native

Résumé en anglais

The possibility to use the peptide bond in dynamic covalent systems is very challenging because of its intrinsic stability. In this work, a novel methodology to exchange peptide fragments in bio-compatible conditions is described. The introduction of small modifications to the N-terminus of a cysteine residue in model peptides allows for the specific activation of that peptide bond for exchange reactions. Through a reverse Native Chemical Ligation (NCL) mechanism, peptide fragments were scrambled in aqueous solution at physiological pH and in the presence of dithiothreitol (DTT), with half-times of equilibration in the 2-10 h range. Additionally, possible biological applications of this new reversible reaction to both peptides and glycopeptides are proposed.

Peptides - Dynamic Combinatorial Chemistry (DCC) - Reversible bond - Native Chemical Ligation (NCL)